Alma Mater Studiorum - Università di Bologna

Dipartimento di Fisica e Astronomia
Corso di laurea magistrale in Astrofisica e Cosmologia

On the CTA capabilities in resolving sources in the Galactic Plane Survey

Tesi di laurea magistrale

Presentata da:
Benedetta Bruno

Relatore:
Chiar.mo Prof. Cristian Vignali

Correlatore:
Dott.ssa Roberta Zanin

Sessione II
Anno Accademico 2020-2021

Ai miei genitori

Abstract

The CTAO will be the next generation of Imaging Atmospheric Cherenkov Telescopes array, breaching the limits of the current facilities and covering energies from 20 GeV up to 300 TeV. One of the major scientific objectives of CTAO will be the Galactic Plane Survey which will provide the first complete very-high-energy scan of the entire Galactic plane. Astronomical surveys of the Galaxy provide essential, large-scale datasets that form the foundation for Galactic science at all photon energies. Moreover, with this complete census of the very-high energy gamma-ray source populations, it will be possible to identify promising targets for follow-up observations, such as PeVatrons, the supposedly cosmic-ray factories which are able to accelerate nuclei at least up to Peta-electronvolt energies. However, the improved sensitivity of CTAO with respect to current facilities will lead to a higher probability of source confusion, a problem that concerns multiple overlapping sources and multiple detections related to a single extended source.

This thesis project aims at studying the specific problem of the impact of the source confusion in resolving sources in the Galactic Plane Survey. The full work is based on simulations and subsequent analysis in a systematic way of different kinds of confused regions, starting with the simplest cases in which two pointlike sources have been simulated with the same spectra, then changing spectral and spatial parameters of these two sources. This will allow us to draw clear conclusions about how source confusion will influence future observations and to fulfil the capabilities of CTAO. It is worth noting that up to now, despite the CTAO community is very wide with more than one thousand scientists signed to the CTAO Consortium, a systematic study like this one - on the scientific impact of the source confusion - has never been carried out.

Here, through the computation of the two sources detection rate, it was possible to see when two sources become distinguishable from each other for every scientific case considered, changing the spatial and spectral parameters of the sources. Moreover, it was possible to compare the parameter values obtained from the fitting process with the simulated ones when the two sources were actually resolved and to see how source confusion affected the sources when a single target was detected. In particular, the presence of a very extended source with a “smaller” one results to be a good example of how a diffuse emission influences the detection of a source.

Sommario

Il CTAO sarà la prossima generazione di Imaging Atmospheric Cherenkov Telescopes, superando i limiti degli strumenti attuali e coprendo energie che vanno da 20 Gev fino a 300 Tev. Uno dei principali obiettivi scientifici del futuro CTAO sarà la *Galactic Plane Survey* che fornirà la prima scansione completa ad altissima energia di tutto il piano galattico. Indagini astronomiche di questo tipo forniscono dati essenziali su larga scala che costituiscono la base per produrre risultati scientifici nell'ambito galattico a tutte le energie. Inoltre, con questo censimento completo delle popolazioni di sorgenti che producono raggi gamma ad altissima energia, sarà possibile identificare oggetti promettenti per osservazioni successive, come i Pevatrons, sorgenti presumibilmente all'origine dei raggi cosmici in grado di accelerare i nuclei almeno fino a energie dell'ordine del Peta-electronvolt. Tuttavia, la maggiore sensibilità del CTAO rispetto alle strutture attuali porterà ad una maggiore probabilità di *source confusion* delle sorgenti, un problema che riguarda casi in cui si hanno molteplici sorgenti sovrapposte e rilevazioni multiple relative ad una singola sorgente estesa.

Nello specifico, questa tesi mira a studiare il problema dell'impatto della *source confusion* nella risoluzione delle sorgenti all'interno della *survey* del piano Galattico. L'intero lavoro è svolto simulando e analizzando diversi tipi di regioni di cielo confuse, partendo dal caso più semplice in cui due sorgenti pointlike hanno lo stesso spettro e cambiando poi i parametri spaziali e spettrali. Questo lavoro svolto in maniera sistematica permette di trarre conclusioni chiare e di completare la caratterizzazione delle *performances* del CTAO. Vale la pena notare che finora, nonostante la comunità del CTAO sia molto ampia con oltre mille scienziati iscritti al consorzio, non è mai stato condotto uno studio sistematico sull'impatto scientifico della *source confusion*.

Attraverso il calcolo del tasso di rilevamento delle due sorgenti, è stato possibile capire quando due sorgenti diventano distinguibili una dall'altra per ogni caso scientifico considerato. Inoltre, è stato possible confrontare i valori dei parametri provenienti dai fit con quelli simulati quando le due sorgenti sono risultate distinte e vedere come la *source confusion* ha influenzato una sorgente quando questa è stata l'unica rilevata. In particolare, la presenza di una sorgente molto estesa con una più piccola è risultato l'esempio perfetto di come un'emissione diffusa influenzi il rilevamento di una sorgente.

Contents

Chapter 1

Introduction to very high-energy astronomy

Gamma-ray astronomy is the astronomical observation of gamma rays, the most energetic form of electromagnetic radiation which can come from extreme astronomical objects, such as neutron stars and pulsars, supernova explosions and regions around black holes. Gamma rays are also observable as a diffuse background radiation found along the plane of the Milky Way.
Almost standardly, we use the definition "high energy" (HE) or "GeV" astronomy for the energy range from $\sim 30\,\text{MeV}$ to $\sim 300\,\text{GeV}$, typically covered by space-based experiments. We will refer to "very high energy" (VHE), or "TeV" astronomy, for the range from $\sim 300\,\text{GeV}$ to $100\,\text{TeV}$ and beyond, typically covered instead by ground-based experiments.

Since it is not possible for objects to get hot enough to produce gamma rays as a thermal emission, these must be produced by a non-thermal mechanism. The mechanisms often rely on the presence of HE sub-atomic particles that are produced by some kind of cosmic particle accelerator. This happens in environments where a small number of particles can be accelerated up to energies very close to the speed of light and carry a significant fraction of the energy available. Since energy is no longer shared roughly equally among particles, as in the case of a "normal" hot environment, these processes are referred to as non-thermal processes (see below, Section 1.1).
These special environments are usually associated with violent events such as explosions, outbursts or powerful jets of material produced close to the giant black holes at the centre of galaxies (see Section 1.2 concerning the classes of gamma-ray emitters). For this

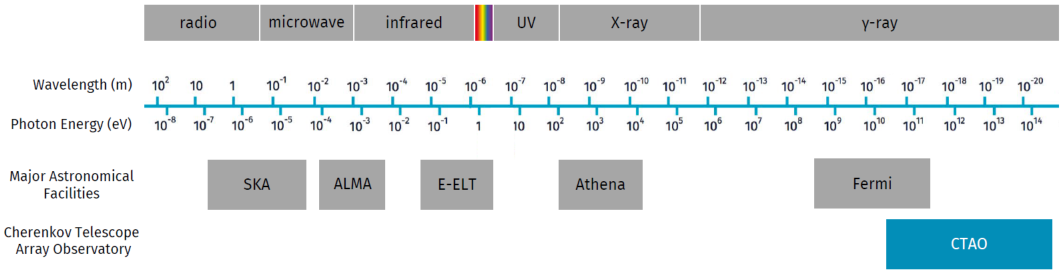


Figure 1.1: The electromagnetic spectrum and how CTAO's energy range will compare to some of the existing and future astronomical instruments. Note that *Fermi*-LAT and CTAO will overlap in the GeV domain. **Credit:** CTA website [38].

reason, gamma rays can be used to trace violent events in the Universe.

Because Earth's atmosphere blocks most gamma rays, observations are generally conducted by high-altitude balloons or spacecrafts. Space experiments cover a very broad gamma-ray energy interval, extending from a few MeV to hundreds of GeV. Significant progress in the understanding of the high-energy sky have been carried out thanks to the *Fermi*-LAT satellite, launched in June 2008, which produced several generations of high-energy gamma-ray source catalogs; the latest one is the *Fermi*-LAT Fourth Source Catalog, 4FGL (for *Fermi* Gamma-ray LAT) [4]. Its energy range extends from 20 MeV to more than 300 GeV: it has an on-axis effective collecting area of $\sim 0.8\,\mathrm{m}^2$ at 1 GeV and a good angular resolution (0.8° at 1 GeV, varying with energy approximately as $E^{-0.8}$ and asymptoting at $\sim 0.1°$ above 20 GeV) for source localization and multiwavelength study [4].

However, beyond a few hundreds GeV, the gamma-ray fluxes are so small that the effective detection area of space-based experiments cannot provide adequate statistics (e.g. the strongest TeV gamma-ray source, the Crab nebula, has an integral flux above 1 TeV that corresponds to $F_\gamma = (2.1 \pm 0.1) \times 10^{-11} \frac{\mathrm{photons}}{\mathrm{cm}^2\,\mathrm{s}}$. Therefore, astrophysical studies of TeV gamma rays only rely on ground-based experiments. A detailed explanation of detection techniques and currently operating facilities will be provided in Chapter 2. The remarkable achievements of these experiments now include the study of morphology, energy spectrum and time variability of all the Galactic and extragalactic source populations known at TeV energies.

The field of gamma-ray astronomy has expanded rapidly during the last decades and with more than 200 known Galactic and extragalactic TeV sources, VHE gamma-ray astronomy has established itself as a branch of astroparticle physics.

Various techniques have been developed to detect VHE gamma rays which allow astronomers to study them indirectly, through the electromagnetic cascade they produce when interacting with the Earth's atmosphere. However, the potential of the imaging atmospheric Cherenkov technique (see Section 2.2) has not yet been fully exploited and significant improvements in sensitivity and resolution over existing instruments can be achieved with an array of Cherenkov telescopes that is both larger (to provide more views of the shower) and more capable (e.g. with wider field of view and higher resolution cameras). Thus, there is both strong scientific and technical motivation for a project like the Cherenkov Telescope Array Observatory (CTAO), the next generation observatory that will further develop the imaging atmospheric Cherenkov technique, breaching the limits of the current facilities (see Section 2.4).

1.1 Production and absorption of gamma rays

TeV gamma rays are copiously produced in environments where effective acceleration of particles (electrons, protons, and nuclei) is accompanied by their intensive interactions with the surrounding gas and radiation fields.

In most cases, electromagnetic radiation processes involving relativistic electrons, in the so-called leptonic model, could explain the photon flux up to the highest energies ($10^{14} - 10^{15}$ eV). An astrophysical accelerator provides a continuous distribution of electron energies where these charged particles are accelerated along a curved path dictated by the magnetic field. Synchrotron radiation is then generated and can be detected with different experimental techniques, from radio-telescopes through X-ray detectors. An important aspect of these observations is that they provide unambiguous evidence for the presence of relativistic electrons in the source regions.

Another important process is the one in which low-energy photons gain energy at the expense of the kinetic energy of ultra-relativistic electrons through collisions is the inverse Compton (IC) scattering. The IC mechanism has the effect of increasing the photon energy and is important in regions where accelerated electrons coexist with a high-energy density of soft photons.

Therefore, synchrotron photons constitute a target for their own parent electron population. The energy distribution of gamma rays produced through the IC mechanism starting from an infrared/X-ray photon population can peak at GeV-TeV energies. The

mechanism described here is the Self-Synchrotron Compton (SSC) in which the processes of synchrotron radiation and IC are strongly correlated.
The *one-zone model* assumes that non-thermal radiations are produced in a single, homogeneous and spherical region in the jet. The emission region moves relativistically toward us and consequently the intrinsic radiation is strongly amplified due to the Doppler boosting. Three parameters are needed to characterise the emission region: the comoving magnetic field, the Doppler factor and the comoving radius of the emission region. With one-zone SSC models it is possible to describe practically all high energy gamma-ray emissions observed from Galactic and extra-galactic sources [27].

In denser regions of the interstellar medium, high-energy protons and nuclei interact with matter through nuclear interactions, often producing neutral mesons, which decay into gamma rays. These processes involving nucleons and mesons are said to be "hadronic".
A direct signature of the presence of accelerated protons in astrophysical environments is provided by the presence of neutrinos and gamma rays. They are mainly generated in the decay of charged and neutral pions, respectively; these mesons are produced in proton-proton collisions:

$$p + p \to \pi^{\pm}, \pi^0, K^{\pm}, K^0, \, p, \, n \, ... \tag{1.1}$$

where "..." represents the presence of higher mass mesons and baryons. Because of its similarity with the process of production of secondary hadrons in a fixed-target accelerator experiment, process (1.1) is usually referred to as the *astrophysical beam dump* mechanism. The cross-section for (1.1) corresponds to about 40–50 mb (1 mb $= 10^{-27}$ cm^2).
A second process that produces secondary mesons is due to high-energy protons interacting with low-energy photons (*ambient photons* γ_ϵ) in the surroundings of sources. This is the so called photoproduction that occurs through the Δ^+ resonance:

$$\begin{aligned} p + \gamma_\epsilon \to \Delta^+ &\to \pi^0 + p \\ &\to \pi^+ + n \end{aligned} \tag{1.2}$$

and cross-section at the resonance of ~ 0.250 mb for each individual channel.
Neutral pions, whose mass m_0 is 135 MeV/c^2, then decay in $\sim 10^{-16}$ s into gamma rays, via the process:

$$\pi^0 \to \gamma\gamma \tag{1.3}$$

The kinematics of the decay shows that, whatever the spectrum of the incident particle, the gamma-ray energy spectrum reaches a maximum at $m_0c^2/2 \approx 68\,\mathrm{MeV}$, just above the threshold of the production reaction, then decreases smoothly, according to a power law whose spectral index is close to that of the incident particle. For many decaying π^0 's, the distribution of the number of photons is constant over the same energy range that can go up to hundreads of TeV.
Although this cross-section $\sigma_{\gamma p}$ is two orders of magnitude smaller than the cross-section σ_{pp} of the beam-dump process, in some astrophysical environments the probability that secondary mesons are produced by reaction (1.2) is much higher than the probability due to (1.1). This is due to the number density of ambient photons, n_γ, that could be much larger than that of environmental matter number density, n, and the event rate for the latter process is $\sigma_{pp} \cdot n \cdot c$, while for the former it is $\sigma_{\gamma p} \cdot n_\gamma \cdot c$.
When a E^{-2} energy spectrum of accelerated protons at the source is considered, the π^0 decay spectrum for an observer within the laboratory frame in the $E^2 \frac{\mathrm{d}N_\gamma}{\mathrm{d}E_\gamma}$ representation rises steeply below $\sim 200\,\mathrm{MeV}$ and approximately traces the energy distribution of parent protons at energies higher than a few GeV. Gamma rays from π^0 decay are also produced by the cosmic ray (CR) diffusion in our Galaxy. The structure resulting from this mechanism is often referred to as the pion-decay bump and uniquely identifies the presence of gamma rays originated by π^0 decays. The discovery of the bump originated by the π^0 decay in the source would identify the presence of HE protons.

An example of a spectral energy distribution (SED or photon energy spectrum) including all of the mechanisms described above is shown in Figure 1.2. It is clear that the SED can be approximated by power laws only in a limited energy range. A photon energy spectrum proportional to E^{γ_p} is referred to as "soft" if γ_p (photon index) is larger than 2 (which corresponds to the decreasing part of the SED), and "hard" in the opposite case.

Discrimination between different models is very challenging. In a hadronic interaction, the secondary photons have an energy by a factor 10 lower than the primary proton. Therefore, the quest for CR sources which are able to accelerate particles in the so-called "knee" energy region requires to survey the gamma-ray sky above 100 TeV. In addition, the IC scattering at these energies is strongly suppressed by the Klein-Nishina effect. Furthermore, where the magnetic field reaches values of $\sim 100\,\mu\mathrm{G}$, leptons are limited by losses and therefore their energies do to exceed a few dozen of TeV.

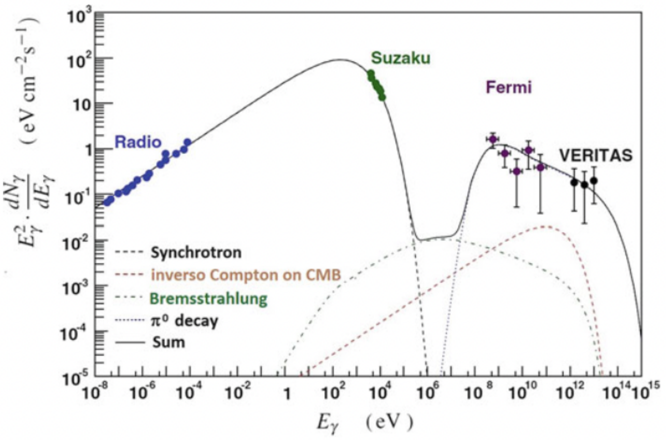


Figure 1.2: Broadband SED data (in radio, X- and γ- rays) and corresponding emission model of Tycho's SNR. Note the shapes of the curves from synchrotron, inverse Compton and π^0-decay: the hadronic contribution to gamma rays seems to be the dominant one in this example. **Credit**: *Fermi*-LAT Collaboration.

For what concerns extragalactic gamma-ray astronomy, the Universe is opaque to gamma rays whenever the energy-dependent photon mean free path is smaller than the distance of the source. The dominant process for the absorption of VHE photons of energy E produced by astrophysical sources is pair-creation:

$$\gamma_E + \gamma_\epsilon \rightarrow e^+ e^- \tag{1.4}$$

on low-energy extragalactic background photons of energy ϵ. These photons extend from the cosmic microwave background (CMB) to the near-ultraviolet (UV) wavelengths. The UV/optical/infrared background radiation is called the *extragalactic background light* (EBL). The cross-section for the pair-production process is described by the Bethe-Heitler formula

$$\sigma(E,\epsilon) \simeq 1.25 \times 10^{-25}(1-\zeta^2)\,\mathrm{cm}^2 \tag{1.5}$$

where $\zeta = \sqrt{1 - \frac{(m_e c^2)^2}{E\epsilon}}$. The cross section in (1.5) is maximized when

$$\epsilon = 2\frac{(m_e c^2)^2}{E} \simeq \left(\frac{520\,\mathrm{GeV}}{E[\mathrm{MeV}]}\right)\mathrm{eV} \tag{1.6}$$

and the threshold energy is a factor of two lower. For energies above 100 TeV, the process (1.6) is dominant for the energies corresponding to photons of the cosmic microwave background [36].

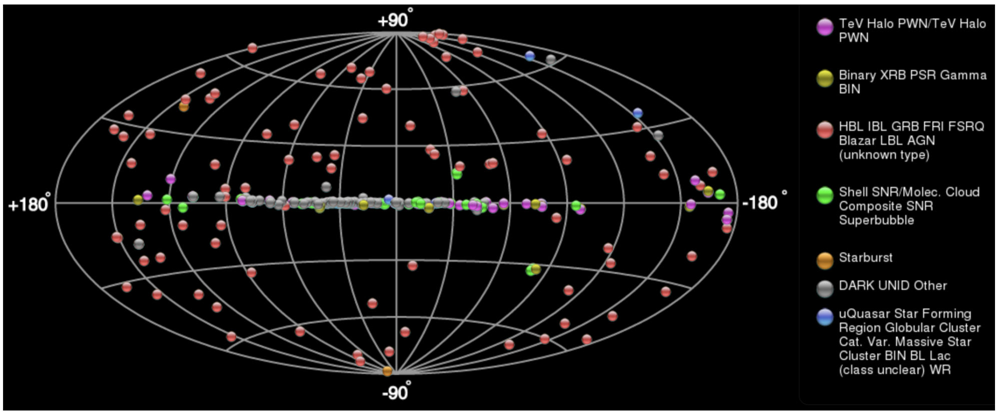


Figure 1.3: Map of the 243 TeV sources (as July 2021) retrieved from the online TeV catalog [23], which displays, with different color codes, the position in Galactic coordinates of the various gamma-ray sources detected from the ground. **Credit**: online TeV catalog.

This absorption effect has been observed on the energy spectra of TeV-emitting active galactic nuclei (AGN). Its estimation allows the density of background photons to be measured. This is an important result, since direct measurements of these cosmic radiation fields are difficult to obtain due to contamination from nearby objects and Galactic light.

1.2 Classes of gamma-ray emitters

All the VHE gamma-ray sources and the associated publications are registered in an online catalogue, TeVCat [23], from which sky maps and characteristic tables can be easily extracted. This database shows that 247 (at the time of writing) highly significant ($\geq 5\sigma$) sources have now been published in referred journals (Figure 1.3). This number has grown over time, from the first discovery in 1989 until now, with a slow evolution until 1996 (only three sources by that time), a fast one during 1996–2004 due to new telescopes coming available, and an even faster rise from 2005 onwards when facilities like H.E.S.S. started operations, as well as MAGIC and VERITAS (these instruments will be introduced in Section 2.3), bringing the source counts to its present level.

Therefore, an important motivation for high-energy gamma-ray astronomy is the study of the phenomenology of these sources. Since non-thermal spectra are a common feature of the observed gamma-sources, detection of TeV sources can provide constraints

on the models of acceleration and radiation processes at extreme conditions.
New emitting sources have been discovered as they do not have an observational counterpart of objects already known at different wavelengths. The fraction of unassociated TeV sources among those firmly identified in a Galactic plane survey is ~25% [3]. These sources are not necessarily dark, i.e. emitting exclusively in the VHE domain; it is also possible that their counterparts were missed by the multiwavelength (MWL) association procedure, for which the angular resolution limit of VHE detections can be relevant, from ~ 0.07° to ~ 0.5°, depending on the detection technique (see Chapter 2). In short, most sources have a firm associations, a plausible or a potential counterpart in other wavelength regimes. Whether there remains a population of truly dark VHE sources in the Galactic plane surveys can only be figured out with deeper MWL studies.

Prior to 2000, the general expectation for the field was to detect the accelerators of the HE CRs, most likely supernova remnants (SNRs) and closeby AGN (e.g. Markarian galaxies [25]). However, the reality over the last two decades has been the discovery of an astonishing variety of VHE gamma-ray sources, with more than half located in the Galactic plane.
The astrophysical gamma-ray sky is usually decomposed into individually-detected sources and diffuse gamma-ray emission. The former, including point sources and extended sources, contains many different source types: Galactic sources like SNRs, pulsar wind nebulae (PWNe), objects showing evidence of both shell and nebular emission (composite objects), binaries, massive stellar clusters, the supermassive black hole Sgr A*, etc., and extragalactic sources like AGN.
The largest class of Galactic TeV emitting sources corresponds to that of pulsars wind nebula, among which the Crab Nebula is the most studied representative.

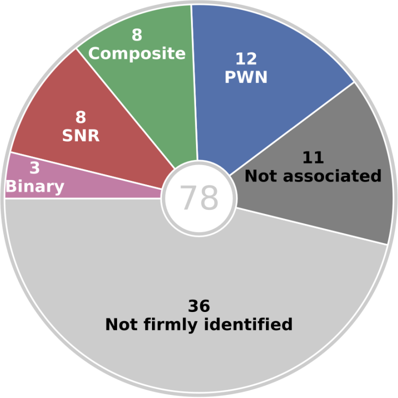


Figure 1.4: Source identification summary pie chart of the HESS Galactic Plane Survey (HGPS) **Credit**: HESS Collaboration [3].

Beyond the Galaxy, almost all known TeV gamma-ray sources are AGN (BL Lac objects, flat-spectrum radio quasars, radio galaxies) and their emissions are thought to originate from one or multiple regions of particle acceleration in the jets. Others extragalactic sources are starburst galaxies and gamma-ray bursts (GRBs) as well as sources like the Galactic ones (SNRs, PWNe) identified in the Milky Way satellites.

Concerning the diffuse Galactic gamma-ray emission, it is mainly attributed to the interactions of energetic CR electrons and nuclei with interstellar nucleons and photon fields in the Galactic plane, providing key insight into the properties of propagation of CRs in the Galaxy. The decay of secondary particles produced in hadron collisions, the IC of the interstellar radiation field by electrons and their bremsstrahlung emission in the interstellar gas are thought to be the main contributors to this diffuse Galactic emission [8]. This results as a power-law spectrum with a photon index $\Gamma \sim 2.7$ ($S \propto \nu^{-\Gamma}$) that is particularly relevant in the GeV-range energies.

1.3 Some open questions

The remarkable achievements of observational gamma-ray astronomy over the last decade, and the recent theoretical and phenomenological studies of acceleration and radiation processes in astrophysical environments, fully justify the further exploration of the sky at high and very-high energies. Recent results have been exciting and have established the field, but there are indications that they represent the tip of the iceberg and that many more sources remain to be discovered. Increasing the number of sources populations will be of foundamental importance for future studies in order to reach an adequate characterisation.

One of the main results from the next-generation detectors will probably be the discovery of new classes of CR sources. The key is related to dedicated surveys, which constitute an unbiased, systematic exploratory approach. Surveys at different extents and depths are amongst the scientific goals of all major planned facilities. Identifying cosmic accelerators in which particles can reach at least the PeV energy range (or “Pevatrons”, 1 PeV = 10^{15} eV) is an important challenge today. About 2% of the cosmic rays are electrons and positrons with a much steeper energy spectrum compared to that of protons or nuclei [32]. Unfortunately, all these particles are charged, and thus continuously deviated by turbulent magnetic fields embedded in interstellar and intergalactic plasmas, so that their direction toward Earth is almost completely uncorrelated with the position of their originating source, except perhaps at extreme energies where fluxes are extremely low ($\sim 10^{-14}\,\mathrm{m}^{-2}\mathrm{sr}^{-1}\mathrm{s}^{-1}$ for energies of 10^{19} eV). Thus, the quest for cosmic accelerators as well as for acceleration mechanisms mostly relies on photons or neutrinos that propagate along a straight line. This explain why the open problem of the identification of Galactic CR sources is connected to the identification of gamma rays originating from the hadronic mechanism. A straightforward test of the acceleration for CRs in SNRs up to PeV energies would be the detection of gamma rays produced through the hadronic mechanism directly from young remnants and/or from dense clouds overtaken by the expanding shells. The main challenge is to distinguish gamma rays emitted through hadronic processes (π^0-decay) from those originating in leptonic processes via SED modelling.

Moreover, there are still many open issues regarding the field of VHE. At around the energies that can be still detectable from space instruments, gamma rays have been observed coming from jets of many black holes, although the exact mechanisms through this emission process occurs are not fully understood. Information from energies above TeV will be crucial to encode details about the physical processes at work in some of the most energetic environments in the Universe.

Regarding observational cosmology, a large amount of our knowledge is based on the diffuse background radiation fields that surround us. A prominent example of such a radiation field is the 2.7 K thermal afterglow of the Big Bang, the CMB. At shorter wavelengths, in the range ultraviolet (UV)−far-infrared, the extragalactic radiation field consists of the accumulated and reprocessed radiation of all starlight produced thus far. This extragalactic background light encodes the integrated history of structure formation and the evolution of stars in the Universe. Understanding these characteristics requires detailed theoretical modeling of all the processes that contribute together with source population studies, and VHE observations of blazars can be used to indirectly constrain these models.

Finally, some of the greatest challenges in the area of fundamental physics are related to gamma-ray phenomena that we look forward to explore. Associated with modeling the formation history of the Universe is the general lack of knowledge regarding the nature of Dark Matter (DM) that can be inferred through observations of VHE gamma rays coming from the annihilation/decay of DM particles. Also, the long travel times of gamma rays from extragalactic sources combined with their short wavelength make them a sensitive probe for energy-dependent variation of the speed of light due to quantum-gravity induced fluctuations of the metric. Then, there is the exciting possibility of finding evidence for Lorentz invariance violation (LIV) and detecting axion-like particles (ALPs) associated with possible quantum gravity effects on space-time at the Planck scale [21].

1.4 Aims of this thesis

The CTAO will transform our understanding of the VHE universe and will address important questions in fundamental physics, as mentioned above. It will also be an important member of the suite of experiments and observatories participating in the expanding areas of multi-wavelength and multi-messenger astronomy.
A significant part about CTAO core programme will comprise surveys of the Galactic plane that are key legacy programmes for all major observatories. In this context, the determination of efficient criteria to identify PeVatron candidates during the survey is essential in order to trigger further dedicated observations.

This thesis project aims at studying the specific problem of the impact of the source confusion, e.g. multiple overlapping sources detected as single source, in resolving sources in the Galactic Plane Survey (GPS, see Chapter 3). It is worth noting that up to now, despite the CTAO community is very wide with more than 1000 scientists signed to the CTAO Consortium, a systematic study on the scientific impact of the source confusion has never been carried out. However, it is of fundamental importance to deal with this problem in order to have the right knowledge and tools to identify sources and describe their morphological and spectral properties. This work will indeed fulfil the characterization of the capabilities of CTAO for the GPS.

The full work is based on simulations with the Science Analysis Tools *Gammapy* adopted by the CTAO to both simulate the gamma-ray sky as expected to be seen and subsequently analyse the simulated data. Simulations cover a wide parameters space, so as to deal with all significant scientific cases and analyse different kinds of confused regions, in a systematic way that allows us to draw clear conclusions on the impact of source confusion in resolving sources in the Galactic Plane.

Chapter 2

Detection techniques and facilities for VHE photons

The atmosphere of the Earth is not transparent to gamma rays, therefore they can be directly detected only by platforms in space. However, beyond 10 GeV the current space facilities cannot provide adequate statistics for comprehensive spectral and temporal studies in the VHE domain.
Due to the constraints on the maximum size and weight of an instrument that can be delivered into space, the effective detection area of any satellite experiment (using current launch vehicles) is limited to the order of a few square meters. Furthermore at higher energies, the photon flux decreases linearly with increasing energy. Thus, to detect the same number of events, the detector aperture $A \cdot T$ (where A is the geometric area and T the exposure time), must increase by a factor of 100. Even with a mission life time of 10 years, this results in a practical limit for space-based observations of astrophysical sources at energies larger than a few hundreds GeV. Instruments with an effective area smaller than $\sim$100 m^2 cannot detect the expected astrophysical flux in the TeV energy region. Even with the advent of new, larger launch vehicles, the prospects for a space telescope with an area significantly larger than $\sim$10 m^2 are highly uncertain.

Fortunately, at these energies alternative methods can be exploited using two different facilities: air shower array, based on the registration of particles coming from atmospheric showers, or Imaging Atmospheric Cherenkov Telescopes (IACTs), which detect the Cherenkov radiation coming from the same showers as particles propagate through the atmosphere. The first has larger collection area due to a combination of higher duty cycle and array footprint; furthermore, the array has a wide field of view (FoV) and a continuous operation capability. IACTs, instead, provide a better angular and energy resolution and a lower achievable energy threshold. Hence, the two methods for ground-based VHE astronomy are complementary.

2.1 Physics of Extensive Air Shower (EAS)

An air shower is an extensive cascade of particles and electromagnetic radiation produced by the interaction of a cosmic ray (called primary) with a nucleus in the atmosphere. In addition to the hadronic component, the decays of short-lived hadrons lead to a shower of particles: photons, electrons, and positrons constitute the electromagnetic (EM) component; muons and neutrinos constitute the penetrating component. All these particles travel at the speed of light in the atmosphere, approximately along the direction of the primary CR.

For photons (and also for electrons and positrons) above a few hundred MeV, the cascade process is dominated by the pair production and the bremsstrahlung mechanisms: an energetic photon scatters on an atmospheric nucleus and produces an e^+e^- pair, which emits secondary photons via bremsstrahlung; such photons produce in turn a pair, and so on, giving rise to a shower of charged particles and photons, degrading the energy down to the critical energy E_c where the ionization energy loss of charged particles starts dominating over bremsstrahlung. In air, the critical energy E_c is about 80 MeV [24].

Gamma rays directly coming from astrophysical sources are less numerous than charged CRs of the same energy, but they can also induce an almost pure EM cascade which reaches higher value of shower maximum depth, i.e. the depth in the atmosphere at which a shower produces the maximum number of particles, and a much lower quantity of muons.

In order to separate the gamma-ray signal from the background due to CRs, a good angular resolution is required for the experiments, and possibly there should be a way to reconstruct their topology, as the air shower originated by a CR is much more widely distributed on the ground compared to that of a gamma ray that results instead slimmer, more concentrated and orientated towards the source (see Figure 2.1 and 2.2). Otherwise, the experiments should be able to detect the presence of muons in hadronic showers, using muon detectors devoted to hadron rejection.

The number of particles in proton-induced events exceeds the number of particles in gamma-induced ones at low altitudes. This implies that, in gamma-ray astronomy,

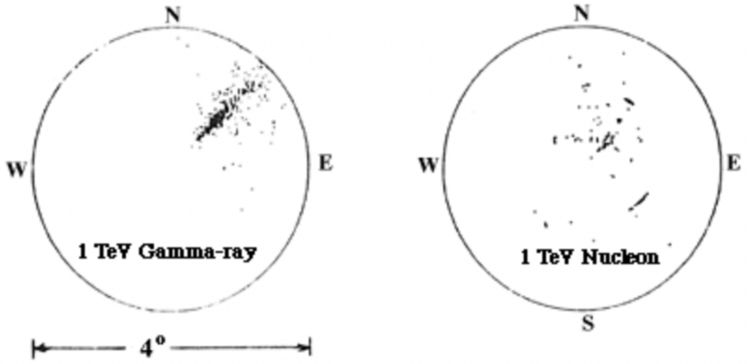


Figure 2.1: The particles in an air shower are much more widely distributed for proton showers (right panel) versus gamma-ray showers (left panel). This is reflected in the corresponding distribution of photons in the detector. **Credit**: NASA's Imagine the Universe.

the trigger probability is higher for the background than for the signal. Also, the small number of charged particles in sub-TeV showers within 150 m from the shower core imposes to locate experiments at extreme altitudes ($>$ 4500 m a.s.l.). At 5500 m a.s.l., 100 GeV gamma-induced showers contain about 8 times more particles than proton-induced showers within 150 m from the core [27].

2.1.1 Cherenkov light

The ultra-high energy particles produced in the shower travel faster than the speed of light in the atmospheric medium. The resulting polarization of local atoms, as the charged particles travel through the atmosphere, causes the emission of a faint, bluish light known as "Cherenkov radiation", similar to the sonic boom created by an aircraft exceeding the speed of sound. The light has its maximum intensity in the wavelength range $300 - 350$ nm, lasts few nanoseconds and is emitted in a coherent cone at an angle with respect to the particle direction given by:

$$\theta_c = \cos^{-1}(\frac{1}{n\beta}) \tag{2.1}$$

from the direction of the particle. Therefore, any secondary charged particle in the shower produces Cherenkov light if its velocity exceeds the threshold $\beta = 1/n$, where n is the refractive index of the medium ($\beta \sim 1$ for ultra-relativistic particles). As the pressure varies with altitude, n decreases with increasing altitude. The angle therefore increases

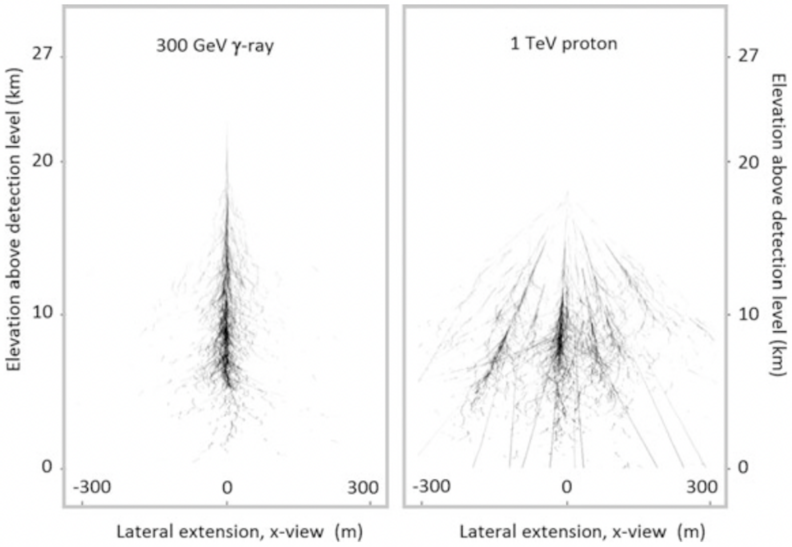


Figure 2.2: Comparison of a pure electromagnetic shower from a 300 GeV gamma ray (left panel) and a shower initiated by a 1 TeV proton (right panel). The plot shows the projection of secondary particle trajectories onto a plane in which the ordinate corresponds to the elevation [36].

when the EM shower propagates from the top of the atmosphere down to ground level, from 0.66° at a height of 10 km to 0.74° at 8 km. This results in a rough focusing of light onto the ground into a ring-like region with radius of $R \sim 10\,\mathrm{km} \times 0.012\,\mathrm{rad} = 120\,\mathrm{m}$ for a typical gamma-ray induced shower.

The number of Cherenkov photons emitted per unit length is $n_c \sim 0.1$ photons cm^{-1} at sea level. Multiplying n_c by the number of particles at maximum (N_{max}) and by the path length of shower particles, the total number of Cherenkov photons turns out to be $N_c \sim 10^6$ for 1 TeV gamma rays [27]. Note that N_c is proportional to E_γ and, hence, it is possible to obtain the energy of the primary gamma ray from the registered signal.

The radiation is too faint and too fast to be detected by the human eye but not too faint for large mirrors and high-speed cameras that can detect the flash of light and image the cascade generated by the gamma rays for further studies of their cosmic sources. The energy of any incident gamma ray is determined from the intensity of the Cherenkov image and its direction of arrival from the image shape and orientation (Figure 2.4).

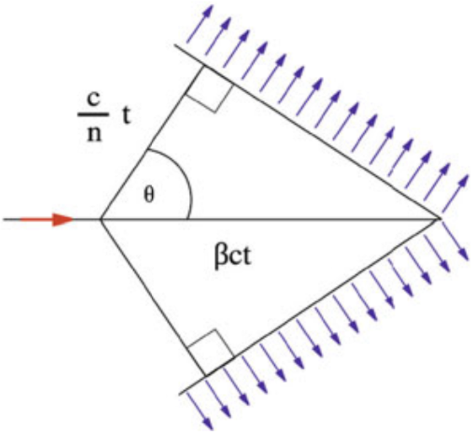


Figure 2.3: Emission of Cherenkov radiation by a charged particle as a cone with aperture $\theta = \cos^{-1}(\frac{1}{n\beta})$. **Credit**: User Harp, Creative Commons Attribution-Share Alike 2.5 Generic license via Wikimedia Commons.

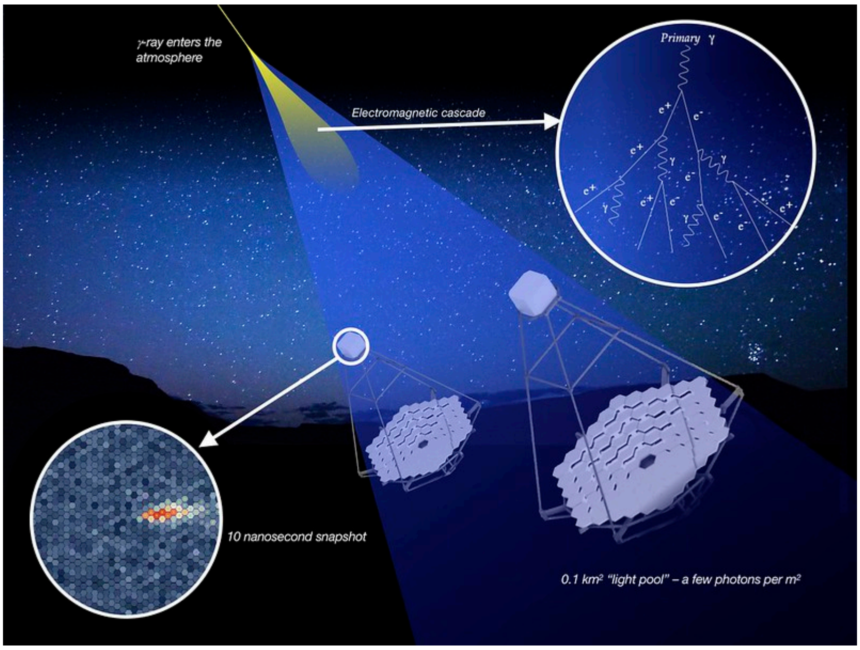


Figure 2.4: Cherenkov effect and how CTA's telescope detect the signal. **Credit**: R. White (MPIK) / K. Bernlohr (MPIK) / DESY.

2.2 Particle sampling arrays and IACTs

The detection tecnique of EAS originally designed for CRs at PeV and EeV energies through the particle sampling arrays can be also implemented for the VHE photons. Experiments presently in operation are large arrays of detectors sensitive to charged secondary particles generated in the atmospheric showers. The energy threshold of such detectors is rather large, e.g. a shower initiated by a 1 TeV photon typically has its maximum at about 8 km a.s.l. [24].
Analyse the particle content of the EAS at the ground can be achieved:

- using a sparse array of scintillator-based detectors or water Cherenkov detectors
- by effectively covering the ground, to ensure efficient collection and hence lower the energy threshold; usually, those experiments are spread over an area of $10^4 - 10^5\,\mathrm{m}^2$ with a spacing between the single detectors of $10 - 20\,\mathrm{m}$.

In order to adopt this technique to gamma-ray astronomy, the energy threshold needs to be reduced by two or three orders of magnitude compared to those EAS designed specifically for CRs. This can be achieved using dense particle arrays located on very high altitudes. Generally, the energy threshold of EAS detectors reaches at most the $0.5 - 1\,\mathrm{TeV}$ range, so they are built to detect HE photons as well as the most energetic VHE gamma rays. At such energies, fluxes are small and large effective areas of the order of $10^4\,\mathrm{m}^2$ are required. It is worth noting that the effective area is the product of the collection area times the detection efficiency and that the collection area can be larger than the area covered by the detector, since one can detect showers partially contained, e.g. when the shower core is observed near the edges of the array.
The sparse sampling sets the energy threshold and determines a poor energy resolution. The direction of the incoming primary particle that induces the shower is reconstructed with the fast timing method making use of the relative times at which the individual detection units are fired by the shower front. Moreover, the energy comes from the integration of the measured density of secondary particles at the detector level, while the angular resolution is limited by the shower fluctuations ($\sim 1^\circ$).
Water Cherenkov detectors can be used in these arrays for particle identification since the photon yield and the emission angle depend on the mass of the particle. Threshold Cherenkov detectors make a yes/no decision based on whether a particle velocity is/ is not above the Cherenkov threshold velocity and, if the momentum has been measured, provides a threshold measurement of the value of the mass. A more advanced version uses the number of detected photoelectrons to discriminate between particle species.

Though, those particle sampling arrays do not have the capability of imaging, in contrast to other instruments that exploit Cherenkov light produced in the air by those ultra-relativistic particles (as it will be explained below).

EAS detectors are capable to detect extended sources and large anisotropy in the gamma and cosmic rays flux at the TeV region of energy. Also, thanks to the duty cycle of 100%, they are the only instruments that can observe transient sources for very long periods.

The imaging Cherenkov technique is the other possible way to study the cascade that gamma rays and CRs produce when interacting with the Earth's atmosphere. This is the basis of how an Imaging Atmospheric Cherenkov Telescope operate.
IACTs detect the Cherenkov light emitted by air shower particles, both coming from the hadronic and the electromagnetic components. The imaging technique relies on the detection on the ground of the images of the Cherenkov light distribution from these cascades. The main goal though is to detect gamma rays coming from the electromagnetic ones. From the measurements, it is possible to determine both the longitudinal and lateral development of the electromagnetic showers, and the arrival direction and energy of the primary gamma rays. In particular, the shower topology and the information on the image shape are used to reconstruct the energy and direction of the primary particle and, since there are more protons hitting the atmosphere than gamma rays, to determine its type.

Imaging Cherenkov telescopes are essentially wide-field optical telescopes consisting of a large reflector of about 10 m radius, reflecting the shower-induced Cherenkov light into a high-speed multi-photomultiplier tube (PMT) camera in the focal plane. Short exposures (less than 30 ns) are required to detect the faint flashes of Cherenkov light against the Poisson fluctuation in the night-sky background.
An IACT array must be operated (as with other telescopes) in almost total darkness, and thus must be installed far from human environments. These telescopes usually operate on moonless nights, thus limiting the duty cycle to 10–15%. High-speed detectors and electronics are required to minimize the integration time, i.e. the amount of time the chip spends in "counting photons." Altogether, the measurements with IACTs are therefore rather challenging and require long observation times.
As with most very large optical telescopes, IACTs typically make use of an altitude-azimuth drive for tracking sources during large exposures.

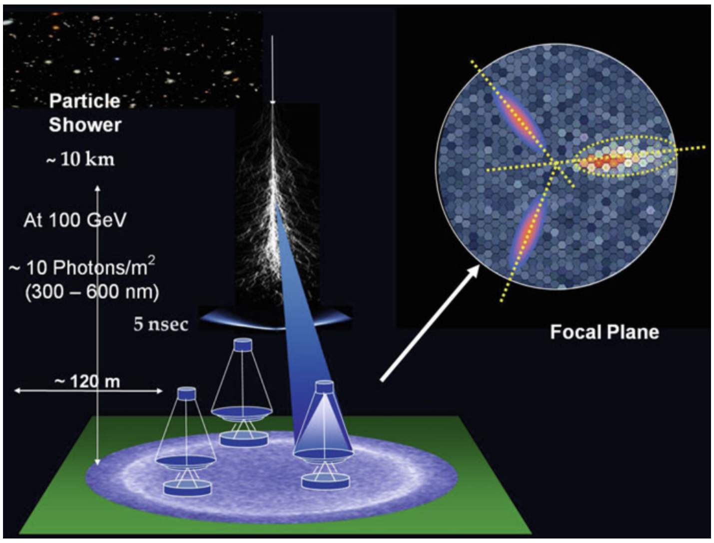


Figure 2.5: Sketch of the imaging atmospheric Cherenkov telescopes with the stereoscopic technique. **Credit**: Modified from an original drawing by Prof. W. Hofmann.

Stereoscopic observations allow background events to be suppressed with higher efficiency (by a factor of about 100) and to reach an angular resolution with a precision of a few arcminutes that allows for the resolution of important substructures of some sources (see Section 2.3 for details about angular resolution of currently running facilities). The successful realization of this technique adopted by IACT experiments has led TeV astronomy to rapidly evolve from an underdeveloped branch of CR studies to a true astronomical discipline. This stereoscopic observation mode allows to intersect the shower axis of two (or more) Cherenkov images of the shower that should have axes pointing toward the arrival direction, the superimposition of which allows the identification of the source (Figure 2.5).

Thanks to the very large collection area, the IACT technique provides large gamma-ray photon statistics even from modest TeV gamma-ray emitters (1-10% of the Crab flux[1]). In combination with good energy and angular resolutions, the gamma-ray photon statistics appears to be adequate for deep morphological, spectroscopic and temporal studies. This also makes the IACT arrays powerful multifunctional and multi-purpose tools for the exploration of a broad range of non-thermal objects and phenomena.

As an example, the stereoscopic observations of air showers with two or more 10 m diameter telescopes located at distances of about 100 m from each other provide a quite

---

[1] The Crab Nebula is the nearby ($\sim$ 2 kpc away) pulsar wind nebula and was the first source detected in VHE gamma rays. As the brightest VHE gamma-ray source visible from both hemispheres, it has become the so-called standard reference in VHE gamma-ray astronomy.

low energy threshold around 100 GeV, effective (by a factor of 100) rejection of hadronic showers and good angular ($\approx 0.1^\circ$) and energy ($\approx 15\%$) resolutions. At energies around 1 TeV, the performance achieves the best sensitivity, with a minimum detectable energy flux of $10^{-13}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$ [27].

Referring to currently running facilities, these arrays are capable to study not only pointlike, but also extended sources with an angular size up to 1 degree or larger. Moreover, the high flux sensitivity and relatively large ($\geq 3^\circ$) field of view of IACT arrays allow effective sky surveys. However, the potential of those arrays is rather limited for the search of very extended structures like the Galactic Plane diffuse emission or the huge radio lobes of the nearby radio galaxy Centaurus A. IACT arrays have great potential for "hunting" of solitary events like the possible VHE counterparts of GRBs, but they need trigger from other wavelenghts. In this regard, the detection technique based on direct registration of particles (leptons, hadrons and photons) of EAS is a complementary approach to the IACT technique [35].

2.3 Currently running facilities

Current IACTs have either a parabolic or spherical mirrors to focus the Cherenkov photons onto a tightly packed array of photomultiplier tubes (PMTs) placed in the focal plane. In both cases, they have many individual mirror segments having a radius of curvature equal to the focal length, placed on an optical support structure.
The energy interval from which the current generation of IACTs are sensitive ranges from 50 GeV to 80 TeV. Their angular resolution is on the order of 0.1°, the energy resolution around 15% and an integral energy flux sensitivity of a few times $10^{-13}\,\mathrm{erg\,cm^{-2}s^{-1}}$. Their sensitivity is sufficient to detect the Crab nebula in about 1 min, and a source with 1% of the Crab flux in 25 h.

HESS This observatory is located in Namibia (altitude 1800 m), in the Southern Hemisphere. It is the IACT with the largest field of view ($\sim 5^\circ$), that better allows for observations and surveys of extended sources, and the only one in the Southern hemisphere with good observation conditions for the Galactic plane. The initial four HESS telescopes (Phase I, completed in 2004) are arranged in the form of a square having a side length of 120 m, to provide multiple stereoscopic views of air showers.
In Phase II of the project, a single huge dish with about $600\,\mathrm{m}^2$ mirror area was added at

	HESS	**VERITAS**	**MAGIC**
Site	Namibia	Arizona	La Palma
Telescopes	2	$4+1$	4
Energy range	$50\,\mathrm{GeV} - 10\,\mathrm{TeV}$	$30\,\mathrm{GeV} - 100\,\mathrm{TeV}$	$85\,\mathrm{GeV} - 30\,\mathrm{TeV}$
θ_{68} at 1 TeV	0.08	0.07	0.08
$\Delta E/E$ at 1 TeV	20%	$15-20\%$	17%
Mirror size (m)	$12, 28$	12	17
Optical FoV diameter	$5°, 3.2°$	3.5°	3.5°

Table 2.1: Indicative performance values of currently operating IACTs. The angular resolution (θ_{68}) and energy resolution ($\Delta E/E$) at 1 TeV are given for 60% containment radius (i.e., the radius containing 60% of the light emitted by a pointilke source). Regarding HESS mirror size and FoV, the two values reported for each parameter correspond to the 12 m telescopes and to the single 28 m telescope.

the center of the array. This is operational since July 2012 and has increased the energy coverage, sensitivity and angular resolution of the instrument [39].

VERITAS Operating at the Fred Lawrence Whipple Observatory (FLWO) in Southern Arizona, USA. It is an array of four 12 m optical reflectors with similar characteristics as HESS-Phase I. These imaging Cherenkov telescopes are deployed such that they have the highest sensitivity from 100 GeV to 10 TeV. The field of view is $\sim 3.5°$ [41].

MAGIC Originally consisted of a single, very large reflector of 17 m diameter installed on the Canary island of La Palma, and with a 3.5° FoV. The first telescope has been fully operational since 2004. In 2009, a second telescope of essentially the same characteristics was added; MAGIC-II was installed at a distance of 85 m from MAGIC-I.
The system of those two telescopes is characterized by the largest collection surface ($\sim 236\,\mathrm{m}^2$) of any existing gamma-ray telescope worldwide [40].

Particle sampling arrays have complementary characteristics with a large FoV and a 100% duty cycle, as it has already been reported. Inaugurated in 2015, an EAS array with a good sensitivity is the High Altitude Water Cherenkov detector HAWC. Furthermore, LHAASO, a new installation in the Northern hemisphere, has been realized to cover a huge area and achieve an impressive sensitivity at energies of several tens of TeV.

	HAWC	**LHAASO**
Site	Mexico	China
Instrumented area	2200 m^2	7800 m^2
Energy range	100GeV − 50 TeV	100 GeV − 1 PeV
FoV (sr)	2	$\pi/3$
Angular resolution	$< 0.5°$ for E>1 TeV	0.45° at 1 TeV
	0.25° for E>10 TeV	$< 0.2°$ for E>6 TeV
Energy resolution above 6 TeV	$\sim 20\%$	$\sim 33\%$
Sensitivity per year at $\sim$ TeV	50 mCrab	9 mCU

Table 2.2: Indicative performance values of currently operating EAS.

HAWC It is a very high-energy gamma-ray observatory located in Mexico at an altitude of 4100 m. It consists of 300 steel tanks of 7.3 m diameter and 4.5 m deep, covering an instrumented area of about 22 000 m^2. Each tank is filled with purified water and contains three PMTs of 20 cm diameter, which observe the Cherenkov light emitted in water by superluminal particles in atmospheric air showers. Photons traveling through the water typically undergo Compton scattering or produce an electron–positron pair, also resulting in Cherenkov light emission. This is an advantage of the water Cherenkov technique, as photons constitute a large fraction of the electromagnetic component of an air shower at ground.

HAWC observes a large ensemble of sources, measuring their spectra and variability to characterize TeV scale acceleration mechanisms: it reaches, in a 1-year survey, a 50 mCrab[2] sensitivity at 5 standard deviations. Due to the large FoV ($\sim$ 2 sr), the experiment can observe diffuse gamma-ray emission from the plane of the Galaxy over a broad range of Galactic longitudes reaching the Galactic center. At present, 30 sources in the TeV catalog have been observed by, or also by, HAWC [13, 36].

LHAASO At higher energies, recently a new project called LHAASO (Large High Altitude Air Shower Observatory) and located at 4410 m above sea level in Sichuan Province, China, took hold very quickly. It consists of an extensive air shower detector array covering an area of 1.3 km^2 (KM2A) with 5195 scintillator counters (ED, 1 m^2 active area) and 1188 muon detectors (MD, water Cherenkov detector with an area of 36 m^2 buried under 2.5 m of dirt). In the center of the array the Water Cherenkov Detector Array (WCDA)

---

[2] The Crab, well-sampled from radio to TeV gamma rays, is used for the cross-calibration of ground-based detectors. 1 mCrab = 5.07 $\times 10^{-13}$ ph $cm^{-2}s^{-1}$ for a minimum energy threshold of 125 GeV.

is placed, covering 78 000 m^2, along with the Wide Field-of-view air Cherenkov/fluorescence Telescope Array (WFCTA) consisting of 18 telescopes.
KM2A is capable of rejecting the CR background by a factor of 10^{-2} at 20 TeV and 10^{-4} above 100 TeV. Along with an angular resolution of $\leq 0.3°$ for gamma-induced showers of up to 40° zenith angle, the KM2A is an essentially background-free gamma-ray detector at energies > 100 TeV. The energy resolution is better than 14% for photons above 100 TeV arriving from a zenith angle of $< 35°$.
WCDA reaches a sensitivity of 9 mCU[3] per year for gamma-ray sources of a few TeV by using three water ponds filled to a depth of 4.5 m, that fully absorb the air shower particles. The shower arrival direction is measured with an angular resolution of about 0.2° above a few TeV. WCDA covers a wide energy interval bridging *Fermi*-LAT and KM2A by having the same sensitivity of *Fermi*-LAT and KM2A at 200 GeV and 20 TeV, respectively. As a detector that continuously surveys 1/6 of the sky at any moment, WCDA is a device designed for monitoring and following up any transient phenomenon that may happen in its energy window.
WFCTA is designed for detection of atmospheric Cherenkov light produced by electrons and positrons from air showers initiated by primary CRs with energy ranging from 50 TeV to 100 PeV. The large field of view of the telescopes, 16° × 16° each, allows the telescope array to cover a wide patch of the sky for diffused CR measurements. The combined operation of these telescopes with KM2A and WCDA enables highly desired spectrum measurements of different CR species in a wide energy range that includes the spectral features known as 'first knee' and 'second knee' [18, 20].

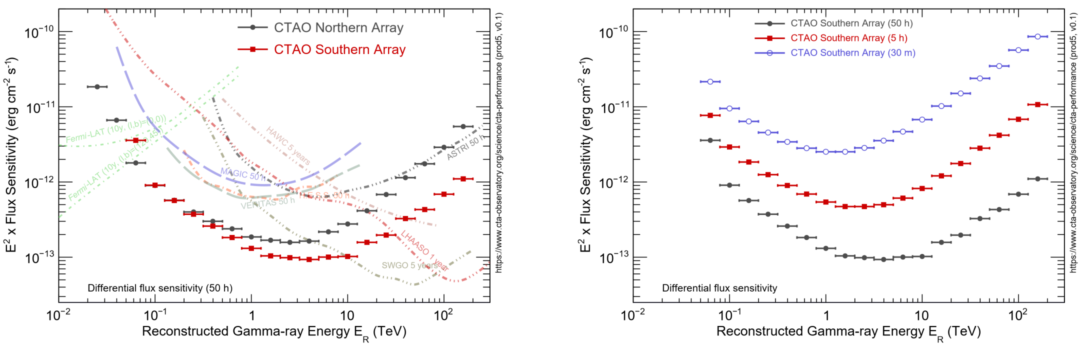
2.4 The future: CTA

The Cherenkov Telescope Array Observatory will be the next generation of IACTs overcoming the limits of current facilities. It has been designed to detect photons from 20 GeV up to 300 TeV, to have an almost full-sky coverage and a larger field of view by about a factor of 2 in diameter.
Current gamma-ray telescope arrays host up to five individual telescopes, whereas CTAO is designed to be the largest ground-based gamma-ray detection observatory of the next decade with several tens of telescopes located in the Northern (CTA-North) and Southern (CTA-South) hemispheres, in the so-called "Alpha Configuration". With a sensitivity improvement by up to a factor 5 to 10, depending on the energy range, of magnitude

---

[3] CU, flux of the Crab Nebula at 100 TeV; $1\,\mathrm{CU} = 6.1 \times 10^{-17}$ photons $\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

with respect to currently operating IACTs, coupled with significantly better angular resolution ($\leq 0.05°$ at $E > 1$ TeV, see Figure 2.7), the array will be used to address many open questions in high-energy astrophysics.
In addition, CTA will explore the ultra-high energy ($E \geq 50$ TeV) window with the highest energy resolution ($\Delta E/E < 0.1$ in $1-10$ TeV range, as shown in Figure 2.8), providing a completely new view of the sky beyond the "edge" of the known electromagnetic spectrum [38].



(a) (b)

Figure 2.6: Comparison of the CTAO performance with existing gamma-ray instruments. **(a)**: The differential sensitivity (on axis and with 50 hr of observations) shown above is defined as the minimum flux needed by CTAO to obtain a 5σ detection of a point-like source, calculated in non-overlapping logarithmic energy bins (five per decade). The analysis cuts in each bin have been optimised to achieve the best flux sensitivity to point-like sources. The optimal cut values depend on the duration of the observation. Note that the curves for *Fermi*-LAT and HAWC are scaled by a factor 1.2 to account for the different energy binning; furthermore, the plotted curves allow only a rough comparison of the sensitivity of the different instruments, as the method of calculation and the criteria applied are not identical. In particular, the definition of the differential sensitivity for HAWC is rather different due to the lack of an accurate energy reconstruction for individual photons in the HAWC analysis. **(b)**: differential flux sensitivity of CTA provided for three different observation times: 0.5, 5 and 50 hours. **Credit**: CTAO webpage [38].

The design was driven by the necessity of a full sky coverage with the study of transient events being the main scientific topic. Access to almost the full sky is necessary as many of the phenomena to be studied by CTA are rare and individual objects can be very important. For example, the most promising galaxy cluster, the brightest starburst galaxy and the only known gravitationally-lensed TeV source are located in the North. The inner Galaxy and the Galactic Centre (GC) are key CTA targets and are located in the South. Full sky coverage ensures that extremely rare but critically important events (for

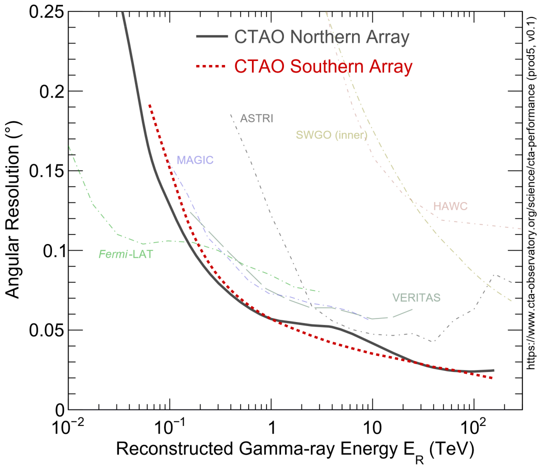
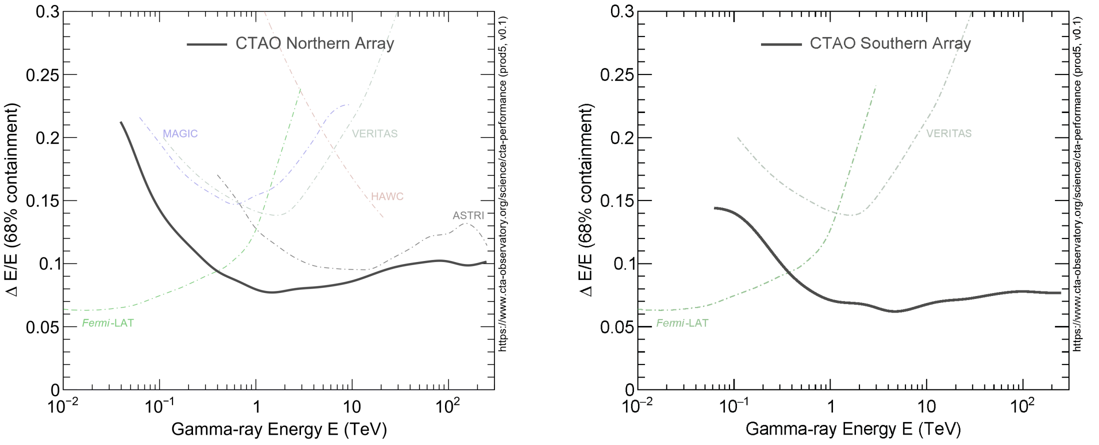


Figure 2.7: The angular resolution vs. reconstructed energy curve (E_r) shows the angle within which 68% of reconstructed gamma rays fall, relative to their true direction. Gamma-ray photon/hadron separation cuts are applied for the MC events used to determine the angular resolution. Dedicated analysis cuts can provide improved angular (or spectral) resolution at the expense of collection area, enabling e.g. a better study of the morphology or spectral characteristics of bright sources. **Credit**: CTAO webpage [38].

example, a Galactic supernova explosion, bright gravitational wave transient, or nearby GRB) will be accessible to CTA.

For what concerns the very wide energy range covered by the CTAO array, it is necessary the use of at least three different telescope types, each of which optimized for a specific energy range: the Large-Sized Telescopes (LSTs), the Medium-Sized Telescopes (MSTs) and Small-Sized Telescopes (SSTs). Because gamma rays with low energies produce a small amount of Cherenkov light, telescopes with large mirrors are required to capture the images. Thus, the large size provides sensitivity at the lowest energies, avoiding significant gamma-gamma absorption by the EBL and allowing access to a wider volume of the Universe (due to the EBL absorption of VHE photons, at 1 TeV it is possible to observe objects at a distance up to hundred of Mpc, at 20 GeV the volume of the visible Universe increseases significantly). The small telescopes spread across a wide area provide sensitivity at the highest energies because very high-energy gamma-ray showers produce a large amount of Cherenkov light over a large area. Furthermore, above 1 TeV, reflector area of single telescopes is not essential while collection area (i.e., the amount of area capable of collecting electromagnetic radiation of the entire array) is fundamental because gamma-ray spectrum decreseases and so does the flux of the ultra high-energy



(a) (b)

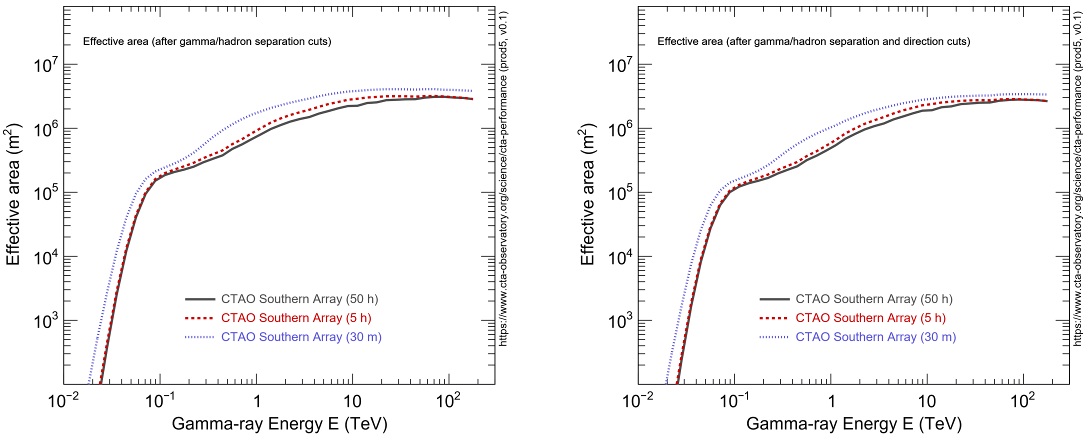
Figure 2.8: The energy resolution $\Delta E/E$ is obtained from the distribution of $(E_R – E_T)/E_T$, where R and T refer to the reconstructed and true energies of gamma-ray events recorded by CTA. $\Delta E/E$ is the half-width of the interval around 0 which contains 68% of the distribution. The plots show the energy resolution of the Northern ((a) panel) and Southern ((b) panel) as a function of reconstructed energy (the result depends only weakly on the assumed gamma-ray spectrum; for the results here it was used $\mathrm{d}N_\gamma/\mathrm{d}E \approx E^{-2.62}$). **Credit**: CTAO webpage [38].

photons. Therefore, the compromise is to use many small reflectors. Ultra high energies are needed to study the extreme accelerators which we know to be present from direct cosmic-ray measurements are present in our galaxy.

A wide energy range maximises the chances of detection of new source classes with unknown spectral characteristics, for example in the search for dark matter, and it is the key for discrimination between scenarios and to identify features.

In the North, the array will be more compact due to the site geography and this will result in a more limited energy range at the ultra high energies, which are those of extreme phenomena of the nearby sources in our Galaxy. Therefore, most of the studies will focus on extragalactic targets and, in the interest of optimisation of the observatory, the CTAO Northern Array will be implemented with only large and medium telescopes. The inner regions of the Galaxy, instead, are better accessible (but not only) from the Southern hemisphere.

The large telescope number and individual wide telescope fields of view result in a CTA collection area which is one or two orders of magnitude larger than current generation instruments at essentially all energies, with substantial benefits for imaging and spectroscopy studies (see Figure 2.9).



(a) (b)

Figure 2.9: Effective collection area plots are provided here only for the CTA South, as in my thesis I will deal with simulations of the Southern array. In both panels, the resulting plots are obtained after gamma-ray photon/hadron separation cuts. **(a)**: The effective collection area for gamma rays from a point-like source shown vs. *E* for cuts optimised for 0.5-, 5- and 50-h observations (no cut in the reconstructed event direction applied). **(b)**: The effective collection area with cuts in the reconstructed event direction. **Credit**: CTAO webpage [38].

The dramatic improvement in the point-source sensitivity of CTAO with respect to current instruments is a consequence of the combination of improved background rejection power, increased collection area and improved angular resolution. The improved background rejection power is achieved primarily through high image multiplicity and is particularly important for the study of extended, low-surface brightness objects and for low-flux objects where deep exposures are required. Figure 2.6 and Figure 2.7 compare the sensitivity and the angular resolution of the CTAO arrays to a selection of existing gamma-ray detectors, respectively.
A summary of CTAO telescopes technical details and characteristics is shown in Table 2.3.

2.4.1 Observatory Operations

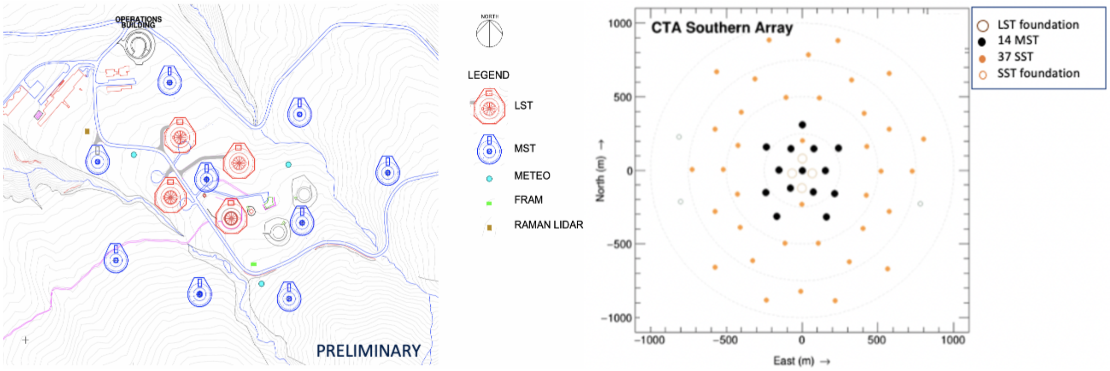
The CTA Observatory is a geographically distributed facility hosting the Headquarters in Bologna (Italy), the Science Data Management Centre in Zeuthen (German) and two array sites. The number of individual telescopes in the CTAO arrays, and the ability to operate multiple sub-arrays independently, provide enormous flexibility of operation.

Also, CTAO will be the first ground-based gamma-ray observatory open to the worldwide astronomical, particle and plasma physics communities as a resource for data from unique high-energy astronomical observations.
Unlike existing ground-based VHE instruments, CTAO will be operated as a proposal-driven open observatory. Observations will be carried out by observatory operators, then the data will be calibrated, reduced and, along with analysis tools, made available to users. After a proprietary period of about one year, data will be made openly available through the CTA data archive.

2.4.2 The telescope array

The larger Southern array will be constructed in the Atacama Desert of Chile at an elevation of 2100 m, close to the existing Paranal Observatory of the European Southern Observatory (ESO). CTA-North Array instead will be built at an elevation of 2200 m on the island of La Palma, Spain, located within the existing Observatorio del Roque de los Muchachos of the Instituto de Astrofisica de Canarias which already hosts the operating gamma-ray observatory MAGIC telescopes, as well as a wide variety of optical telescopes of various sizes.
In the past, the "Omega Configuration", formerly referred to as the baseline configuration with 118 telescopes divided into both sites, was considered. This refers to the full-scope configuration that could be deployed in the operation and enhancement phase, depending on the available funds. The "Alpha Configuration" is the current official configuration and is not a subset of the "Omega Configuration" in terms of telescope positions.
The approved distribution of telescopes for the Southern and Northern Arrays of the CTA Observatory consists of 4 LSTs and 9 MSTs covering an area of $\approx 0.25\,\mathrm{km}^2$ in the CTAO Northern Array, 14 MSTs and 37 SSTs covering $\approx 3\,\mathrm{km}^2$ in the CTAO South Array. The definition of these configurations approved in June 2021 is the result of a meticulous optimization process for each array's scientific capabilities, which implies the specialization of the Northern array in extragalactic sources (low and medium CTAO's energy range) and that of the Southern array in Galactic targets (medium and high CTAO's energy range) for the first construction phase.

While the individual telescopes may vary in size and design, CTAO telescopes will be constructed and will perform similarly. Each telescope will have a mount that allows it to rapidly point towards targets (pointing time of $\leq 90\,\mathrm{s}$ to any point in the sky with $> 30°$



(a) (b)

Figure 2.10: Proposed layout as the "Alpha Configuration" for the three sizes of telescopes (LSTs, MSTs and SSTs) in CTA-North (a) and in CTA-South (b). **Credit**: CTAO webpage [38].

elevation) and will comprise a large segmented mirror to reflect the Cherenkov light to a high-speed camera that can digitize and record the image of the shower.

Large-Sized Telescopes (LST) Telescopes with large mirrors are required to capture images from gamma rays with low energies that produce a small amount of Cherenkov light. The centres of both arrays will thus be covered by four LSTs in order to provide the full-system sensitivity between 20 and 150 GeV. The LST is an alt-azimuth telescope with a 23 m diameter parabolic reflective surface, supported by a tubular structure made of reinforced carbon fibre and steel tubes. The reflective surface of 400 m^2 collects and focuses the Cherenkov light into the camera, where photomultiplier tubes (PMTs) convert the light in electrical signals that can be processed by dedicated electronics. The camera has a total FoV of about 4.3° and has been designed for maximum compactness and low weight, cost and power consumption while providing optimal performance at low energies. Although the LST will stand 45 m tall and weight around 100 tonnes, it will be extremely responsive, with the goal of being able to re-position within 30 seconds. Both the re-positioning speed and the low-energy threshold provided by the LSTs are critical especially for CTA studies of extragalactic transients, high redshift AGN and GRBs. The LSTs will expand the science reach to cosmological distances and fainter sources with soft energy spectra.

Medium-Sized Telescopes (MST) The MST will be CTAO's "workhorse", with sensitivity in the core energy range of CTA, from about 150 GeV to 5 TeV and they will be

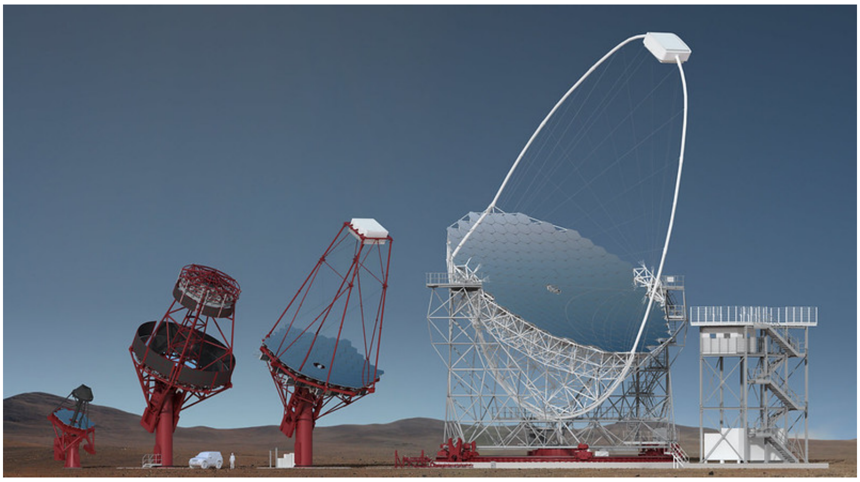


Figure 2.11: From left: the SST prototype design, the two medium-sized prototype designs (SCT and MST) and the LST prototype design. **Credit**: Gabriel Pérez Diaz, IAC.

distributed around the LSTs. The mirror will be about 12 m in diameter and will have two different camera designs that use PMTs. The MST cameras will have a large FoV of about 8°, enabling the MSTs to take rapid surveys of the gamma-ray sky. The MST is a modified Davies-Cotton telescope with a reflector diameter of 12 m on a polar mount and a focal length of 16 m. It will have up to 90 hexagonal-shaped mirrors that are aligned with an active mirror control assembly to provide a uniform reflector.

A dual-mirrored version of the MST, the **Schwarzschild-Couder Telescope (SCT)**, is proposed as an alternative type of medium telescope that will eventually be considered for a future upgrade of the CTAO Southern Array. The SCT's two-mirror optical system is designed to better focus the light for improved imaging detail and improved detection of faint sources. The SCTs have improved angular resolution value as a result of a smaller point spread function (PSF) and the very large number of camera pixels, based on silicon photomultipliers (SiPMs), covering approximately 8° FoV. Both the 9.7 m diameter primary and 5.4 m secondary mirrors are segmented and have active alignment. The SCT uses the same positioner as the MST and the camera technology developed for the SCT is also used in the SST-2M GCT (see below description of the SST).

Small-Sized Telescopes (SST) SSTs will outnumber all the other telescopes planned to be spread out over $\sim 3\,\text{km}^2$ in the Southern array only. As already stated, this is due to the fact that very high-energy gamma-ray showers produce a large amount of Cherenkov light over a large area, and the SST's smaller mirror is sensitive to the highest

	LST	**MST**	**SST**
Optimized energy range	20 GeV − 3 TeV	80 GeV − 50 TeV	1 TeV − 300 TeV
Energy range (in which subsystem provides full system sensitivity)	20 GeV − 150 GeV	150 GeV − 5 TeV	5 TeV − 300 TeV
Number of telescopes	4	23	37
Primary reflector diameter	23.0 m	11.5 m	4.3 m
Effective mirror area	370 m^2	88 m^2	8 m^2
Positioning time to any point in the sky ($> 30°$ elevation)	30s	90s	60s
Pointing precision	$<$ 14arcseconds	$<$ 7arcseconds	$<$ 7arcseconds
Optical FoV diameter	4.3°	7.7°	10.5°

Table 2.3: Summary of some mentioned CTAO telescopes technical details.

energy gamma rays (between a few TeV and 300 TeV). The SSTs' wide coverage and high sensitivity improve CTA's chances of detecting the highest energy gamma rays.
The CTA-SST design is based on the ASTRI design, a dual-mirror Schwarzschild-Couder aplanatic configuration. Thanks to its small plate scale, it uses a novel compact camera based on SiPM sensors with a FoV of approximatly 10°. The 4.3 m diameter primary mirror is segmented into hexagonal facets and the 1.8 m secondary mirror is monolithic.

2.4.3 Synergies

The scientific output of CTA can be very much strengthened by operating in synergy with other facilities at different wavelengths and messengers. Indeed, multi-wavelength (MWL) and multi-messenger (MM) studies using CTAO provide added value to the science cases in two main ways: to understand the origin of cosmic rays and the extreme physical environments that produce them, it is necessary to study non-thermal signatures that span many orders of magnitude in frequency in the broad-band spectral energy distribution (SED) of a given object; information on the nature of gamma-ray emitting sources can be provided by MWL observations, enabling, for example, the object class, environmental conditions or the distance to be established. Moreover, messengers like neutrinos and gravitational waves (GW) provide crucial informations about astrophysical sources from which they come as they are not affected by deviations due to e.g. the magnetic field.

Radio to Sub-millimeter The combination of radio measurements with those at very high energies can provide limits on the electron density independent of assumptions about magnetic field strengths and can help determine which of several competing non-thermal processes dominate at the highest energies. Radio measurements also provide important magnetic field constraints via Faraday rotation and provide the ephemerides of known pulsars, to guide the search for potential gamma-ray pulsations with CTA.
Gamma-rays are tracers of the interstellar gas and CTAO's sensitivity to diffuse emission around accelerators makes mapping of this gas over wide areas absolutely essential to enable identification of sources in the Galactic plane and within other large-scale surveys such as that of the LMC. Sub-millimeter wavelengths thus complement CTA science by offering a detailed understanding of the environment into which shock waves propagate and through which accelerated particles are transported and interact.
Eventually, mm/sub-mm observations along with CTAO can be used to directly study the relation between near event horizon physics and cosmic-ray acceleration and non-thermal processes in microquasar's jet as well as in the innermost regions of nearby AGN.

Infra-red (IR)/Visible (V) through Ultra-violet (UV) The last few years have revealed that many compact, high-energy sources emit detectable levels of synchrotron emission in the IR/V, which can also display very fast variability. These sources are also high energy gamma-ray emitters, making IR/V a new frontier for MWL exploration [21]. In addition, IR/V studies of non-radiative shocks in supernova remnants can provide useful constraints on non-linear particle acceleration and its influence on shock heating. Finally, IR/V observations provide an interesting perspective in the case of gamma-ray binaries, as properties of circumstellar discs may directly affect inter-wind shocks and lead to light-curve evolution.
Furthermore, polarisation offers an ideal method for isolating the synchrotron/non-thermal component in cases of mixed emission. Polarimetry can be also be used to provide new insights to broad-band SED correlations, on top of deriving magnetic field parameters in jets' studies, also useful to improve the modelling of SEDs and emission-region localisation.
Finally, the ultra-violet (UV) domain probes synchrotron emission of electrons which

have comparable energies to those emitting inverse-Compton emission detected by CTAO. As such, simultaneous UV observations of bright AGN, blazars and other variable objects can be extremely useful, as long as they are not too absorbed by interstellar gas.

X-ray and γ-ray Phenomena which result in high enough temperatures for thermal X-rays to be produced are very often associated with shock waves, accretion or high velocity outflows, and hence with particle acceleration and gamma-ray emission. This thermal X-ray emission from gamma-ray sources provides valuable information about plasma properties (e.g. temperatures, densities) and energetics (e.g. outflow/shock velocities).
Non-thermal X-ray emission also provides a natural tracer of locations of extreme particle acceleration. The combination of VHE gamma-rays and X-rays is extremely powerful for constraining magnetic field strengths, the electron to proton ratio of the accelerated particles and the particle energy distributions.
The hard X-ray / soft gamma-ray domain ($0.1 - 10\,\mathrm{MeV}$) represents a very useful window for the non-thermal spectra of astrophysical sources, but is at higher gamma-ray energies that the synergies with CTAO are stronger and the instrumental performance better matched to CTAO. The GeV domain is dominated by pion decay, bremsstrahlung and inverse Compton emission, and in combination with the TeV range can help identify the dominant radiative mechanisms.

Ground-based VHE gamma-ray instruments As has already been mentioned, several ground-based VHE gamma-ray instruments, such as particle sampling arrays, may be operational at the same time as CTAO. None of these instruments are direct competitors, but rather provide complementary performance.
Also one or more of the current generation of IACT arrays may continue operations into the contruction phase of CTAO. Use of these telescopes for monitoring could be considered, under suitable agreement between the telescope and CTAO. In particular, for cases when the CTAO sites are at different longitudes than current IACTs, these can extend monitoring of bright flaring sources to periods before and after the CTA observations, even if with different sensitivity (refer to Figure 2.6a).

Neutrinos and Gravitational Waves Essentially all mechanisms invoked for the production of gamma rays will also produce high-energy neutrinos, and unlike charged cos-

mic rays, both point back to their sources, so that theoretically it is possible to identify the sources of the incoming particles. There is thus strong complementarity to observations with these two messengers.
Neutrino telescopes, using e.g. reconstructed muon tracks, are less able to precisely pinpoint the origin of neutrinos. However, the possibly neutrinos detected are the only completely unambiguous tracer of hadronic acceleration, even up to high redshifts and beyond PeV energies, a combination that is not possible for gamma-ray telescopes due to photon-photon absorption. CTAO is the ideal instrument to follow up on any VHE neutrino clustering, i.e. detection of numerous neutrinos from the same sky direction, necessary to localise and characterise the VHE accelerators.

Besides with neutrinos, gravitational waves are the newest messengers to be considered in the MM analysis. Mergers of BHs, NSs and mixed systems are generators of those events that could be detectable by ground-based GW interferometers out to a few hundred Mpc [1], with the expectation of several to a hundred GW transients per year [2]. However, until the advent of third-generation detectors, the localisation errors will be large and asymmetric (at best, source localisation accuracy is several tens of square degrees), especially if compared to CTA's field of view.

Chapter 3

Galactic Plane Survey

Astronomical surveys of the Galaxy provide essential, large-scale datasets that form the foundation for Galactic science at all photon energies.
The current generation of ground-based instruments has revealed that more than half of the detected TeV sources cluster along the Galactic plane (90% of the Galactic sources lie at $|b| \leq 3.5$ degrees in latitude [3, 19]). PWNe are the most numerous source population, followed by SNRs and binaries. However, most Galactic plane sources remain unidentified for three main reasons: they have multiple associations at lower energies which cannot be disentangled; they consist of several still unresolved sources; they are completely "dark", with no counterpart at any other wavelength, as already stated in Section 1.2.

The CTA Galactic Plane survey (GPS) will lead to a significant improvement with respect to current IACTs in the understanding of our Galaxy at TeV energies, providing the first complete very-high-energy scan of the entire Galactic plane with a sensitivity better than 4 mCrab [43]. This will allow scientists to perform spectral and morphological studies with unprecedented precision resolving spatial sub-structures on arcminute scales, as well as reconstructing previously unmeasurable spectral features.

3.1 A Key Science Project

The Key Science Projects (KSPs) are proposals that consist of large projects ensuring that the key science issues for CTAO are addressed in a coherent fashion and provide legacy science data products. Among the different proposed KSPs [21], four surveys are planned to be part of CTA's Core Programme:

- the Galactic Plane survey;
- a deep survey with better sensitivity with respect to the overall GPS (few mCrab) of the inner Galaxy and the Cygnus region;
- the Large Magellanic Cloud survey;
- the extragalactic survey, covering $1/4$ of the sky to a depth of $\sim 6\,\mathrm{mCrab}$.

Among those, the GPS will fulfil a number of important science goals, including:

- providing a complete census of the VHE gamma-ray sources, allowing, for the first time at these energies, to conduct population studies of SNRs and PWNe mainly through the detection of hundreds of new sources;
- identifying a list of promising targets for follow-up observations, such as new gamma-ray binaries and PeVatron candidates, that are of key importance in the search for the origin of cosmic rays;
- determining a precise characterization of the diffuse emission in the $0.1 - 100\,\mathrm{TeV}$ energy range across the entire Galactic plane;
- producing a multi-purpose, legacy dataset comprising the complete Galactic plane at very high energies that will have long-lasting value to the entire astronomical and astroparticle physics communities;
- likely reveal many new and unexpected phenomena.

The GPS will be carried out using both the Northern and the Southern arrays, and will provide complete coverage of the Galactic plane. The exposure time will be most likely non-uniform across the plane and will be determined by the scientific goals to be addressed in different regions. Because of the knowledge from the current Galactic VHE source population and from other HE surveys, the inner Galactic region (-60° $\leq l \leq$ 60°) will be most likely allocated significantly more observation time than other regions, allowing to reach a sensitivity of a few mCrab.

As for the pointing patterns, a single-row and a double-row patterns strategy has been considered as well as a triple-row pattern and a non equilateral double-row pattern with independent spacing in longitude and latitude. In particular, the last one, with a pattern step s $=$ 2.25° (h $=$ 1.95°), has proven to provide the best sensitivity in the Galactic

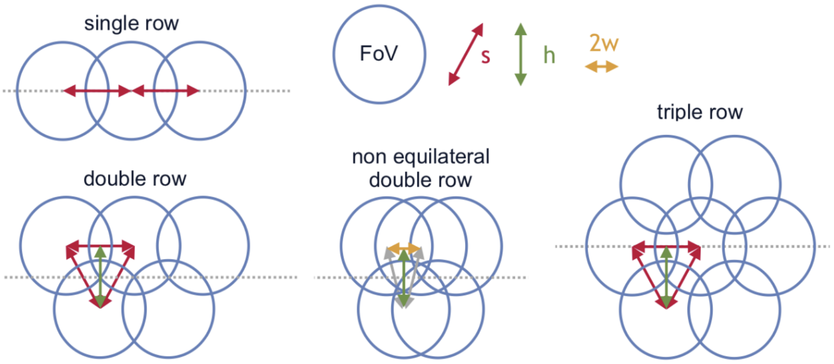


Figure 3.1: Schematic view of the pointing patterns considered for the GPS. Dashed lines represent the Galactic equator. Circles represent the CTAO FoV for an individual pointing. Red lines show the pattern step (s), i.e., the distance between pointing directions of adjacent pointings for single row and equilateral patterns. Green lines show the latitude spacing (h) for patterns with multiple rows. The orange line shows twice the longitude spacing (w) for the non equilateral double row pattern. For easier comparison with the other patterns, the step of the non equilateral double row pattern is defined as $s = \sqrt{4/3}\,h$, the step of an equilateral double row pattern with the same latitude spacing h. Figure taken from [34].

plane and almost as good as the triple-row pattern at higher latitudes [34]. The pointing patterns are illustrated in Figure 3.1.

From the surveys that have already been completed at VHE, we know that PWNe are the largest source class, followed by SNRs and gamma-ray binaries, as already stated. About two-thirds of known sources are not yet firmly identified; most have multiple plausible associations that are challenging to disentangle, although some appear to be dark accelerators that are not detected at lower energies. Only a handful of the VHE Galactic sources are point-like in nature (largely binary systems). The large majority of the so far detected sources have extended VHE emission, with a typical angular size of $\sim 0.1^\circ - 0.2^\circ$ (in radius), and a few are considerably larger than this, although this average extension is most likely an observational bias due to HESS and VERITAS limited gamma-FoV[1], the only two IACTs among the currently running ones that carried out a part of the GPS; the worsening of the acceptance of these facilities, i.e. the effective area integrated over solid angle, as increasing the distance from the pointing direction leads to difficulties in identifying sources that extends more than 1°. The reconstructed spectra are generally well fit by power-law spectral models in the energy range between several

---

[1] Related to the entire array, it is defined as the distance from the pointing direction center to the point where sensitivity worsens by a factor of 2.

hundreds of GeV and several tens of TeV, with typical differential photon indices in the range of $\Gamma \sim 2.3$ - 2.7 [21].

3.2 Data Challenge: results

In 2018 the CTA Consortium run the first internal Science Data Challenge. The goal was to enable the CTA Consortium Science Working Groups to derive science benchmarks for the Key Science objectives.

The first CTA Data Challenge (DC-1) was focused on the validation of the CTAO Science Analysis Tools prototypes and had the goal to prepare for the analysis of real CTAO data and to validate features and completeness of science tools packages. Moreover, it was an opportunity to push the development of new algorithms and ensure the compatibility of CTAO data with multiple tools and the external scientific community. It also resulted in numerous images, spectra and light curves, that can be used to illustrate CTA's science case and that enriched the outreach material.

The GPS was one of the four KSPs that were simulated using realistic pointing patterns.

Recently, a new sky model has been developed for the simulation and analysis of the GPS, including improved physically-motivated models for Galactic source populations and interstellar emission, advance on the optimization of the survey strategy and development of pipelines to derive source catalogues tested on simulated data [34]. Sources were modelled using measurements from past and current IACTs based on the compilation provided by gamma-cat [22] along with the three main classes of Galactic VHE gamma-ray emitters (PWNe, SNRs and gamma-ray binaries) based on physical models inferred by observations and theory. Moreover, sources detected by *Fermi*-LAT based on the 3FHL catalogue [11] and by HAWC based on the 2HWC catalogue [5] not included in gamma-cat, were considered. The same amount of known and already included bright sources have been removed from the synthetic populations of PWNe e SNRs. Detailed information about how emission from SNRs, PWNe and binaries were modelled and simulated can be found in [34].

Through the $\log N - \log S$ distribution of the VHE sources, i.e. the cumulative source counts as a function of source photon flux integrated at energies above 1 TeV, it is then possible to estimate the number of VHE sources expected to be detected by CTAO. Figure 3.2 illustrates the obtained $\log N - \log S$. The vertical shaded band indicates the sensitivity target for the CTAO GPS program [21]. The two dominant source classes detected

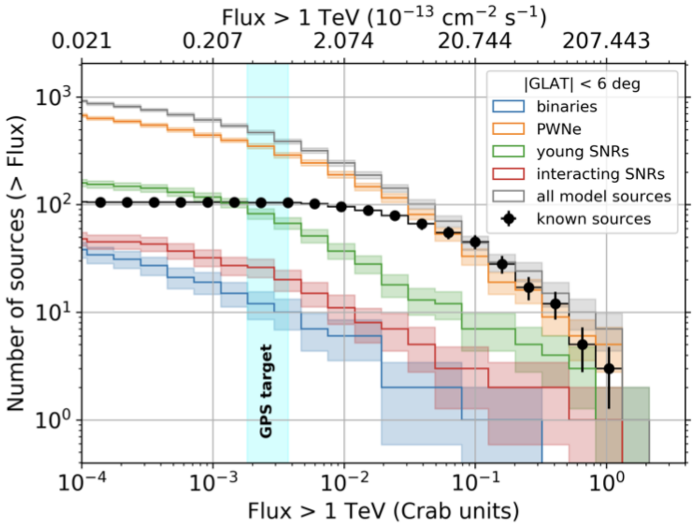


Figure 3.2: Cumulative source counts as a function of source photon flux integrated at energies above 1 TeV, as measured by existing and past IACTs (from gamma-cat), from *Fermi* LAT (3FHL), and HAWC (2HWC), and as predicted by the population synthesis for Galactic gamma-ray sources. Known or synthetic sources are only included if their distance from the Galactic equator is $< 6°$. The vertical shaded band indicates the sensitivity target for the CTA GPS program. Figure taken from [34].

in the survey are PWNe and SNRs; CTA GPS may be able to increase the population size of PWNe about seven times the current sample and by a factor larger than two the SNRs observed at TeV energies.

Figure 3.3 shows a realistic simulated image of the GPS obtained from the simulated data built in the catalog of [34]: sources of the entire Galactic plane are located at latitudes $|b| < 6°$ and for energies between 0.07 and 200 TeV.

Furthermore, also the interstellar high-energy emission along the Galactic plane need to be modelled and analysed. Galactic interstellar medium diffuse model is based on predictions from codes that solve the cosmic ray transport equations and calculate the related multiwavelength emission. It includes emission from interstellar gas, produced by hadronic interactions of cosmic-ray nuclei and by electron/positron Bremsstrahlung. It also accounts for the IC radiation produced by the interactions of cosmic-ray electrons and positrons with low-energy photons. Figure 3.4 shows the counts map of the obtained diffuse model.

The development of pipelines for the GPS catalog is one of the most significant achieve-

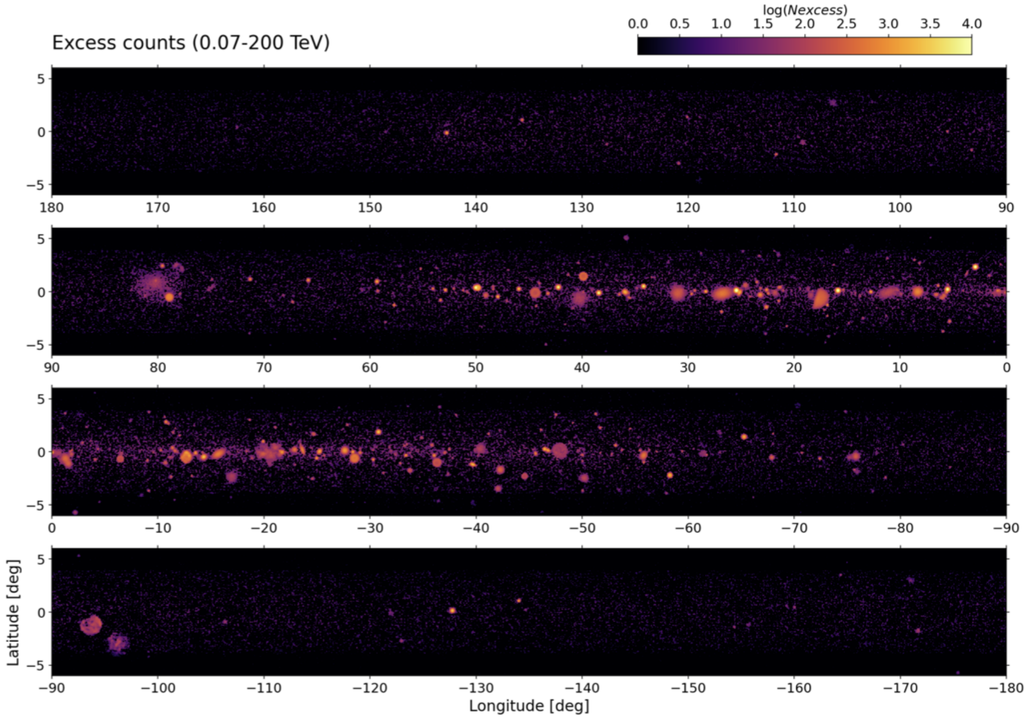


Figure 3.3: Excess counts above the instrumental background model for the full GPS survey in the energy range $0.07 - 200$ TeV. This simulated CTA image of the entire Galactic plane adopt the proposed GPS observation strategy in [34]. Figure taken from [34].

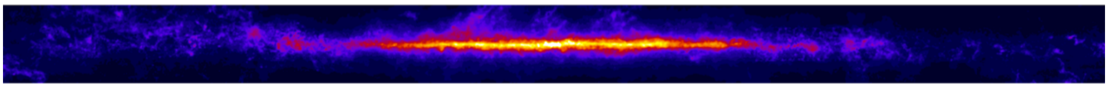


Figure 3.4: Simulated CTA image of the Galactic interstellar medium diffuse emission modeled including both emission from the interstellar gas and the inverse Compton component. Figure taken from [43].

ments of the DC-1. The obtained results are fairly good and promising: about 250 sources has been detected and 80% ($> 5\sigma$) of them has been correctly identified [28]. However, these results are obtained for the original full-scope "Omega Configuration" and need to be revisited in the future at the light of the recently defined layout, the "Alpha Configuration". Preliminary results show that the downsizing project leads to a 10% reduction in the GPS source detection.

3.3 Source Confusion

The improved sensitivity of CTAO with respect to current generation IACT instruments will result in a large number of sources detected but will also lead to a higher probability of source confusion. The source confusion problem concerns multiple overlapping sources decteted as a single one and multiple detections related to extended structures. In contrast with surveys at lower energies [4], the processes involved in the radiation of VHE photons are generally driven by either large timescales (e.g., for instance, of continuous injection of particles) or by the interactions with the CR pool permeating the Galaxy or with nearby sources like molecolar clouds extending several parsec as in the Galactic centre ridge [6] or large SNR shells [3]. The large size of these structures, along with the large number of sources that will be detected, projected on the sky plane, will contribute to an extremely large source confusion. The analysis techniques currently used in the field of the Cherenkov/VHE astronomy are severally limited by systematic errors when deriving information about large-scale emission, such as the interstellar diffuse emission, larger than the gamma-ray FoV, usually lower than the optical FoV reported in Tables 2.1 and 2.3, so that even sources extended few degrees ($\sim 1^\circ - 2^\circ$) can lead to those systematic errors. This limitation, combined with the superposition of sources, will prevent the maximum exploitation of the instrument, limiting the minimum flux and therefore, scientific results that will be derived.

Therefore, source confusion is expected to be a significant problem along the Galactic plane and, in particular, in the inner regions of the Galaxy. The main issues to be considered are the unknown shapes of the sources, the unknown level of diffuse emission, the high source density in the inner Galaxy (so that many sources that will be detected by CTAO will overlap), and the dependence of source identification on the analysis methods. The added value of MWL information and the improved angular resolution of CTA will help in resolving many instances of source confusion, but a large fraction of the sources detected in the GPS will naturally be relatively weak and thus inherently very

difficult to resolve and identify.
Initial studies have been carried out with the goal of estimating the expected level of source confusion in CTA GPS. In one of these investigations, a Galactic population of sources was simulated based on our current understanding of the VHE source population: an extrapolation of the source counts distributions (i.e. $\log N - \log S$) of source photon indices and sizes consistent with existing data were adopted, along with a spatial distribution of sources around the Galactic Centre. No diffuse emission was included (except for the GC ridge) and two different extrapolations of the $\log N - \log S$ distribution were used to bracket the range of the expected source density. A position in the sky was considered confused if there was more than one simulated source within a radius of 1.3 times the CTA angular resolution. Using these assumptions leads to an approximate lower limit to the amount of source confusion. The estimated confusion lower limits range from 13-24% at 100 GeV to 9-18% at 1 TeV, for the region $|l| < 30°$ and $|b| < 2°$ [21].

My thesis project aims at studying the specific problem of the impact of the source confusion for the GPS in a systematic way that has never been done before, analysing a wide paramters space. The goal is to obtain general results that allow to draw clear conclusions for future CTAO observations of the confused region that has been considered here (see Section 4.3).

3.4 PeVatrons

CRs are primarily energetic nuclei; their spectrum, observed on the Earth, shows two major spectral features called the knee and the ankle, although also a second knee is present as the possible presence of a second steepening of the spectrum. The CR spectrum (Figure 3.5) is dominated by protons up to the first knee which is located at an energy of a few PeV, where the differential power-law spectrum steepens from $E^{-2.7}$ to $E^{-3.0}$. Up to the first knee, they are believed to have origin in our Galaxy and, in order to maintain the CR intensity at the observed level, they must provide $\sim 10^{41} \mathrm{erg\,s^{-1}}$ in the form of accelerated particles.
Sources that are able to accelerate CR up to the first knee are dubbed as Pevatrons. They are closeby sources whose importance is related to the origin of the most energetic CRs in our Galaxy.

PeVatron sources are expected to have hard power-law spectra (i.e. not much steeper

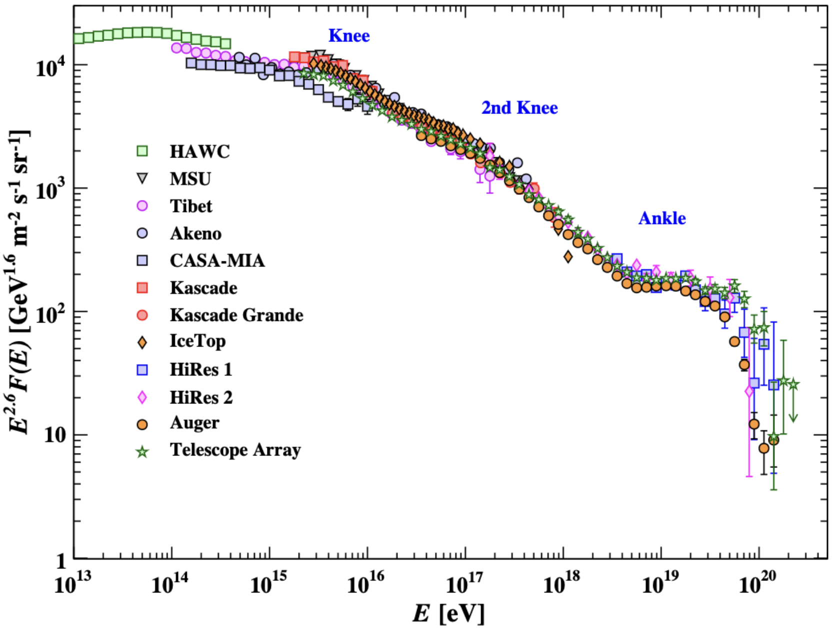


Figure 3.5: All-particle spectrum as a function of E (energy-per-nucleus) from air shower measurements. The characteristic knee is visible at an energy of a few PeV while at the energy of $\sim 10^{19}$ eV, the spectrum becomes flatter in correspondence to the second transition point, called the ankle. The possible presence of a second steepening of the spectrum near 8×10^{16} eV is the so-called "second knee". **Credit:** PDG ([44]).

than E^{-2}) that extend up to energies of $\sim$ 50 TeV and beyond. This is the VHE domain above which the problematic ambiguity between leptonic and hadronic emission mechanisms is resolved due to the Klein-Nishina effect, where the cross-section for IC electron and CMB photon interactions decreases very quickly. Indeed, a hadronic component could pop up above the leptonic component for energies $E > 50$ TeV. Theoretically, SNRs provide adequate conditions to have efficient CR acceleration through Diffusive Shock Acceleration (DSA, [14]), i.e. the mechanism through which CRs gain energy by repeatedly passing back and forth between the upstream and downstream plasmas. SNRs are promising PeVatron candidates since they satisfy the energetic requirements if they can convert $\sim$ 10% of their kinetic energy into accelerated particles. The acceleration of Galactic CRs in SNRs and their subsequent interactions between hadrons and ambient matter, and between leptons and radiation, make these objects bright in gamma rays. The amount of relativistic particles present in the acceleration region increases with time as the SNR passes through its free expansion phase, and reaches a maximum in the early stages of the Sedov phase; correspondingly, the peak in gamma-ray luminosity

typically appears $10^3 - 10^4$ years after the supernova explosion. Indeed, over the last 20 years, many young and middle-aged SNRs (an aged SNR is $\sim 10^6$ years old) have been detected in the GeV and TeV bands. Most of them show a shell-like morphology, supporting the general prediction of DSA and, thus, establishing the SNR shocks as effective particle accelerators. The spectral feature of the pion bump has been detected at sub-GeV energies in a handful of aged SNRs, proving the hadronic origin of subGeV-GeV spectral component, but what remains still uncertain is how the highest energies are produced by leptons or hadrons and how hadrons are accelerated. There is still no spectral signature that would clearly distinguish reaccelerated CRs from particles injected from the thermal pool and accelerated at the shock [16].
Therefore, despite the extensive experimental and theoretical studies over the last several decades, the SNR paradigm of the origin of Galactic CRs at the highest energies is still to be confirmed. Indeed, all SNRs seen at TeV energies show cut-off in their VHE gamma-ray spectra below few tens of TeV and thus, currently, it is still unclear whether or not SNRs can actually act as CR PeVatrons.

The quest for PeVatrons requires high sensitivity at $10 - 100\,\mathrm{TeV}$ and can be carried out in two main ways:

- **Source class approach**: select a source (e.g. the GC ridge region) or a source class (e.g. SNRs) and investigate the feasibility of such class as contributor to the cosmic-ray flux;

- **Survey approach**: look for powerful gamma-ray emitters in the multi TeV range and investigate how these different particles factories contribute to the spectrum of cosmic rays (the contribution of different sources or source classes should be different).

The dataset provided by the future CTAO GPS will be used to look for VHE gamma-ray sources which are PeVatron candidates. Their best fit positions and extensions are expected to be spatially coincident with dense gas regions, which can provide target gas material for accelerated CRs to interact with via proton-proton interactions followed by a subsequent π^0 decay. Deeper pointed follow-up observations of promising individual candidates will be required to confirm and further characterize the nature of PeVatrons. Additionally, it is also suggested to search for diffuse gamma-ray emission in the vicinity ($\lesssim 100\,\mathrm{pc}$) of prominent gamma-ray bright SNRs. This is motivated by the belief that PeV particles are expected to quickly escape the SNR shock [14, 33] and then propagate diffusively in the ambient medium surrounding the remnant; the interaction of such runaway PeV particles with the ambient gas produce gamma rays with a characteristic hard spectrum extending up to $\sim 100\,\mathrm{TeV}$ [29].

Although supernova remnants remain prime candidates as suppliers of Galactic CRs, massive stars with powerful winds have been proposed as a viable alternative to supernova remnants [17] primarily as contributors to the knee region around 1 PeV. In fact, the conversion of few % of a massive star cluster (MSC) kinetic energy into accelerated particles can easily account for the gamma-ray emission observed in coincidence with some MSCs, such as Cygnus OB2 [9, 12] and Westerlund 1 [7], in a scenario where gamma-rays would result from pions produced in collisions between the relativistic hadrons and the ambient gas. In addition, the gamma-ray spatial profile suggests that CRs are continuously produced during a period compatible with the typical age of a MSC.

A preference for young massive star clusters as proton PeVatrons over supernova remnants has recently been argued in the context of the $1/r$-type (where r is the distance from the cluster) spatial distributions of parent protons, derived from the observations of extended TeV gamma-ray sources associated with luminous stellar clusters, in particular with Cygnus OB2 [10]. In this scenario, among the recently announced new sources observed by the partially completed KM2A of LHAASO, the positional coincidence of LHAASO J2032+4102 with the Cygnus Cocoon that surrounds Cygnus OB2, and with photons exceeding 1 PeV emitted from it, can be treated as evidence of the operation of massive stars as hadronic PeVatrons, although the emission is also compatible with a leptonic model of the Klein-Nishina tail.

Regardless of the nature of objects associated with the UHE sources, the photons detected by LHAASO in less than one year of operation has revealed many hot spots as clusters of gamma rays in specific directions of the sky, with the detection of more than 530 photons at energies above 100 TeV and up to 1.4 PeV from 12 UHE gamma-ray sources with a statistical significance larger than 7σ. This proves the existence of Galactic PeVatrons although the only identified source is the Crab Nebula with an energy spectrum reaching 1 PeV which provides also the first model-independent evidence that it operates as an electron PeVatron [18].

Despite these recent results, a facility like CTAO is still needed because of its unrivalled angular and energetic resolution and flux sensitivity. Detection prospects comparing multiple predictive models with known gamma-ray sources, including the ultra high-energy sources recently detected by LHAASO, have been already performed [33].

Therefore, understanding the capabilities of CTA to detect those sources, with related confusion, is of extraordinary importance for the Galactic science cases. Source confusion resulted to be the main problem in the DC-1 performed, with multiple overlapping simulated sources that leads to incorrect source modeling in the resulting catalog, and thus no suitable association. As a consequence, assessing how source confusion will impact on the future CTAO observations of detected sources needs to be investigated.

Chapter 4

Simulations of VHE sources

Unidentified sources can be associated with multiple objects so that one is not able to pinpoint the precise origin of the VHE emission because of the source confusion induced by the relatively broad PSF of VHE gamma-ray instruments compared to those at other wavelengths. Moreover, multiple overlapping sources can be detected as a single source. In order to shed light on the capabilities of CTAO for the GPS, the goal of this thesis is to simulate and, subsequently, analyse in a systematic way different kinds of complex regions to obtain a firm assessment on the impact of source confusion in resolving sources along the Galactic plane. Therefore, the full work is based on full-sky simulations performed by using *Gammapy*. This is the official CTAO Science Analysis tool package that in this thesis work is used to both simulate the gamma-ray sky as expected to be seen by CTAO as well as to analyse the simulated data.

At first, simulations consider the easiest case: the modelization and subsequent chacacterization of two nearby sources. The source model consists of a spatial template and spectral template: for the first one, two types have been considered, either a point-like model or a 2-D Gaussian shape for the extended sources; concerning the spectral model, a power-law emission has been considered. The parameter space varies according to the values assumed by photon index, flux normalization, distance between the sources, etc. so as to cover all possible relevant scientific cases. All the models are described in details in Section 4.3.

These two sources are simulated considering only the instrument background level that results from the information contained in the CTAO instrument response function (IRF, see below Section 4.1). It was decided to neglect the diffuse emission from the in-

terstellar medium in the simulations to avoid the choice of a particular model among the variety of them which are present in literature. Instead, very extended sources have been considered (e.g. Gaussian spatial models with $\sigma \geq 1°$) in order to reproduce the interstellar emission of two sources by simply simulating one big source. Obviously, it is possible to plan future similar studies including the interstellar emission model (IEM).
As a consequence of the IEM absence, the choice of a pointing direction is meaningless: since the interstellar emission is relevant in the GC and becomes less and less important moving away from it, there will be no difference among the simulations performed in any sky direction.

Moreover, the sky models simulated are convolved with the IRFs corresponding to a zenith angle observation of 20° on-axis and for the Southern site only. With a large array like CTAO the collection area does not essentially change with a particular zenith angle value; the energy threshold, instead, is affected by this choice but for the analysis presented here is not a concern. Consider an optimized angular resolution goes at the expense of the sensitivity: the difference in sensitivity reach up to 20%. Furthermore, the choice of an averaged azimuth does not affect the analysis at this energies.

4.1 Instrument Response Functions

The IRF provide a mathematical description that links the reconstructed photon arrival direction p and energy E of an event to the true incident direction p_{true} and energy E_{true} of a photon. Given a gamma-ray sky flux model $\Phi(p, E)$ (unit: $\mathrm{m^{-2}s^{-1}TeV^{-1}sr^{-1}}$), the expected number of detected events is:

$$N(p, E)\mathrm{d}p\mathrm{d}E = t_{obs} \int_{E_{true}} \mathrm{d}E_{true} \int_{p_{true}} R(p, E|p_{true}, E_{true}) \times \Phi(p_{true}, E_{true})\mathrm{d}p_{true} \tag{4.1}$$

where t_{obs} is the observation time and $R(p, E|p_{true}, E_{true})$ (unit: $\mathrm{m^{-2}TeV^{-1}}$) describes the response of the instrument to an observation by linking the reconstructed quantities to their true values.
For CTA, as well as for the other IACTs such as HESS, VERITAS and MAGIC, IRFs[1] are factorised into the product of three indepent functions:

$$R(p, E|p_{true}, E_{true}) = A_{eff}(p_{true}, E_{true}) \times PSF(p|p_{true}, E_{true}) \times E_{disp}(E|p_{true}, E_{true}) \tag{4.2}$$

---

[1] Publicly available IRFs: https://www.cta-observatory.org/science/cta-performance/

where:

- $A_{eff}(p_{true}, E_{true})$ is the effective collection area of the detector (unit: m^2, see Figure 4.1a). It is the product of the detector collection area times its detection efficiency at true energy E_{true} and position p_{true}.
- $E_{disp}(E|p_{true}, E_{true})$ is the energy dispersion (unit: TeV^{-1}, see Figure 4.1b). It gives the probability to reconstruct the photon at energy E when the true energy is E_{true} and the true position is p_{true}. Gamma-ray instruments consider the probability density of the migration $\mu = \frac{E}{E_{true}}$, i.e. $E_{disp}(\mu|p_{true}, E_{true})$.
- $PSF(p|p_{true}, E_{true})$ is the point spread function (unit: sr^{-1}, see Figure 4.1c). It gives the probability of measuring a direction p when the true direction is p_{true} and the true energy is E_{true}. Gamma-ray instruments consider the probability density of the angular separation between true and reconstructed directions $\delta p = p_{true} - p$, i.e, $PSF(\delta p|p_{true}, E_{true})$.

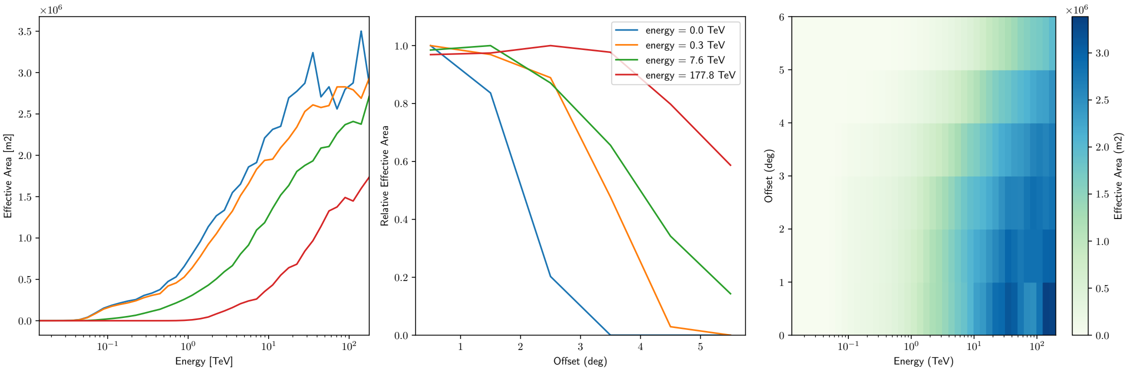
In addition, the IRFs contain the description of the background rates as a function of energy and position within the field of view (Figure 4.2). The background rate is mostly composed by cosmic-ray hadrons and electrons that survive the gamma-ray selection criteria.

CTAO performance values are derived from Monte Carlo simulations based on the `CORSIKA`[2] air-shower simulation program and telescope simulation tool `sim_telarray` [15]. `CORSIKA` is a publicly available, open-source code used by all the current IACT arrays, and represents a standard tool in the wider astroparticle physics community.
The simulations performed here uses the latest (at the time of writing) high-level CTAO IRFs (current version: `prod5-v1`) corresponding to an ideal Southern Array of the CTAO Alpha Configuration with good and stable atmospheric and instrumental conditions, gamma selection cuts that have been optimized to maximize the angular resolution[3] and a zenith angle of 20° and exposure time of 50 h.

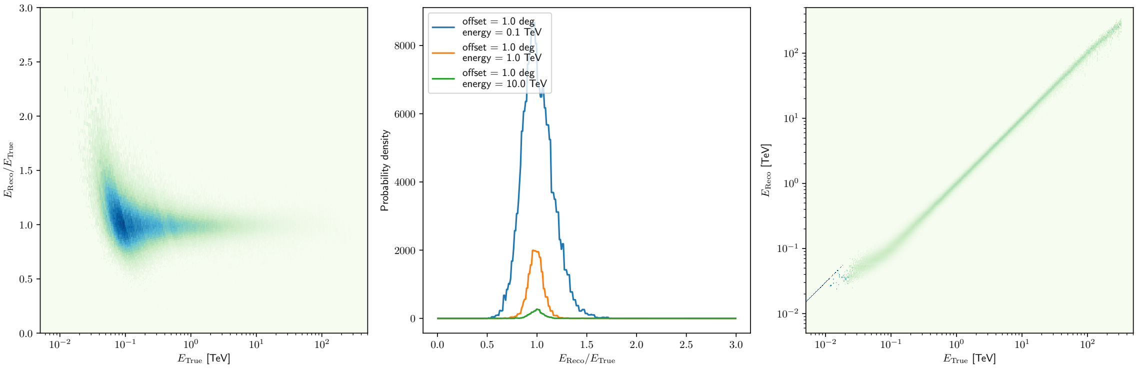
---

[2] https://www.iap.kit.edu/corsika/index.php

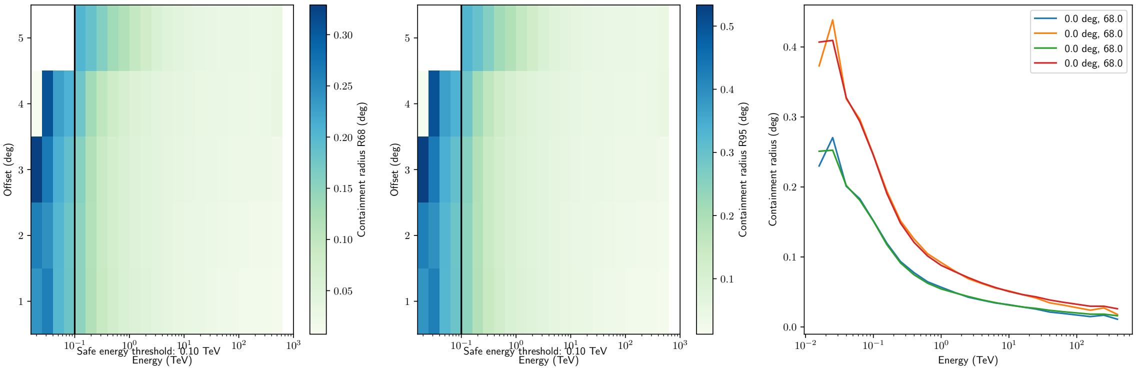
[3] The IRFs are actually optimized to both maximize the angular resolution and minimize the sensitivity, but the angular resolution was the only mentioned because it is the one that counts the most.



(a) Effective area as a function of true energy (left), offset angle (center) and both (right) from the FoV center ($A_{eff}(p_{true}, E_{true})$), following the notation in IRF theory. Colors in left and center panels refer to different values of photon energy.



(b) Probability density of the energy migration $\mu = \frac{E}{E_{true}}$ as a function of true energy and offset angle from the FoV center ($E_{disp}(E_{true}, \mu | p_{true})$), as in the IRF theory (center). Left and right panels show the dependence of the reconstructed energy E_{Reco} on the true energy E_{True} (left) and the E_{Reco} (right) from E_{True}.



(c) The radially symmetric probability density of the angular separation between true and reconstructed directions $\delta p = p_{true} - p$, as a function of true energy and offset angle from the FoV center ($PSF(\delta p, E_{true} | p_{true})$), as in the IRF theory, and for two values of containment radius (68% and 95% of the PSF). As can be seen, above a certain energy threshold, the containment radius does not vary significantly for any offset angle from the FoV center.

Figure 4.1: CTAO IRFs. All the images come from the latest available irfs file, `prod5-v1`, with an optimized analysis regarding the angular resolution, a zenith angle of 20° and an exposure time of 50 h.

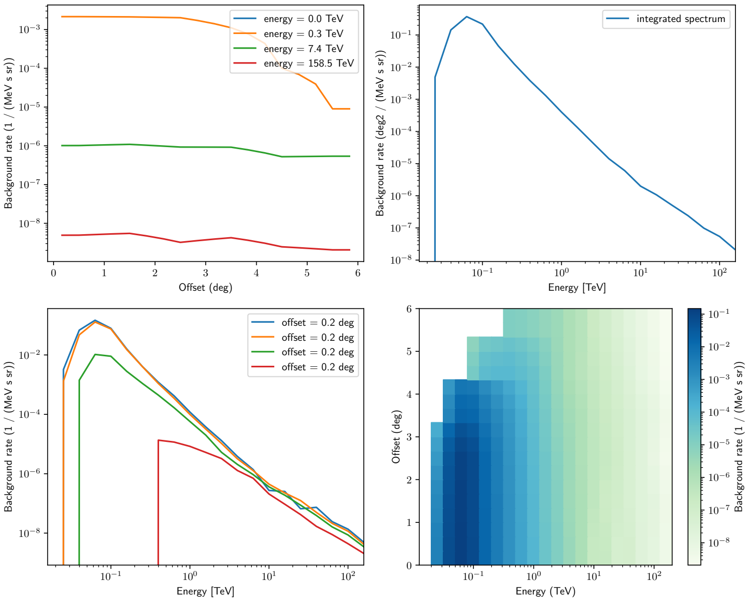


Figure 4.2: Background rate per solid angle as a function of reconstructed energy and detector coordinates. Left, upper panel shows the dependence on the offset for given energy values that can be inferred from the legend; the dependence from the reconstructed energy is shown as an integrated spectrum in the right, upper panel and for given detector coordinates in the left, bottom panel. The bottom, right panel shows the overall dependence from both the variables as an histogram. All the images come from the latest available irfs file, `prod5-v1`, with an optimized analysis regarding the angular resolution, a zenith angle of 20° and an exposure time of 50 h. As can be seen from the right, bottom panel, the background rate is more important at low energies and small offset angle from the FoV center.

4.2 Scientific tools and packages description

Gammapy [26] is an open-source Python package for gamma-ray astronomy built on *Numpy*, *Scipy* and *Astropy*. The CTAO has chosen Gammapy as core library for its Science Analysis tools, but it is also recommended by the HESS collaboration to be used for science publications, and is already widely used in the analysis of other existing gamma-ray instruments, such as MAGIC, VERITAS and HAWC.

This is a very young project organised in sub-packages containing many classes, functions and some global variables needed to analyse gamma-ray event data, but it is still under heavy development, fixing issues and adding new functionality for CTAO.

Gammapy supports binned simulations, i.e. Poisson fluctuation of predicted counts maps or spectra, as well as event sampling to simulate DL3 (data level 3 FITS files) events data, consisting of event lists and corrisponding IRF. Full-sky simulations for different

predefined sky models are simulated and analysed in a sistematic way by using the latest version of the Science Tools, `gammapy-0.18.2`, released November 19, 2020, whose documentation is fully available online [30]. The detailed steps of the simulation study, including source description given by spatial and spectral models, is explained below, while the analysis and the results obtained will be covered by next chapters.
To simulate a specific observation, it is not always necessary to simulate the full photon list. In many cases, simulating directly a reduced binned dataset is sufficient: the IRFs reduced in the correct geometry are combined with a source model to predict an actual number of counts per bin. The latter is then used to simulate a reduced dataset using Poisson probability distribution.

Above all, a brief description of some `gammapy` sub-packages with their Python classes and functions used in the pipeline code is presented.

`gammapy.data` This sub-package is used for data and observation handling. It currently contains the `EventList` class, as well as classes for IACT data and observation handling, but for the aim of this work it was used to create in-memory observations, i.e observations that are saved temporally and not locally, with the `Observation` class.

`gammapy.datasets` This contains classes to handle reduced gamma-ray data for modeling and fitting. The `Dataset` class bundles reduced data, IRFs and model to perform likelihood fitting and joint-likelihood fitting. All datasets contain a `Models` container with one or more `SkyModel` objects that represent additive emission components.
To model and fit data in Gammapy, it is necessary to create a `Datasets` container object with one or multiple `Dataset` objects. Gammapy has built-in support to create and analyse different types of datasets, but only `MapDataset` and `FluxPointsDataset` were used in this thesis work.
The map datasets represent 3D cubes (`WcsNDMap` objects) with two spatial and one energy axis. For 2D images, the same map objects and map datasets are used, with the energy axis having only one energy bin. The `MapDataset` contains a counts map and the fit statistic used is Cash. The `FluxPointsDataset` class contains a set of estimated flux points and a spectral model, and the fit statistic used is χ^2 (a better explaination of both fit statistic, Cash and χ^2, is present in Section 5.1).

`gammapy.makers` Classes that can be used to perform data reduction tasks from DL3 data to binned datasets are contained in this sub-package. All data reduction steps can be combined in a single loop to run a full data reduction chain using the `MapDatasetMaker`. This class makes maps for a single IACT observation, filling the corresponding `counts`, `exposure`, `background`, `psf` and `edisp`.

To exclude a data range from a `MapDataset`, that is associated with high systematics on instrument response functions, Gammapy provides a `SafeMaskMaker` class to compute safe data range masks according to certain criteria. The method used (or a combination of those) can be specified by a keyword according to the chosen criterion (for example a maximum offset cut). The `SafeMaskMaker` does not modify any data, but only defines the `MapDataset.mask_safe` attribute. This means that safe data range can be defined and modified in between the data reduction and fitting.

`gammapy.modeling` This sub-package contains all the functionality related to modeling and fitting data. This includes spectral, spatial and temporal model classes. It is possible to define either via a YAML configuration file, i.e. a file format in which the model can be described according to certain standards readable through a script (see Code Listing 4.1), or via Pyhton code the `Models` to use, which is a list of `SkyModel` objects representing additive emission components, usually sources or diffuse emission, although a single source can also be modelled by multiple components. The `SkyModel` is a factorised model with a `SpectralModel` component and a `SpatialModel` component. The models used in this thesis work will be described in detail in Section 4.3.

The package contains also the `Fit` class that provides methods to fit, i.e. optimise parameters and estimate parameter errors and correlations. It interfaces with a `Datasets` object, which, in turn, is connected to a `Models` object, which has a `Parameters` object, containing the model parameters. Currently `iminuit`[4] is used as modeling and fitting backend.

Code Listing 4.1: Example of a YAML configuration model file.

```
components:
- name: model1
  type: SkyModel
  spectral:
    type: PowerLawSpectralModel
    parameters:
    - name: index
      value: 2.0
    - name: amplitude
```

---

[4] https://iminuit.readthedocs.io/en/latest/about.html

```
          value: 1.0e-12
          unit: cm-2 s-1 TeV-1
       -  name: reference
          value: 1.0
          unit: TeV
          frozen: true
  spatial:
       type: PointSpatialModel
       frame: galactic
       parameters:
       -  name: lon_0
          value: 0.0
          unit: deg
       -  name: lat_0
          value: 0.0
          unit: deg
```

`gammapy.estimators` This submodule contains algorithms and classes for high level flux and significance estimation such as flux maps, flux points, flux profiles and flux light curves. All estimators allow to estimate fluxes in bands of reconstructed energy. The core of any estimator algorithm is hypothesis testing: a reference model or counts excess is tested against a null hypothesis (see Section 5.1).

From this package, `FluxPointsEstimator`, `FluxPoints` and `TSMapEstimator` have been used. The first is a class that estimates flux points for a given list of datasets, energies and spectral models; the amplitude of the reference spectral model is fitted within the energy range defined by the energy group and this is done for each group independently. The amplitude is re-normalized using the "norm" parameter, which specifies the deviation of the flux from the reference model in this energy group. `FluxPoints` is the flux points container returned by this fitting process.

`TSMapEstimators` allows to compute TS (short for test statistic) map from a `MapDataset` by fitting by a single amplitude parameter on each pixel as described in Appendix A of [37]; the fit is simplified by finding roots of the derivative of the fit statistics.

4.3 Outline of the simulations pipeline

The code produced for the thesis work is entirely written in Python 3.7 through the open-source web application Jupyter Notebook with natively implemented functionalities and using `Gammapy` software package utilities. In the following, it is explained how simulations have been performed with all the models that can be reasonably considered for a first-order characterization of the gamma-ray sources and by specifying how the parameters space has been explored.

Fake observations are created with `Observation` that contains the observation duration, pointing information and the CTAO IRFs (that are loaded with `irf.load_cta_irfs`). A reasonable livetime exposure for the CTA GPS is 10 h. Then, it is necessary to define a map geometry for binned simulations and an "empty" `MapDataset` is created with the `MapDatasetMaker` class.
As can be seen in Code Listing 4.2, the energy range covered is from 0.5 to 200 TeV generated with 10 samples. This results in approximately 3 bins per decade.

Code Listing 4.2: Example of setting process to simulate.

```python
#load irf
irf = load_cta_irfs("$PATH_TO_IRF_FILE/irf_file.fits")

# Define the observation parameters
livetime = 10.0 * u.hr #Livetime exposure of the simulated observation
pointing = SkyCoord(5, 0, unit="deg", frame="galactic")
# Create an in-memory observation
obs = Observation.create(pointing=pointing, livetime=livetime, irfs=irf)

# Define map geometry for binned simulation
energy_reco = MapAxis.from_edges(
np.logspace(-0.3, 2.3, 10), unit="TeV", name="energy", interp="log"
)
geom = WcsGeom.create(
skydir=(5, 0), #Sky position of map center
binsz=0.01, #Map pixel size in degrees
width=(5., 5.), #width of the map in deg
frame="galactic",
axes=[energy_reco], #List of non-spatial axes
)

```

```
# Make the MapDataset
empty = MapDataset.create(geom, name="dataset-simu")
maker = MapDatasetMaker(selection=["exposure", "background", "psf", "edisp"])
maker_safe_mask = SafeMaskMaker(methods=["offset-max"], offset_max=2.0 *
    u.deg)

#Make map dataset with MapDataset and Observation
dataset = maker.run(empty, obs)
dataset = maker_safe_mask.run(dataset, obs)
```

Regarding the background, no models were assumed, and the only component used in the simulation consists of an instrumental background model (approximating the residual background from interactions of diffuse cosmic-rays with the atmosphere) defined in the creation of the “empty” dataset necessary for every simulation and comes directly from the IRFs file.

Then spatial and spectral models are defined and, for each source, two models are unified together as a `SkyModel` and collected into a list, `Models`.

4.3.1 Spatial models

Spatial source components are either modelled as point-like or Gaussian-like sources (`PointSpatialModel` and `GaussianSpatialModel`). Regarding the first, center position is provided in a given coordinate frame as input parameter. The Gaussian model instead requires also a value for the length of the major semiaxis (σ, in angular units). However, default values are also provided so that one can use these models with reasonable initial values to fit datasets models.
Examples of these two models definition are shown in Listing 4.3.

Code Listing 4.3: Examples of defining spatial models of pointlike and gaussian sources located in the Galactic Center.

```
spatial_model1 = PointSpatialModel(
lon_0 = "0.0 deg", lat_0 = "0.0 deg", frame = "galactic"
)

spatial_model2 = GaussianSpatialModel(
lon_0 = "0.0 deg", lat_0 = "0.0 deg", frame = "galactic", sigma = "0.2 deg"
)
```

4.3.2 Spectral models

Among all the spectral models available, for this work the source events are simulated following a simple power-law (PL) spectral model given by the formula

$$\Phi(E) = \Phi_0 \cdot \left(\frac{E}{E_0}\right)^{-\gamma} \tag{4.3}$$

where Φ_0 and γ are the flux normalization (at $E_0 = 1\,\mathrm{TeV}$) and the photon index, respectively. As before there are default values, but for simulations purpose these parameters were set for every scientific case considered.
An example on how to define the `PowerLawSpectralModel` is shown in Listing 4.4.

Code Listing 4.4: Examples of defining spectral models for sources with PL and ECPL model.

```
spectral_model1 = PowerLawSpectralModel(
index = 2.0,
amplitude = "1e-12 cm-2 s-1 TeV-1",
reference = "1 TeV"
)
```

For each science case considered, a .yaml file with all the model information was created. These models are assigned to the dataset according to the different science case considered each time and the number of counts for the given model are faked with Poisson fluctuations to produce a simulated counts map.

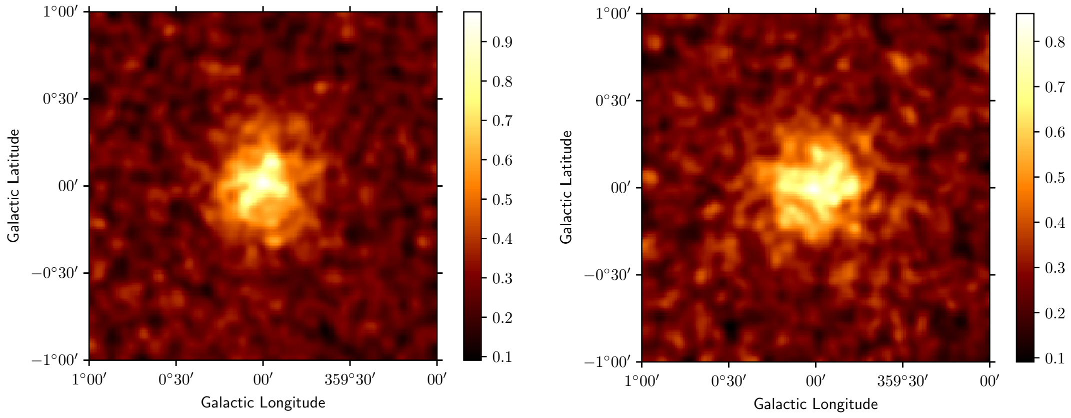
The capability of disentangling two sources clearly depends on the distance between them, which is connected to the angular resolution of the instrument, but also to the flux at a given energy, which translates into a dependence from the photon index and flux normalization. This has been studied starting from the easiest case, i.e. two pointlike sources, then one of the two sources has been considered extended and finally both of them have been modelized with a `GaussianSpatialModel`. A primary source was set to have a fixed position at the center of the FoV, and a fixed spectral model with photon index and flux normalization equals to 2 and $1 \times 10^{-12}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$, respectively. Only when considering very extended sources with $\sigma \geq 1°$, the flux normalization of the primary source at 1 TeV has been changed to $1 \times 10^{-13}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$. Those type of sources are indeed simulated in order to reproduce a kind of realistic diffuse emission, therefore the flux normalization has to be lowered because it would not be plausible to have such extreme fluxes for this kind of background.

For the second source, the performed simulations covered a wide range of the parameters:

- the Galactic longitude of the second simulated source, that is the distance between the two sources, is varied from 0.02° to 0.1°, with a step of 0.02°, when both sources are pointlike; from 0.05° to 0.4° with a step of 0.05° and from 0.1° to 0.8° with a step of 0.1° when it is present only one Gaussian source; from 0.2° to 1.0° with a step of 0.2° and with an additional value of 1.3° when both sources are Gaussian.
- σ length for the Gaussian models is varied covering the following values expressed in degrees: 0.2, 0.3, 0.5, 1.0, 3.0.
- Flux normalization at 1 TeV is varied from 4×10^{-14} to $1 \times 10^{-12}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$, with variable step and with a variable lower boundary depending on the spatial extension of the source.
- Photon index is varied from 1.5 to 2.5, with a step of 0.5.

Then, the simulated datasets can be analysed as all standard ones. To begin, it is possible to have a simple plot of the smoothed counts map and, using interactive widgets, to explore the non spatial axes, to get a first impression of what has been simulated. In Figure 4.3 are shown some zoomed-in map examples with non-spatial axes summed over, while in Figure 4.4 is shown the difference between the presence or not of a very extended source to reproduce a diffuse emission background.

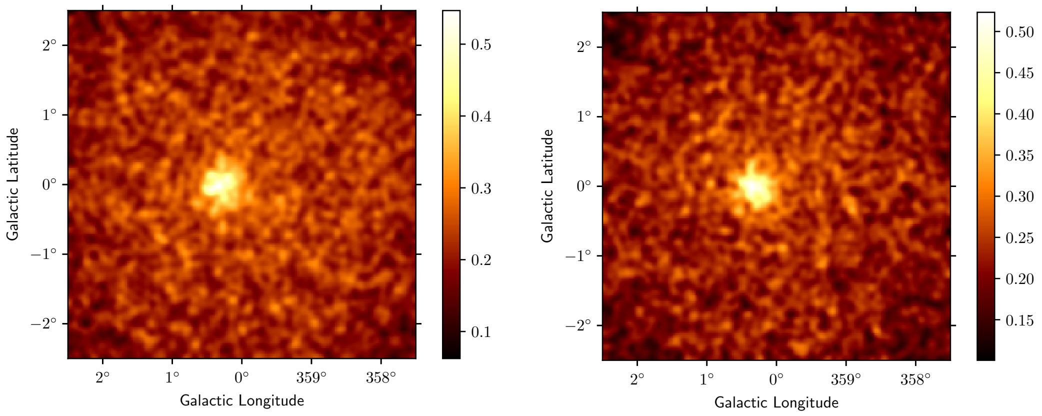
For each flux normalization value, pointlike sources have been simulated for every photon index; for each width of the Gaussian, extended sources have been simulated for every possible flux normalization and every photon index. Note that these values have been assigned to the second source, as the source at the pointing center has always the same spectrum, as already stated, to partially simplify the amount of work. This results in a limited cases study that can be easily extended in future works. Moreover, all these kinds of confused regions have been simulated varying the distance between the two sources, with steps that have been already reported above and are summarized in Table 4.1. All the scientific cases simulated can be found in Appendix A in a schematic way.



(a) (b)

Figure 4.3: Counts map of two science cases considered among all the simulated datasets. In both panels, the two simulated sources are both Gaussian with $\sigma = 0.2°$, with the first source that has a photon index and flux normalization of 2.0 and $1 \times 10^{-12}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$, respectively. The maps are displayed with non-spatial axes summed over. **(a)**: The second source has a photon index and flux normalization equals to 2.0 and $1 \times 10^{-12}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$, respectively, while the distance between the two sources is 0.3°. **(b)**: The second source has a photon index and flux normalization equals to 2.5 and $1 \times 10^{-12}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$, respectively, while the distance between the two sources is 0.3°.
From maps like those presented here, it is difficult to determine if it is present more than one source in the FoV, as the two Gaussian overlap consistently.

During the simulation process, a preliminary analysis of the extreme cases in terms of flux normalization was necessary to set a lower flux limit for pointlike sources. This can been estabilished producing 2-dimensional significance maps and running a peak finder to detect source candidates, as shown in Listing B.1 in Appendix B. The lower limit was then set to be $4 \times 10^{-14}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$. Because of the worsening of the flux sensitivity for extended sources, which is $\propto \sqrt{\sigma_{PSF}^2 + \sigma_{gauss}^2}$, as the Gaussian width increases, also the lower flux limit increases, as can be seen clearly from the Tables in Appendix A. Subsequently a limited range of cases can be tested, as the source cannot be detected significantly ($> 5\sigma$) by CTA below these limits.



(a) (b)

Figure 4.4: **(a)**: Counts map of a science cases considered among all the simulated datasets in which the primary source is modeled through a Gaussian with $\sigma = 1.0°$, as examples of the reproducible background through very extended sources, with photon index and flux normalization of 2.0 and $1 \times 10^{-13}\text{cm}^{-2}\text{s}^{-1}\text{TeV}^{-1}$, respectively. The second source is a Gaussian with photon index and flux normalization equals to 2.0 and $1 \times 10^{-12}\text{cm}^{-2}\text{s}^{-1}\text{TeV}^{-1}$, respectively, while the distance between the two sources is 0.3°. The map is displayed with non-spatial axes summed over. **(b)**: Counts map of the same scientific case in which, instead, the first source is absent: this shows the difference in considering those extended sources as a model of background. It can be seen that this second map has slightly less background noise, compared to the first one.

Models		**Parameters**	**Boundaries**	**Step**
Spatial	Pointlike	Longitude	0.02° – 0.1°	0.05°
	Gaussian	Longitude	0.05° – 0.4°	0.05°
			0.1° – 0.8°	0.1°
			0.2° – 1.0°	0.2°
		Sigma	0.2° – 3.0°	variable
Spectral	Power-Law	Index	1.5 – 2.5	0.5
		Flux Normalization	$4 \times 10^{-14} - 1 \times 10^{-12}$ $\text{cm}^{-2}\text{s}^{-1}\text{TeV}^{-1}$	variable

Table 4.1: Summarizing table of models adopted to perform simulations along with boundary and step values assigned to the parameters related to the second source.

Chapter 5

Analysis of simulations

After simulating all the cases of interest, the subsequent analysis is presented here, preceded by some notions of statistic. It is of particular importance the description of a method to distinguish if one or more sources are detected. Each simulated case concerns always two overlapping sources, so that there is some level of source confusion, and then with an increasing distance it is possible to see whether the sources are resolved distinctly by CTAO and at which distance value.

It is worth noting that each test case has been simulated and subsequently fitted with an appropriate model 100 times in order to obtain an adequate statistic; the estimate of the parameter values and their errors is indeed performed considering the Gaussian distribution of those coming from the 100 fits, since a single fit does not provide relevant statistical information. Note that the fit is 3-D as the spatial and spectral features are fitted simultaneously.

The pipeline code is partially reported in this Chapter to provide further insights on how the analysis was performed and to stress which `gammapy` sub-packages, functions and classes were adopted.

5.1 Statistics

The Gammapy routine that performs the statistical algorithms used for fitting datasets and to estimate parameter values will be described here.

`gammapy.stats` contains statistical estimators, fit statistics and algorithms commonly used in gamma-ray astronomy. It has to deal with events counts in a given time window with statistical Poissonian errors on the counts.

Gamma-ray measurements are counts containing both signal and background events. The significance of a signal over the background is estimated through Poisson likelihood functions, the fit statistics. In Gammapy, they are all log-likelihood functions normalized like chi-squares, i.e. if L is the likelihood function adopted, they follow the expression $2 \times \log L$ [30].

As default, Gammapy uses `Cash` statistics which is equivalent to Chi-Square statitistics for counts/bin larger than 10. The number of counts, n, is a Poisson random variable of mean value $\mu_{sig} + \mu_{bkg}$. The former is the expected number of counts from the source (the signal), the latter is the number of expected background counts. Writing the likelihood L and applying the expression above, $2 \times \log L$, the following formula for the Cash fit statistic is obtained:

$$C = 2 \times \left(\mu_{sig} + \mu_{bkg} - n \times \log\left(\mu_{sig} + \mu_{bkg}\right)\right) \tag{5.1}$$

The Cash statistic is implemented in `cash` and is used as a `stat` function by the `MapDataset`. This statistic function is at the heart of the model fitting approach in Gammapy. It is used to estimate the best fit values of model parameters and their associated confidence intervals.

Estimating TS The probability of the collected data during a given observation to be drawn from a particular model M is usually described by the maximum likelihood function $L_i(M)$. A classical approach to modeling and fitting relies on the hypothesis testing. One wants to estimate whether an hypothesis H_1 is statistically preferred over the reference, or null-hypothesis, H_0. The maximum log-likelihood ratio test [31] provides a way to estimate the *p-value* , i.e. the probability of obtaining test results at least as extreme as the results actually observed (under the assumption that the null hypothesis is correct), of the data following H_1 rather than H_0, when the two hypotheses are nested. The ratio is noted as:

$$\lambda = \frac{\max L(X|H_1)}{\max L(X|H_0)}$$

The Wilks theorem [42] shows that, under some hypotheses described in detail in the Paragraph below, $2 \log \lambda$ asymptotically follows a χ^2 distribution with n_{dof} degrees of freedom, where n_{dof} is the difference between the free parameters of H_1 vs. H_0. Using this definition, the fit statistics $-2\log\lambda$ is simply the difference of the fit statistic values for the two hypotheses, usually referred to as ΔTS. Hence, ΔTS follows the χ^2 distribution with n_{dof} degrees of freedom [30]. This can be used to convert ΔTS into a "classical significance" with the implementation shown in Code Listing 5.1.

Code Listing 5.1: Converting ΔTS into a classical significance σ

```python
from scipy.stats import chi2, norm

def ts_to_sigma(ts, df=1):
"""Convert delta ts to sigma"""
p_value = chi2.sf(ts, df=df)
return norm.isf(0.5 * p_value)
```

As an example, if only one degree of freedom is considered, the TS value relates to a Gaussian significance and one can estimate the statistical significance in terms of number of σ as $\sqrt{\Delta TS}$. In case the excess ΔTS were negative, which can happen if the background is overestimated, the following convention would be used:

$$\sqrt{\Delta TS} = \begin{cases} -\sqrt{\Delta TS} & \text{if excess} < 0 \\ \sqrt{\Delta TS} & \text{else} \end{cases}$$

Wilks' theorem For completeness, the Wilks' theorem is here illustrated as explained in [42]:

A population K is assumed in which a variate x has distribution function $f(x, \theta_1, \theta_2, \ldots, \theta_h)$, *which depends on the parameters* $\theta_1, \theta_2, \ldots, \theta_h$. *A* ***simple hypothesis*** *is one in which the* θ_i *have specified values. A set* Ω *of admissible hypotheses is considered one in which is consistent with a set of simple hypotheses. Geometrically,* Ω *may be represented as a region in the h-dimensional space of* θ_i. *A set of* ω *of simple hypotheses is specified by taking all simple hypotheses of the set* Ω *for which* $\theta_i = \theta_{0,i}$, *where* $i = m+1, m+2, \ldots, h$.
A random sample O_n *of n individuals drawn from K is considered.* O_n *may be geometrically represented as a point in an n-dimensional space of the x's. The probability density function associated with* O_n *is:*

$$P = \prod_{\alpha=1}^{n} f(x_\alpha, \theta_1, \theta_2, \ldots, \theta_h) \tag{5.2}$$

Let $P_\Omega(O_n)$ *be the least upper bound of P for the simple hypotheses in* Ω *and* $P_\omega(O_n)$ *the least* *upper bound of P for those in ω. Then*

$$\lambda = \frac{P_\omega(O_n)}{P_\Omega(O_n)} \tag{5.3}$$

is defined as the likelihood ratio for testing the composite hypothesis H that O_n comes from a population with a distribution characterized by values of the θ_i for some simple hypothesis in the set ω. When H is true, then O_n is from some population of the set described above.
It is often useful when dealing with large samples to have an approximation to the distribution of λ. Consider an approximation for those cases in which ***optimum*** *estimates of θ_i exist, that is consider the existence of functions $\tilde{\theta}_i(x_1, \ldots, x_n)$ (maximum likelihood estimates of the θ_i) such that their distribution is:*

$$\frac{|c_{ij}|^{1/2}}{(2\pi)^{h/2}} e^{-1/2\sum_{i,j=1}^{h} c_{ij} z_i z_i} (1+\phi) dz_1, \ldots, dz_h \tag{5.4}$$

where $z_i = (\tilde{\theta}_i - \theta_i)\sqrt{n}$, $c_{ij} = -E\left(\frac{\partial^2 \log f}{\partial\theta_i \partial\theta_j}\right)$, with E denoting mathematical expectation, ϕ is of the order of $1/\sqrt{n}$ and $|c_{ij}|$ is positive definite.

Given this, Wilks' theorem enunciates is:

Theorem. If a population with a variate x is distributed according to the probability function $f(x, \theta_1, \theta_2, \ldots, \theta_h)$ such that optimum estimates $\tilde{\theta}_i$ of the θ_i exist and are distributed in large samples according to (5.4), then when the hypothesis H is true that $\theta_i = \theta_{0,i}$, $i = m+1, m+2, \ldots h$, the distribution of $-2\log\lambda$ (where λ is given by (5.3)) is, except for terms of order $1/\sqrt{n}$, distributed like χ^2 with $h - m$ degrees of freedom.

5.2 Fitting pipeline

The spatial and spectral analysis is performed with 3-D fitting using the `Fit` class twice for each simulated dataset: first, the fit assumes that only a single source is present, either pointlike or extended, then the process is repeated assuming the presence of two sources, as they have been actually simulated. This procedure allows us to compare the TS of the two fitting results and evaluate if CTAO is able to significantly detect the two sources or not at increasing distances between them in all the possible confused regions that have been simulated.
It is worth noting that the 3-D fitting procedure is very time-consuming and, for this reason, the scientific tested cases were limited.

To provide adequate statistics, it is necessary to fit the same dataset numerous times, as already stated; however, failed fits are not uncommon and subsequently it is necessary to fit a sufficient high number of times to obtain adequate statistics even if the failed fits are excluded from the analysis. Therefore, the fitting process is performed with a loop cycling over each dataset, in order to fit 100 times the same simulated source with the same model (therefore, at the end of the process there are 200 fitting results because of the two models for each simulated dataset).

An example of the fitting process with only one dataset concerning an extended source with a pointlike is shown in Code Listing 5.2. Inside the loop, the `MapDataset` is at first simulated, according to the model previously assigned, then is fitted with a single Gaussian spatial model and with two Gaussian spatial models (for all the sources considered, the spectral model is always a PL emission model). The Gaussian model has been replaced by the pointlike spatial model only with datasets in which the sources have been simulated as pointlike.

Code Listing 5.2: Example of the fitting process applied to a single dataset with a previously assigned model consisting of Gaussian and pointlike sources. The loop is performed to obtain adequate statistic and properly estimate parameter values and errors.

```
for m in range (0, 100):
    dataset.fake()

    ###     Fitting with one Gaussian
    #Copy the original dataset to preserve it
    mapdata = dataset.copy(name=namefits)
    mapdata.models = dataset.models[0].copy()

    fit = Fit([mapdata])
    result = fit.run()
    # checking if the fit has run correctly
    if result.message != ’Optimization terminated successfully.’:
        fail1.append(’failed’)
    else: fail1.append(’not’)
    # saving results
    result_1gauss = result_1gauss.append(result.parameters.to_dict())
    stat1.append(result.total_stat)

```

```
    ###  Fitting with two Gaussian
    #Copy the original dataset to preserve it
    mapdata = dataset.copy(name=namefits)
    mapdata.models = Models([dataset.models[0].copy(),
        dataset.models[0].copy()])

    fit = Fit([mapdata])
    result = fit.run()
    # checking if the fit has run correctly
    if result.message != ’Optimization terminated successfully.’:
        fail2.append(’failed’)
    else: fail2.append(’not’)
    # saving results
    result_2gauss = result_2gauss.append(result.parameters.to_dict())
    stat2.append(result.total_stat)

    # saving results locally
    result_1gauss.to_csv("result_1gauss_"+namefits+".csv")
    result_2gauss.to_csv("result_2gauss_"+namefits+".csv")

###     Estimating TS and hypothesis testing
ndof = 5
source_detected = ([])
cases = len(stat1)

# Here I compare every couple of fit, even if one has failed
# (in the analysis of the results, I will then keep only the valid results)

for l in range(0, cases):
    ts = stat1[l]- stat2[l]
    sigma = ts_to_sigma(ts, ndof)
    if sigma > 5:
    source_detected.append(2)
    else: source_detected.append(1)

# saving statistical results of each fit performed
results_stat= pd.DataFrame()
results_stat["total_stat1"]= stat1
results_stat["total_stat2"]= stat2
results_stat["source_detected"] = source_detected
```

```
results_stat["failed_1"]= fail1
results_stat["failed_2"]= fail2

# saving statistical results locally
results_stat.to_csv("stat_gausspoint_"+namefits+".csv")
```

Once a fitting loop is over, it is necessary to save locally in .csv files all fit results in order to analyse them at a later stage. Moreover, it is of fundamental importance to check if the fit has converged successfully for the two assumed models, that are the null hypothesis with one source and the other hypothesis with two of them, otherwise no comparison and hypothesis testing is possible. Considering only successfully fits, ΔTS was computed for them and Code Listing 5.1 was adopted to convert ΔTS in classical significance σ.

Directly from a single fitting process, it is possible to obtain a sky map of the excess counts (with respect to the background) and, more interestingly, the fitted spectrum with its residuals for a given dataset, using the `TSMapEstimator` and `FluxPointsEstimator` classes, respectively. Some of the resulting spectra will be presented in the next Chapter.

Chapter 6

Results

When all the data related to the scientific cases are collected, it is possible to analyse them. Here, results and summarizing plots, coming from the post-fitting analysis performed after the fitting pipeline explained above, will be presented in three Sections, paying particular attention to the source confusion problem. First, the focus is on showing when two sources have been detected or not, with detection maps as confidence level that will be shown in almost every scientific case. Then, the different representation of data and how those are eventually affected by the source confusion problem will be presented, depending on how many sources have been detected.

Numerous plots are produced to show results in different ways, therefore only some of them are shown in this Chapter, in order to increase the focus on the main results. In particular, the majority of the plots related to how the spectral and spatial parameters were estimated by the fits considering the model with a single source and the model with two sources are shown in Appendix D, along with further explanations distinguishing cases in which sources have been simulated (both pointlike, extended and pointlike and both extended).

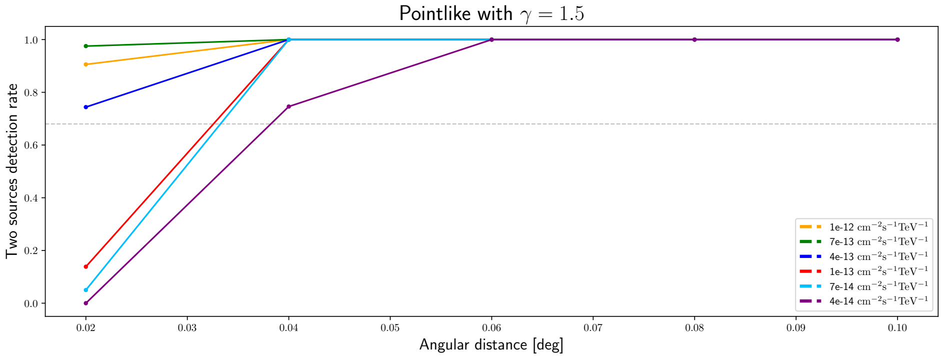
6.1 Two source detection power

The pipeline produced for each science case 100 fits with the one source model and 100 fits with the two sources model. For each pair of fits we calculated the two sources detection significance ΔTS. These values define the significance of the hypothesis that two sources are present over the null hypothesis (one source). The threshold value was set as $\Delta TS = 25$. We computed the two sources detection rate, i.e. a percentage of fits pairs in which the $\Delta TS > 25$, among all the successfully converged fits. When larger than 68%, this percentage (or rate) provides a confidence level of the measurements and allows us to produce the detection maps of the two sources detection power. It is worth noting that some parameter combinations never reach a sufficient number ($\gtrsim 20$) of fits with a "successful" result; therefore, the high number of failed fits implies an exclusion of those particular scientific cases from the comprehensive analysis.

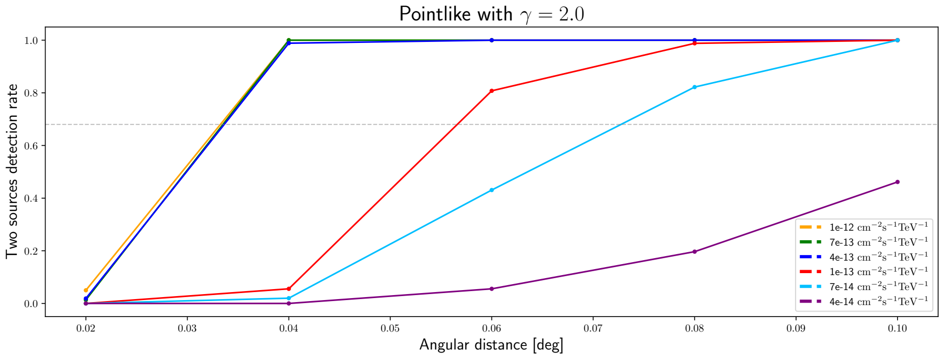
Figures 6.1, 6.2, 6.3, 6.4 and 6.5 show the two sources detection rate for all the considered scientific cases, except for the simulations concerning a Gaussian with $\sigma = 1.0°$ at the center of the FoV and a Gaussian with $\sigma = 0.2°$, because in these scientific cases the threshold was never exceeded (see Section 6.3).
It is possible to say that, when the sources are both pointlike, increasing the distance between the two targets quickly leads to the capability of distinguishing them, for each pair of spectral parameters, except for the cases where lower flux limits are reached and thus compromising the results. Figure 6.1 shows in three panels, depending on the photon index, the trend of the two sources detection rate for different flux normalization values as a function of the distance between the sources. The threshold, displayed through the dashed grey line in the plots, is nearly always exceeded, but it is necessary to have larger angular distance values when the photon index increases. Therefore, harder spectra with lower flux normalization are more difficult to be distinguished when compared to a source with a flat spectrum and flux normalization similar to that of the Crab nebula.

Note that, for some couples of spectral parameters, it is more difficult to distinguish the two sources, and the rate is never above the threshold of 68%: this is due to the lower limit in flux that is actually higher and slightly higher than $4 \times 10^{-14} \mathrm{cm}^{-2} \mathrm{s}^{-1} \mathrm{TeV}^{-1}$ for $\gamma = 2.5$ and $\gamma = 2.0$, respectively. Further information about the estimated parameters when both sources were considered pointlike are presented in Appendix D, Section D.1.

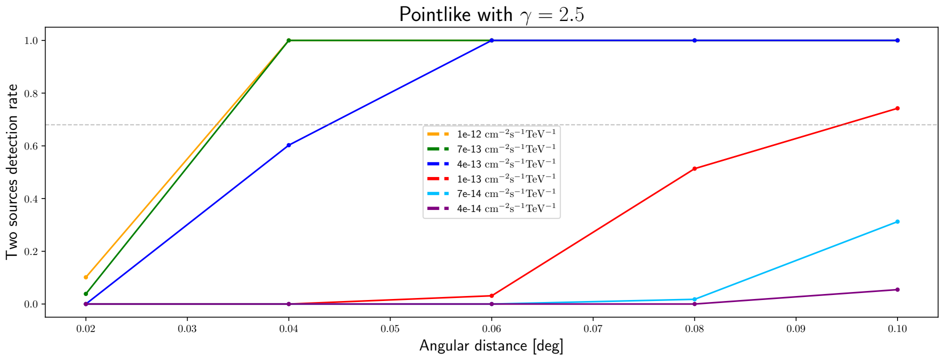
When an extended source is taken into account, it will be more challenging to both obtain successfully converged fits and identify the two sources.
Figure 6.2 shows the scientific cases concerning a Gaussian with $\sigma = 0.2°$ and a pointlike: 87 out of 96 cases had enough converged fits to be considered. From those plots, it can be seen that increasing the distance between the two targets does not correspond to a gradual improvement in the capability of distinguishing them, for each pair of spectral parameters. In fact, the threshold is always exceeded at each distance value in Figure 6.2a, except in the case of the lowest flux normalization value for which it is necessary to have larger distances to distinguish the sources. Furthermore, it is evident from Fig-



(a)



(b)



(c)

Figure 6.1: Trend of the capability of distinguishing sources at different flux normalization values as a function of the distance between the pointlike sources. The three panels refer to photon index 1.5 (a), 2.0 (b) and 2.5 (c). The grey dashed line shows the threshold above which the sources are considered distinguished. Note that for most cases the threshold is always overcome at some longitude value that increases when the photon index γ is higher and the flux normalization is lower. Harder spectra need larger distance values to go above the threshold.

ure 6.2b and 6.2c that more difficulties arise when the distance between the sources has values around the σ of the Gaussian model and that harder spectra with lower flux normalization can not be distinguished at any distance value.

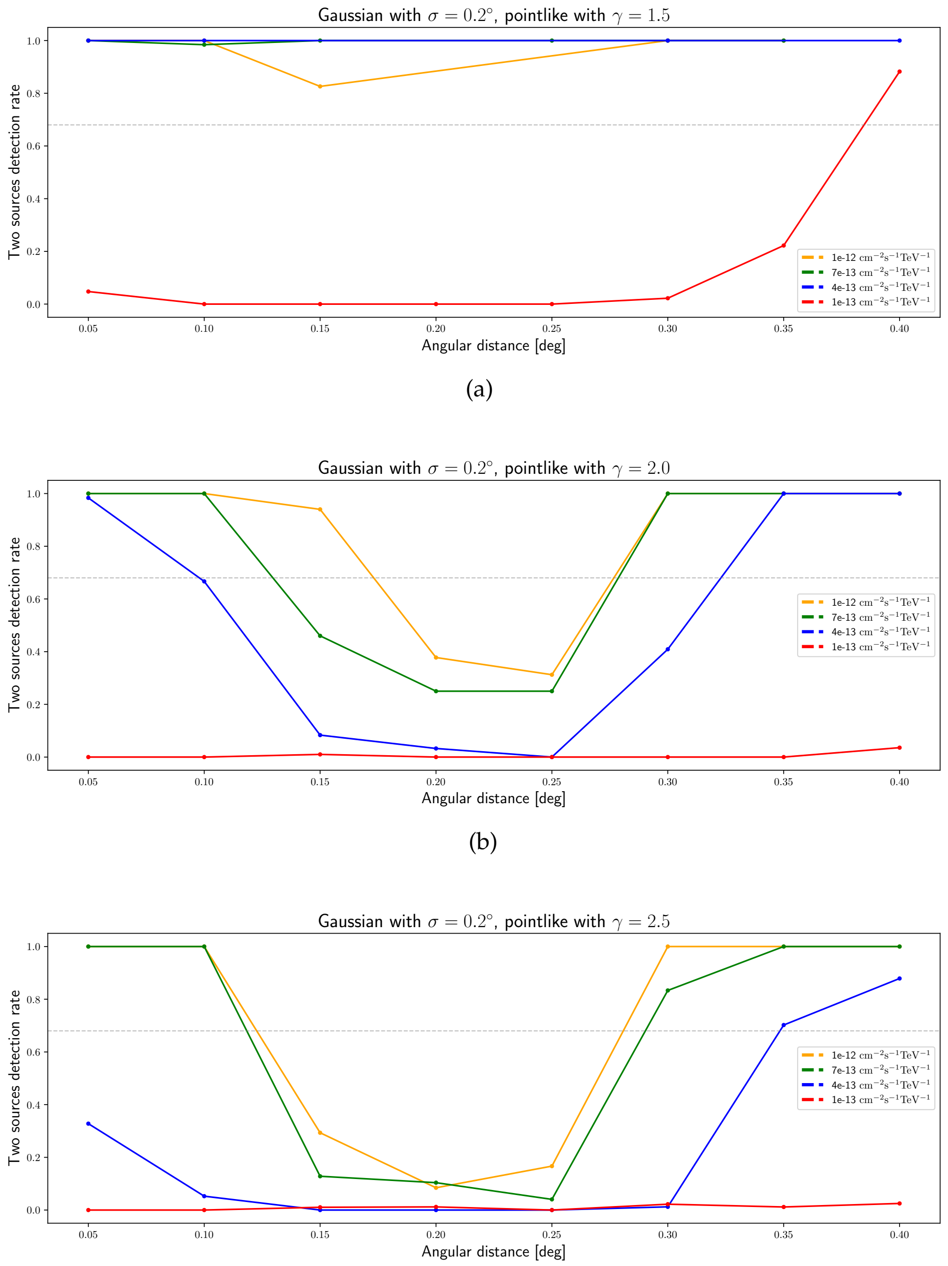
This trend in the percentage of distinguished sources is present also when the Gaussian has $\sigma = 0.3°$ (Figure 6.3), for which 83 out of 96 scientific cases were considered. When $\sigma = 0.5°$, more difficulties arise when the photon index is $\gamma = 2.0$ and there are even less converged fits (63 out of 96 cases with sufficient converged fits were analysed); therefore, this trend may be present, but is not clearly visible as for the other σ values since the majority of the scientific cases with angular distance values around 0.5° are not reported because of the failed fits (Figure 6.4).
Additional plots related to the spectral and spatial information coming from the fits with a Gaussian and a pointlike are displayed in Appedix D, Section D.2.

Regarding the confused region in which two Gaussian sources, both with $\sigma = 0.2°$, have been simulated, plots in Figure 6.5 were produced. Here, the results are similar to the two pointlike cases: increasing the distance between the two targets quickly leads to the capability of distinguishing them. However, the scale of distances is very different, as for two extended sources larger distance are necessary to have a confidence level of the detection rate higher than 68%. Moreover, it is possible to detect the two sources as separate with lower flux limits that are higher than for the pointlike. Further plots and information can be found in Appendix D, Section D.3.

6.2 The case of two sources detected

From each dataset with a number of fits terminated successfully $\gtrsim 20$, spatial and spectral parameter values and errors are extracted: from the package `scipy.stats`, it is possible to consider their distributions of values and errors as Gaussian and to adopt the `norm` class. Then, the mean from the first and the standard deviation from the second have been derived in order to obtain the measurements of a certain parameter. This implies, in cases where the number of fit terminated successfully is low ($20 - 40$), that the parameter values estimated will end up with larger errorbars due to the lower statistics available.

When two sources are resolved, it is necessary to check if spatial and spectral fitted pa-

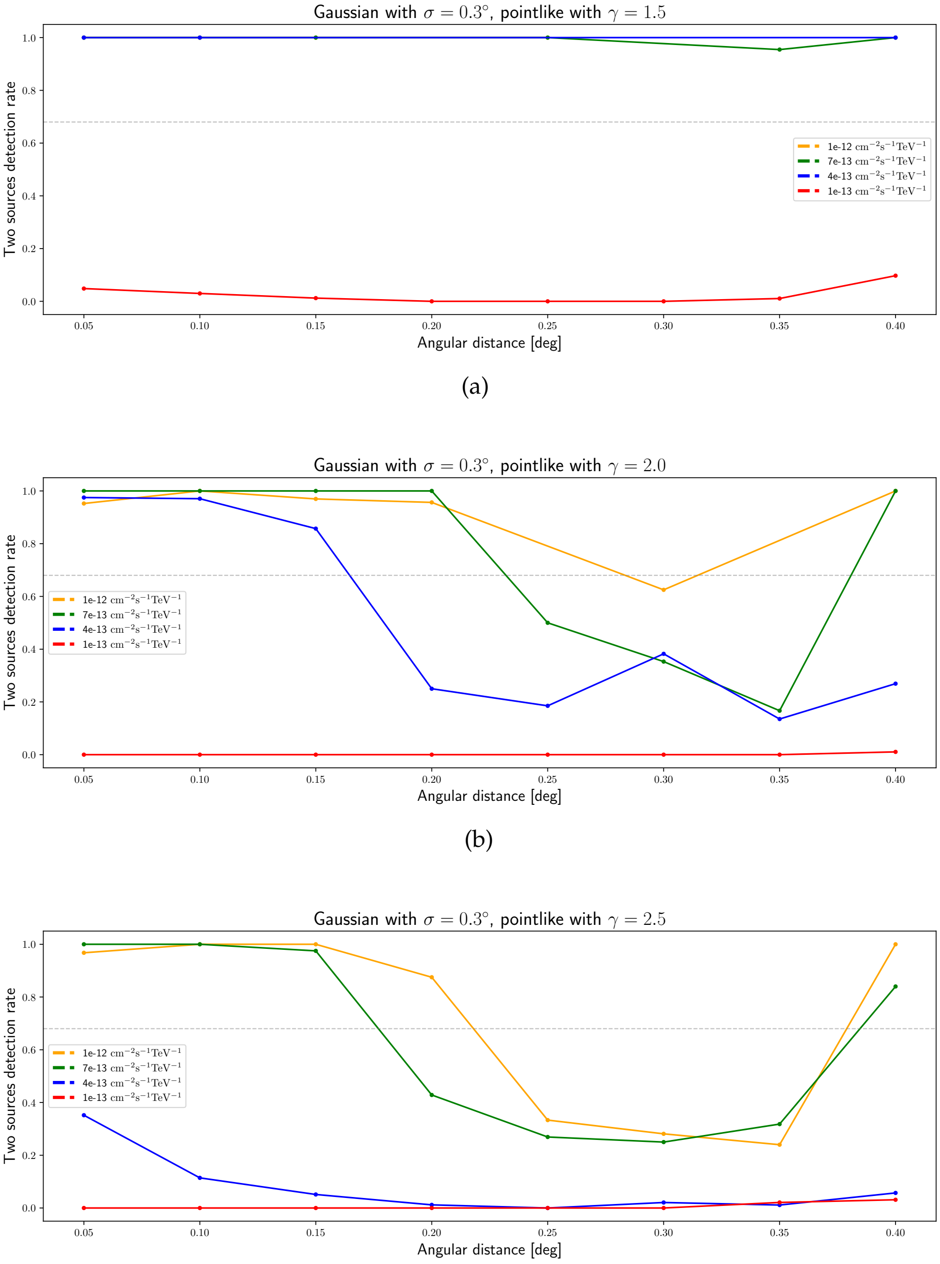


(a)

(b)

(c)

Figure 6.2: Trend of the capability of distinguishing sources at different flux normalization values as a function of the distance between the sources under the assumption of a Gaussian source with $\sigma = 0.2°$ in the center of the FoV and a pointlike source. It is clearly visible the falldown of the rate when the distance value is around 0.2° for photon index equal to 2.0 and 2.5; thus, even if the two sources are distinguished at lower and higher distance values (for most of the flux normalization values), there is a range in which the distinction is more problematic.

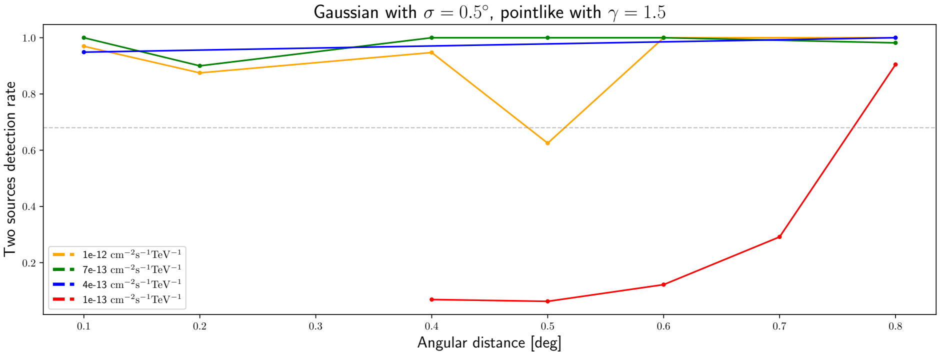


(a)

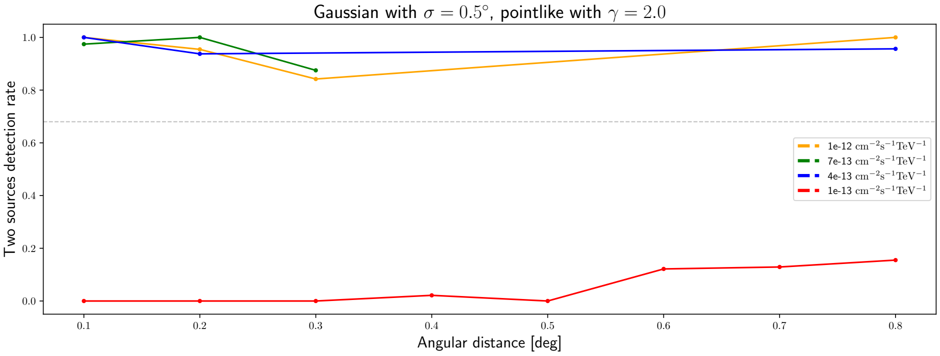
(b)

(c)

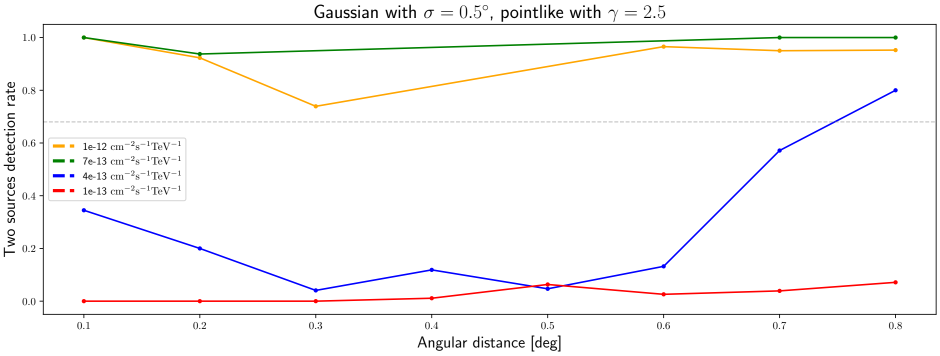
Figure 6.3: Trend of the capability of distinguishing sources at different flux normalization values as a function of the distances between the sources under the assumption of a Gaussian source with $\sigma = 0.3°$ in the center of the FoV and a pointlike source. Also here, it is clearly visible the trend for which this rate has a falldown when the distance values is around 0.3° for photon index equal to 2.0 and 2.5. Moreover, it can be seen that for the lower values of flux normalization, especially for harder spectra, the rate is almost always below the threshold.



(a)

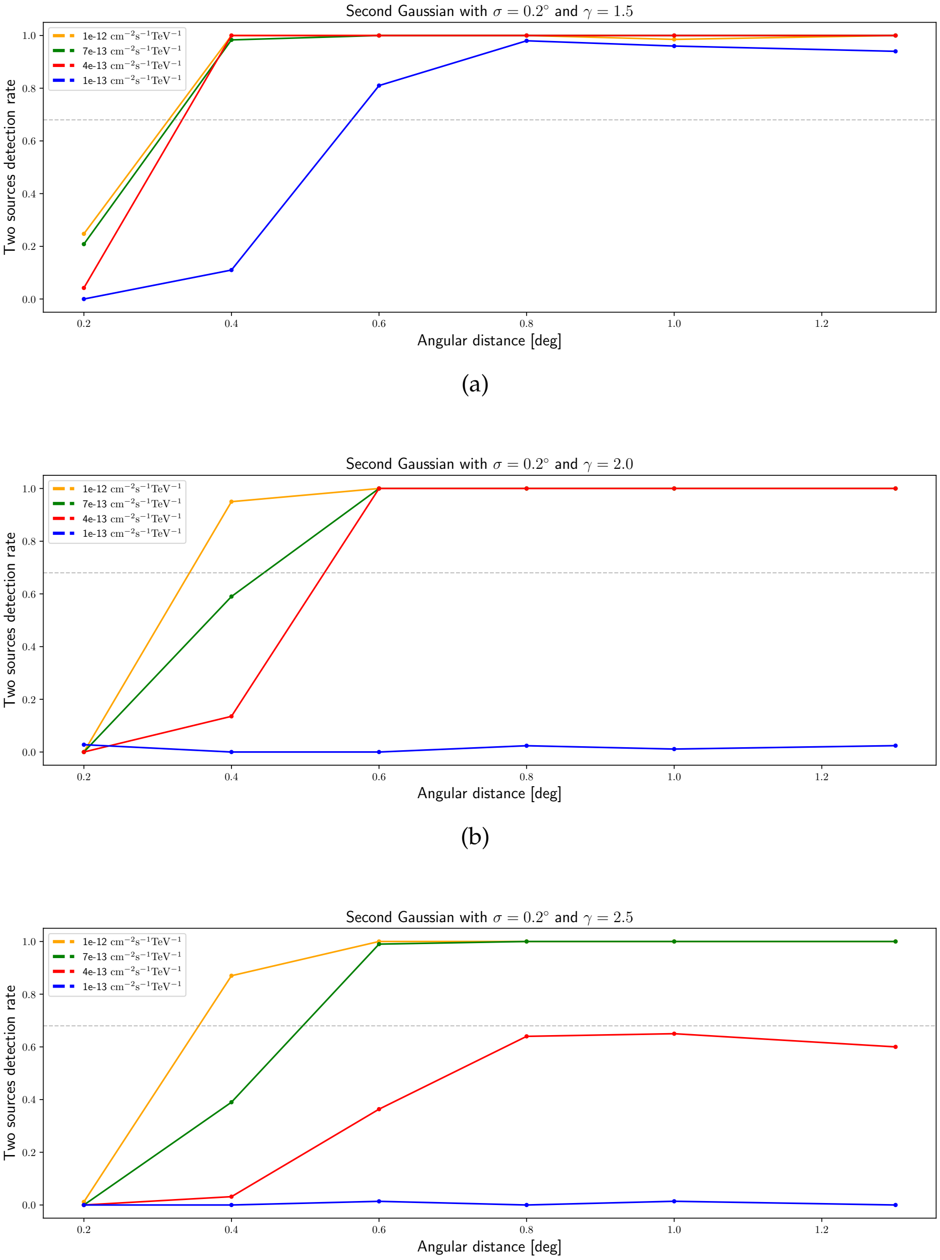


(b)



(c)

Figure 6.4: Trend of the capability of distinguishing sources at different flux normalization values as a function of the distance between the sources under the assumption of a Gaussian source with $\sigma = 0.5°$ in the center of the FoV and a pointlike source. It is still visible for some flux normalization and photon index the trend for which this rate has a fall for distance values around 0.5°. Note that more problems arise during the fitting process around this value, so that there is not sufficient statistics and it is not possible to have clear information where there are no points in the plots. Moreover, it can be seen that, for the lowest value of flux normalization, the rate is nearly always below the threshold.



(a)

(b)

(c)

Figure 6.5: Trend of the capability of distinguishing sources at different flux normalization values as a function of the distance between the sources for the scientific cases regarding both Gaussian sources with $\sigma = 0.2°$. Note that the threshold is almost always overcome at some distance value that increases when the photon index γ is higher and the flux normalization is lower, as in the pointlike cases but with a differente distance scale.

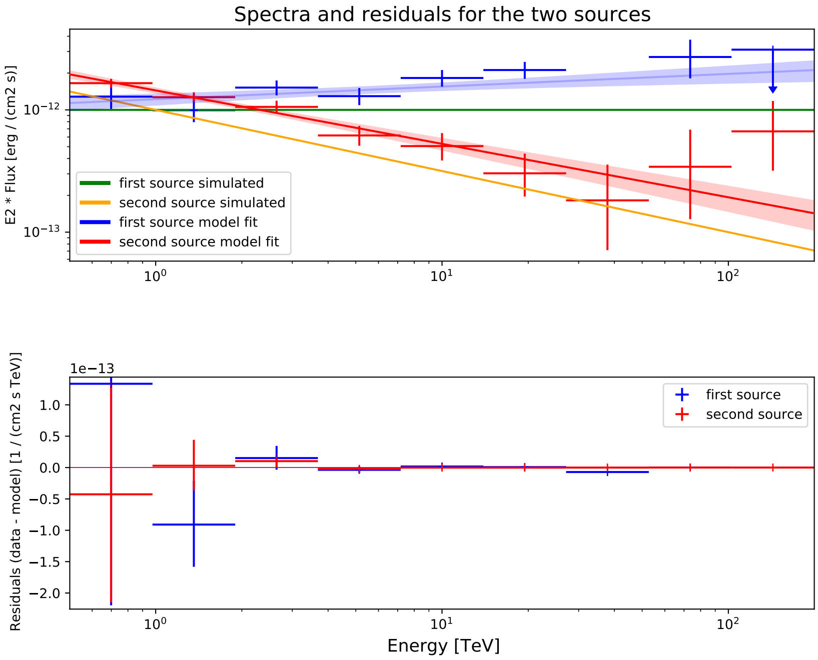


Figure 6.6: Example of spectra and residuals (data − fitted model) coming from a single fit adopting the two Gaussian sources model (blue and red lines); the shadowed regions show the 3σ interval within which the models can extend. These are data that refers to the simulated scientific case in which the first source is a Gaussian with $\sigma = 0.2°$ and the second source is a pointlike with flux normalization, photon index and distance from the first of $1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, 2.5 and 0.1°, respectively (green and orange lines). It is visible that the fitted spectra and the simulated ones are not so consistent with each other, thus, even if the two sources have been distinguished, their spectral features are still affected by the source confusion problem.

rameters were consistent with the simulated values within statistical uncertainties. Figure 6.6 shows an example of the resulting two sources spectra regarding the scientific case in which the first source is a Gaussian with $\sigma = 0.2°$ and the second is a pointlike with flux normalization, photon index and distance from the first of $1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, 2.5 and 0.1°, respectively. The two sources are resolved with a confidence level of 100%: this means that, in all of the analyzed fits, the two sources are resolved with a significance $> 5\sigma$, therefore the model with two sources is significantly better with respect to the one source model considering all the fits terminated successfully. However, from Figure 6.6 it is possible to observe that the fitted spectra show slightly different trends from the simulated ones, thus their spectral features are still affected by the source confusion problem, even if the two targets have been distinguished.

Examples of the Gaussian distributions for the pointlike spectral parameters are shown in Figure 6.7.

Also through 2D histograms, it is possible to display together the photon index and flux normalization at 1 TeV of the fitted sources, as shown in Figure 6.8 in which the scientific case presented is the same as before. Obviously, from these plots, only qualitatives conclusion can be inferred as no errors are taken into account.

Another way to analyse a single dataset is via confidence ellipses; in particular, the combination of two parameters can be displayed along with some statistical information, like means, standard deviations and the Pearson Correlation Coefficient (PCC) of two variables, i.e. a measure of linear correlation between two sets of data[1]. To compute those plots, a plotting function code has been used and can be found in the online *Matplotlib* documentation[2]; the full code implemented is explained in Appendix C, while some examples are shown in Figure 6.9, with the same scientific case as before. The radii of the ellipses are controlled by the number of σ. The percentage values corresponding to 1, 2 or 3 σ do not coincide to 68%, 95% and 99.7%, if the 2-D data are normally distributed like in these examples. Thus, a value of 1 makes the ellipse enclose 39.3% of the points (1σ in 1-D contain 68% of the data), while a value of 3 makes the ellipse enclose 99.4% of the points. In both panels of Figure 6.9, the correlation results to be positive and this is more evident considering the first source (left panel).

6.3 The case of one source detected

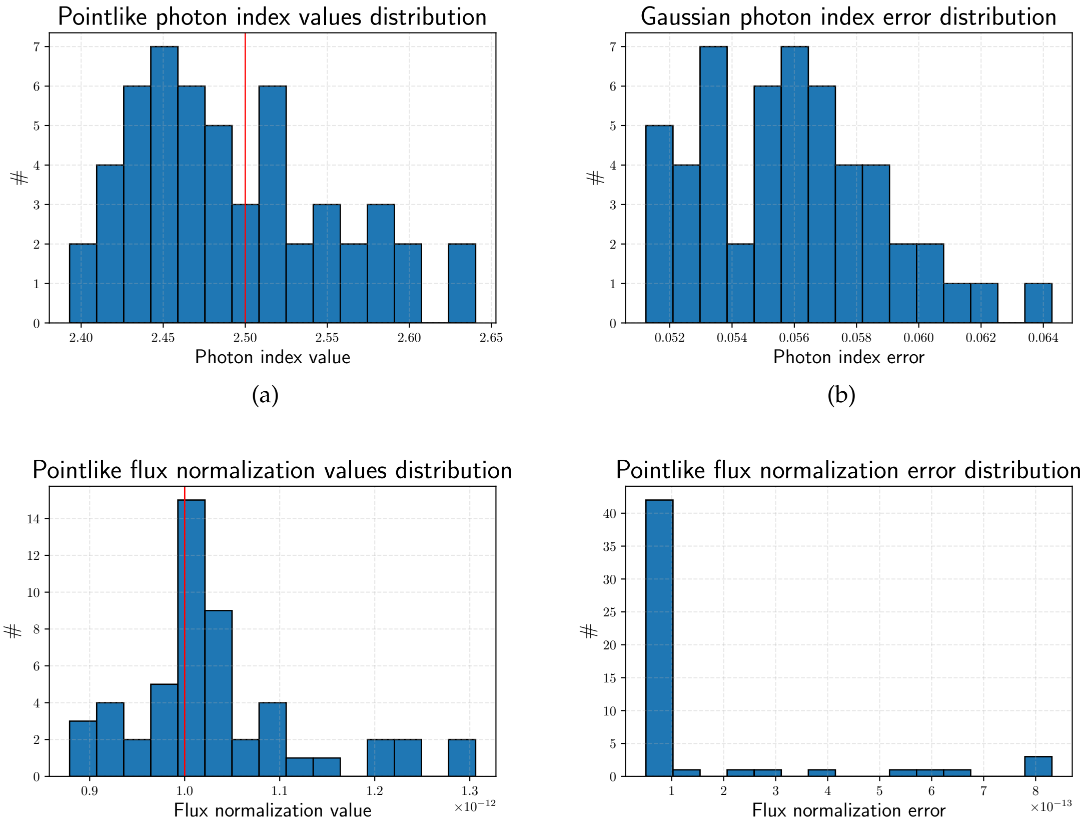
When ΔTS is lower than 25, only a single source is detected. If the two sources detection rate is lower than 68%, considering all the converged fits, it is more statistically likely that the two simulated sources were not distinguished. This is the classical example of a source confusion problem as the fitted parameters of a single source results influenced by the combination of two contributions; therefore, the derived parameters are different from the simulated ones.

Figure 6.10 shows an example of the resulting spectrum from a single fit with the one

---

[1] It is the ratio between the covariance of two variables and the product of their standard deviations; as such, it is essentially a normalised measurement of the covariance, such that the result always has a value between -1 and 1.

[2] https://matplotlib.org/devdocs/gallery/statistics/confidence_ellipse.html#sphx-glr-download-gallery-statistics-confidence-ellipse-py



(a) (b)

(c) (d)

Figure 6.7: Photon index Gaussian distributions concerning values (left panel) and errors (right panel). This figures refer to the second source of the scientific case in which the first source is a Gaussian with $\sigma = 0.2°$ and the second is a pointlike with flux normalization, photon index and distance from the first of $1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, 2.5 and 0.1°, respectively. Taking the mean from the values distribution and the sigma from the errors distribution, the resulting average photon index is 2.493 ± 0.003, while the flux normalization is $(1.03 \pm 0.22) \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$.

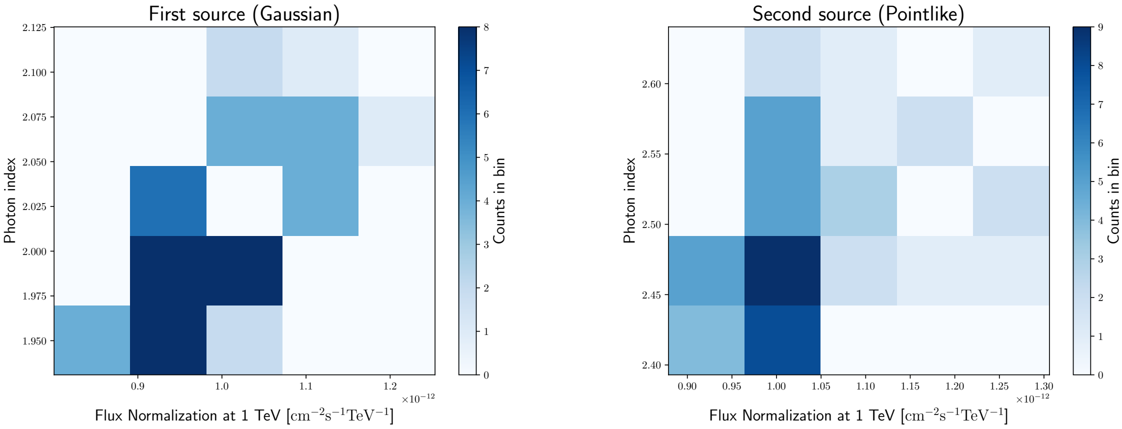


Figure 6.8: 2D histograms related to the scientific case in which the first source is a Gaussian with $\sigma = 0.2°$ and the second is a pointlike with flux normalization, photon index and distance from the first of $1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, 2.5 and 0.1°, respectively. The spectral features are shown for the source located at the center of the FoV, the so-called "first source" (left panel) and for the second source (right panel). It is possible to see that the ranges of the values of the parameters in which these are most likely are not so wide.

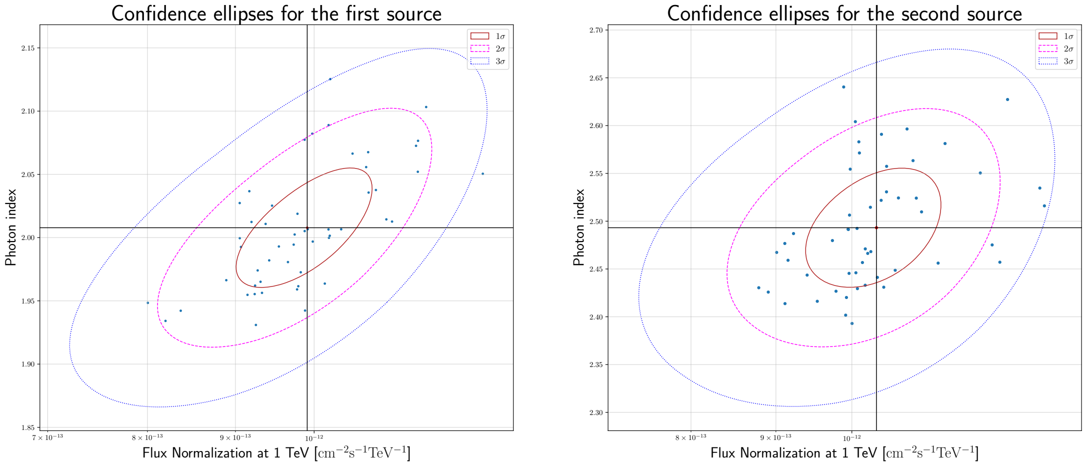


Figure 6.9: Confidence ellipses related to the fit with two sources and to the scientific case in which the first source is a Gaussian with $\sigma = 0.2°$ and the second is a pointlike with flux normalization, photon index and distance from the first of $1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, 2.5 and 0.1°, respectively. The left panel shows the resulting parameters for the so-called first source, the right panel shows the second one. In both panels, the red point at the intersection of the two perpendicular black lines corresponds to the mean values of the two parameters.

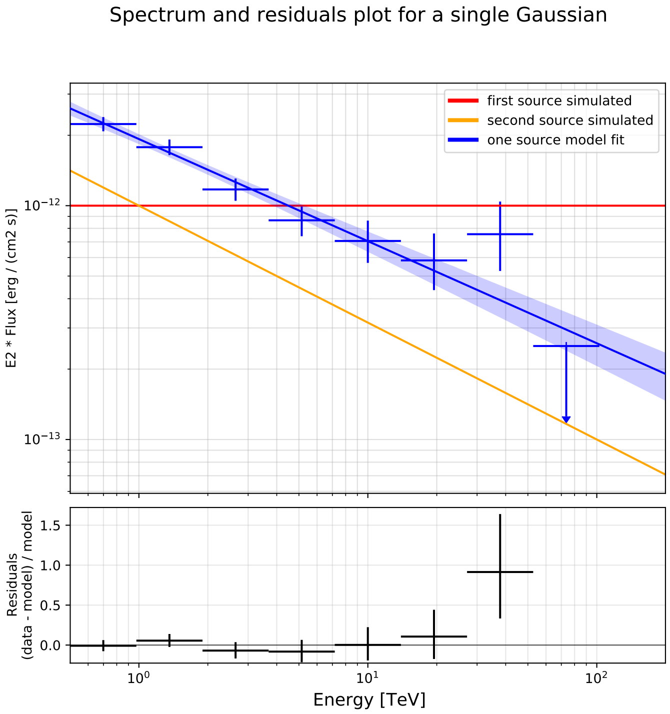


Figure 6.10: Example of spectrum and residuals ((data−fitted model)/fitted model) coming from a single fit with the one source model. These are data that refer to the simulated scientific case in which the first source is a Gaussian with $\sigma = 0.2°$ and the second is a pointlike with flux normalization, photon index and distance from the first of $1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, 2.5 and 0.15°, respectively. The blue line and datapoints refers to the fitted model with one source, while the red and orange lines are the models with which the first and second source have been simulated. The shadowed grey region shows the 3σ interval within which the fitted model can extend. The impact of the source confusion is clearly visible as the blue model is a combination of the two simulated ones.

source model regarding the scientific case in which the first source is a Gaussian with $\sigma = 0.2°$ and the second is a pointlike with flux normalization, photon index and distance from the first of $1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, 2.5 and 0.15°, respectively. This is one of the cases in which the detection rate has a falldown. The two sources are resolved with a significance $> 5\sigma$ with a 29.3% of confidence level: this means that the model with one source most likely provides a better solution compared to the two sources model (because the rate is lower than 68%).

The source confusion problem resulted more evident when a small source was considered along with a very extended source resembling a sort of diffuse emission. This was simulated when both sources were assumed to be Gaussian, one at the center of the FoV, with $\sigma = 1.0°$ and flux normalization and photon index of $1 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and 2.0, respectively, and a second Gaussian with $\sigma = 0.2°$. In all the considered combination of parameters of the second target, the two sources detection rate resulted always lower than the threshold (among the 45 out of 54 scientific cases in which the converged fits were $\gtrsim 20$), therefore the model with a single source is preferred over the other.

In particular, the detected source shows a distance from the center of the FoV and spectral features similar to that of the second source, but somehow distorted due to the combination of the two contributions. This is the example of what could be observed in the presence of a diffuse emission due to the interstellar medium or a halo in a MSC, for example.

In the case of the second target that has flux normalization and photon index equals to $4 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and 2.5, respectively, with a distance beween the sources that is 0.4°, it is possible to plot the spectrum resulting from the successfully terminated fits with a single source and its confidence level between 3σ in terms of flux normalization and photon index. This is shown in Figure 6.11, where it is evident that the flux normalization is overestimated due to the source confusion problem (note that this Figure is different from Figure 6.10 because here we consider all the converged fits among the 100 run). Here, the simulated spectrum is within the confidence level of 3σ but, in the end, what can be inferred is that, especially regarding the lower flux normalization, the detected source suffers from the effect of the simulated background with the extended source at the center of the FoV.

All the spectral features resulting from the fitting process of all the scientific cases are shown in summarizing plots in Figure 6.12. It is clear that the spectral parameter values are largely affected in the vast majority of cases by the superposition of the two sources: the flux normalization is overestimated in most cases, while the photon index shows particular inconsistency with the simulated values when the spectrum is harder.

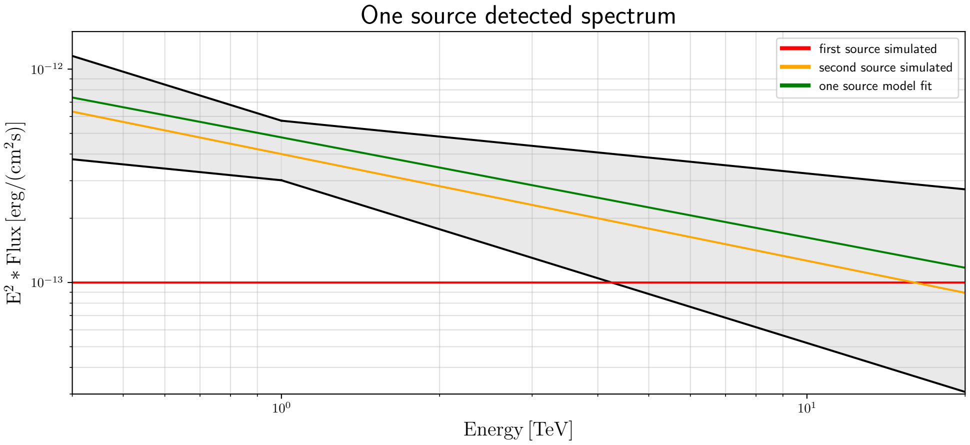


Figure 6.11: Example of a fitted spectrum resulting from all the converged fits among the 100 runs with a single source. The scientific case considered here is that of a Gaussian with $\sigma = 1.0°$ at the center of the FoV and a Gaussian with $\sigma = 0.2°$, flux normalization and photon index equal to $4 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and 2.5, respectively, and a distance between the sources that is 0.4°. The green line shows the spectrum with the mean fitted values, while the shadowed grey region shows the 3σ interval within which the model can extend. The red and orange lines are the models with which the two sources have been simulated.

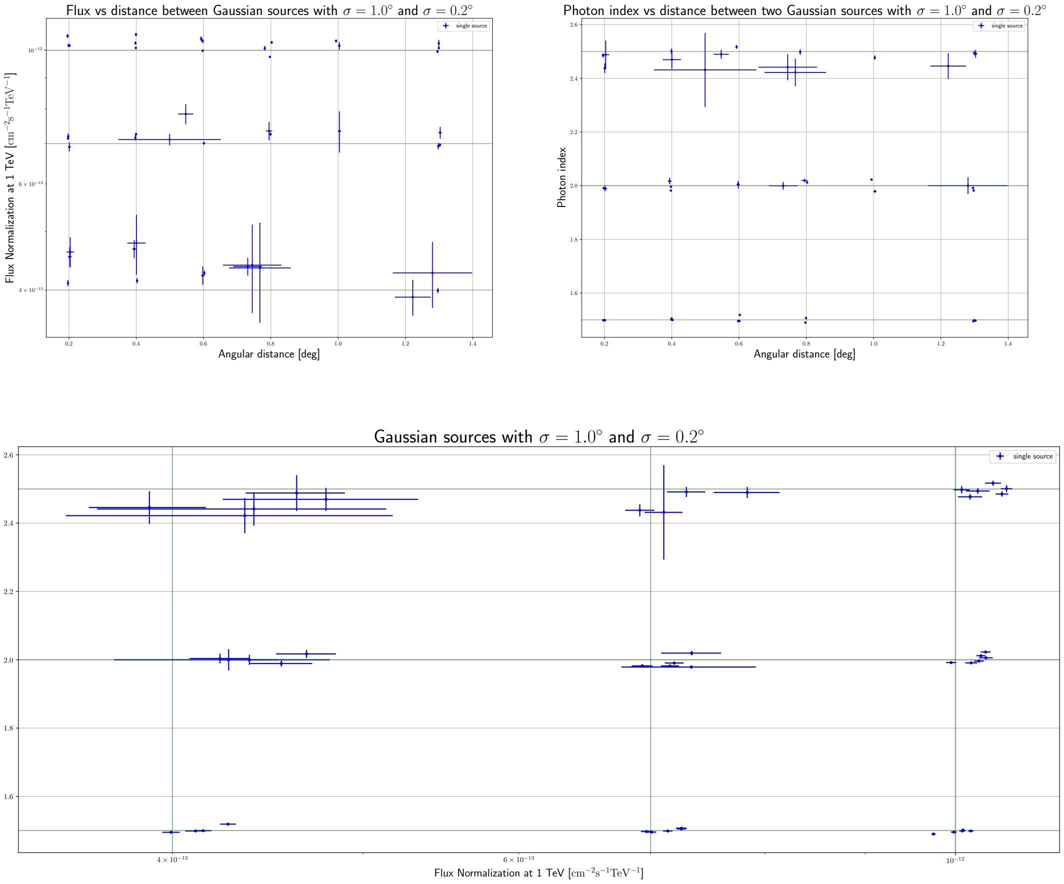


Figure 6.12: Flux normalization and photon index as a function of longitude (top) and photon index as a function of flux normalization (bottom). These are results of 45 out of 54 scientific cases including a Gaussian source, with $\sigma = 1.0°$ and flux normalization and photon index of $1 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and 2.0, respectively, and a Gaussian with $\sigma = 0.2°$. The excluded cases are that with $\lesssim 20$ of successfully terminated fits. These data refers to the fit in which a single source has been detected as the detection rate was always lower than 68%. Finally, the green lines refer to the simulated values of the spectral parameters to show the agreement between these values and the fitted ones. Note that for lower flux normalization and harder spectra, the errorbars are larger and more visible because the number of successfully converged pair of fits is lower.

Chapter 7

Conclusions

Observations with the current generation of VHE gamma-ray telescopes have revealed an astonishing variety of particle accelerators in the Milky Way, such as SNRs, PWNe and binary systems. The upcoming CTAO will enable a survey of the entire Galactic plane in the energy range from a few tens of GeV to 300 TeV with unprecedented sensitivity by up to a factor 5 to 10, depending on the energy range, with respect to currently operating facilities, and significantly better angular resolution ($\lesssim 0.05°$ at $E > 1\,\mathrm{TeV}$), breaching the limits of the current facilities. Despite the improved angular resolution, CTAO will be largely affected by source confusion, a problem that concerns multiple ovelapping sources and multiple detections related to a single extended source. Moreover, source confusion is particularly relevant in the Galactic Plane due to the presence of the diffuse emission, i.e. the sum of the IEM and contribution of unresolved sources, and high source density. In this context, it is of crucial importance to deal with it, and in this work I tackled in a systematic way the revolving capabilities of CTAO, as first step in the study of the source confusion problem.

To reach this goal, different science cases were simulated using the approved CTAO Science Analysis tool *Gammapy*, taking into accont two sources each time. The targets have been modeled adopting two spatial models, pointlike and Gaussian, and with a powerlaw spectral model, varying the model parameters in order to obtain a reasonable differentiation of the scientific cases that may be considered. Then, the performed analysis relied on a 100-fitting loop, repeated twice to test both the model with one source (null hypothesis) and that with two sources. Being the two models nested, the significance of the two sources model with respect to that of one source model is obtained by computing the ΔTS, i.e. the difference between the test significance of each model. In this way it is possible to compare the two hypotheses and define the two sources detection rate, i.e. the percentage of fits pairs in which the ΔTS resulted > 25, among all the successfully converged fits. The two sources detection rate was displayed for every scientific case in which the statistics was sufficient (number of pair fits successfully converged $\gtrsim 20$) and the threshold was exceeded for some combination of spectral parameters and for distances between the sources. Hence, it is possible to say if we can actually resolve two sources with a confidence level of at least 68% and for which combination of spectral parameters and angular distance, concerning the cases in which both of them are pointlike, in which a Gaussian with sigma 0.2°, 0.3° or 0.5° and a pointlike are present and when both sources are Gaussian with $\sigma = 0.2°$.

The detection rate increases, as expected, at increasing distances between the sources for the science cases in which the two sources have been simulated with the same spatial model, pointlike or Gaussian, but with a different scale of distances; we note that, for extended sources, the minimum angular distance at which the two sources are resolved with a confidence level of 90% is larger than for pointlike sources (Figures 6.1 and 6.5). Harder spectra with lower flux normalization provide more challenging cases in terms of capability of distinguishing the sources; however, it is worth noting that, for the lowest values of the flux normalization, there is the detectable flux limit problem, so that the second source can not be detected at all.

The science cases adopting a Gaussian with σ equal to 0.2° or 0.3° plus a pointlike show instead a different behaviour: for the majority of the spectral parameter combinations, the two sources are resolved at short distances, from the very first step; then, when the distance between the sources has values similar to the σ of the Gaussian, a falldown is present and the two sources cannot be resolved anymore; at higher values, both sources are again detectable as distinct targets (Figures 6.2, 6.3). A similar trend was found for the cases in which a Gaussian with $\sigma = 0.5°$ and a pointlike source were present, but for some combinations of spatial and spectral parameters the failed fits dominated over the successfully converged ones and this led to a limited scientific case study (Figure 6.4). This is a counterintuitive behaviour that would require further studies before establishing a firm conclusion.

If the capability of resolving the two sources is the first step in the process of addressing the source confusion, the capability of estimating parameters in agreement with the simulations is the second one. The latter gives us the possibility to estimate the uncertainties on the parameter determination.

In the case when two sources are detected, we checked the consistency of the estimated parameters with the simulated ones within the computed statistical uncertainties. In particular, there is larger dispersion around the original values when a Gaussian and a pointlike are present (Figures D.3, D.4 and D.5); furthermore, the errobars are larger because there were less successfully converged fits. When the two sources are both pointlike or both Gaussian with $\sigma = 0.2°$, the spectral and spatial parameters result to be in agreement with the simulated ones. The only exception is for the sigma values of the Gaussian at the center of the FoV that, in some cases, is overestimated, reaching values of nearly $\sigma = 0.4°$ for some parameters combination (Figures D.1, D.2, D.6 and D.7).

When a single source is detected, instead of two that have been simulated, we have the classic source confusion problem. This resulted in all the scientific cases in which we had the falldown of the two sources detection rate, when a Gaussian and a pointlike have been considered: the obtained spectra were clearly affected by the contribution of the two simulated sources, especially regarding the flux normalization. Interesting cases occured when a very extended source was simulated, as it were a "background" emission. Indeed, for all the cases in which a Gaussian with $\sigma = 1.0°$ was considered along with another smaller Gaussian with $\sigma = 0.2°$, the detection rate never exceeded the threshold, and the spectral parameters of the resulting single source are a combination of the two simulated ones (Figures 6.12 and 6.11). These are good examples of how a diffuse emission can probably influence future observations. To overcome this issue, it is necessary to infer the contribution of unidentified sources through extrapolation coming from simulations of synthetic source populations, which are based on real data. Studies like this lead to the results of the 1-DC, including physically-motivated models for Galactic source populations, like the one that can be found in [34]. Furthermore, the IEM contribution needs also to be considered: currently, our knowledge of it comes from codes solving the cosmic ray transport equations and taking into account the gas maps of the interstellar medium through which it is possible to make predictions [43].

7.1 Improvements and future prospects

The study presented in this master thesis can be improved, on the one hand by enlarging the explored scientific cases to make them more realistic, having either Gaussian spatial morphology or more complex spatial models, and, on the other hand, by improving the statistics of the fitting pipeline by producing 1000 fits instead of 100. For the Gaussian spatial model, different σ values can be considered, even reaching $\sigma \sim 10°$ to reproduce further additional large-scale structures that cover the entire FoV, as a diffuse emission. By doing so, it can be possible to obtain enough pair of fits to analyse also the more difficult cases in which failed fits are dominant.

This has not been carried out here because the fitting process with *Gammapy*, especially with extended sources, is very time-consuming. For example, $8 - 12$ hours are necessary to obtain all the results coming from the 200 fit procedure of a single scientific case, given flux normalization, photon index and distance values. The parallelization of the code has been used to speed up the procedure, with access to a cluster consisting of 4 computing nodes with 64 cores each (managed by a schedular "IBM loadleveler") for a total of 32GB of RAM, but additional improvements can be adopted to better parallelize the working code. As stated in Section 4.2, *Gammapy* is a young project that is still under heavy development, thus improvements from this point of view can also be considered, in order to carry out further works like the one presented in this thesis. Moreover, limiting the parameter values can be another way to speed up the fitting process.

Two additional aspects can be investigated as follow-up work of this thesis: the usage of ECPL model to fit the data, to see if it can be statistically preferred over a PL model in a given energy interval for those VHE sources that seem to require it; the introduction of the spectral analysis of smaller energy ranges within 0.5 and 200 TeV, that is very important to understand where source confusion can be, at least partially, overcome, and how the spectral analysis can be affected, depending on the energy interval.

Furthermore, the usage of a realistic diffuse model by assuming a theoretical IEM can be matter of in-depth investigations. This would introduce the dependence on the specific position in the sky; furthermore, it could lead to draw different conclusions for confused regions depending on the pointing direction that can be chosen toward the Galactic Center or its anticenter.

Appendices

Appendix A

Scientific cases

Here, all the scientific cases considered to simulate different kinds of confused regions are listed in a systematic way. It is recalled that, to perform the sources simulation, the two spatial models considered are the pointlike and Gaussian templates and that the spectral model assumes a PL emission. Furthermore, the first source is the one simulated at the center of the pointing direction with a fixed spectrum in all the simulations, while the spatial and spectral parameters of the second source are changed in each case.
The following Tables show the flux normalization at 1 TeV and photon index values of the so-called second source, i.e. the source that fall in the FoV centered on the primary target, along with the width of the Gaussian when an extended sources is present. Moreover, the two sources detection rate, i.e. the percentage by which the two sources have been distinguished considering all the fit terminated successfully according to the analysis, is reported for each step in distance between the sources considered (five distances for the pointlike, eight for the cases in which one Gaussian is present, six when both are Gaussian).

Two Pointlike sources

Flux normalization $\left[\mathrm{cm^{-2}s^{-1}TeV^{-1}}\right]$	Photon index	Two sources detection rate (for every distance step)
1×10^{-12}	1.5	90.6%, 100%, 100%, 100%, 100%
	2.0	5%, 100%, 100%, 100%, 100%
	2.5	10.1%, 100%, 100%, 100%, 100%
7×10^{-13}	1.5	97.5%, 100%, 100%, 100%, 100%
	2.0	1.4%, 100%, 100%, 100%, 100%
	2.5	3.8%, 100%, 100%, 100%, 100%
4×10^{-13}	1.5	74.4%, 100%, 100%, 100%, 100%
	2.0	1.9%, 98.9%, 100%, 100%, 100%
	2.5	0%, 60.3%, 100%, 100%, 100%
1×10^{-13}	1.5	13.8%, 100%, 100%, 100%, 100%
	2.0	0%, 5.6%, 80.8%, 98.8%, 100%
	2.5	0%, 0%, 3.1%, 51.4%, 74.2%
7×10^{-14}	1.5	5%, 100%, 100%, 100%, 100%
	2.0	0%, 2%, 43.1%, 82.2%, 100%
	2.5	0%, 0%, 0%, 1.8%, 31.3%
4×10^{-14}	1.5	0%, 74.6%, 100%, 100%, 100%
	2.0	0%, 0%, 5.6%, 19.7%, 46.2%
	2.5	0%, 0%, 0%, 0%, 5.5%

Table A.1: Characterization of the second source as a pointlike for the scientific cases considered in this work. For each entry, the distance between the sources has been varied from 0.02° to 0.10° with a step of 0.02°; the percentage refers to these five steps of distance. Note that for the last couple of spectral parameters it is even more difficult to distinguish the two sources: this is due to the lower limit in flux that is actually higher than $4 \times 10^{-14}\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ for $\gamma = 2.5$.

Pointlike and Gaussian sources

Sigma σ [deg]	Flux normalization $\left[\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}\right]$	Photon index	Two sources detection rate (for every distance step)
0.2°	1×10^{-12}	1.5	100%, 100%, 82.6%, 80%*, 100%*, 100%, 100%, 100%*
		2.0	100%, 100%, 94%, 37.8%, 31.3%, 100%, 100%, 100%
		2.5	100%, 100%, 29.3%, 8.5%, 16.7%, 100%, 100%, 100%
	7×10^{-13}	1.5	100%, 98.4%, 100%, 76.9%*, 100%, 100%, 100%, 100%*
		2.0	100%, 100%, 46%, 25%, 25%, 100%, 100%, 100%
		2.5	100%, 100%, 12.8%, 10.4%, 4.1%, 83.3%, 100%, 100%
	4×10^{-13}	1.5	100%, 100%, 100%, 100%*, 66.7%*, 100%, 100%*, 100%
		2.0	98.4%, 66.7%, 8.3%, 3.3%, 0%, 40.9%, 100%, 100%
		2.5	32.8%, 5.3%, 0%, 0%, 0%, 1.2%, 70.2%, 87.9%
	1×10^{-13}	1.5	4.8%, 0%, 0%, 0%, 0%, 2.2%, 22.2%, 88.2%
		2.0	0%, 0%, 1%, 0%, 0%, 0%, 0%, 3.6%
		2.5	0%, 0%, 1.1%, 1.2%, 0%, 2.2%, 1.2%, 2.5%
0.3°	1×10^{-12}	1.5	100%, 100%, 100%, 71.4%*, 100%, 0%*, 87.5%*, 100%*
		2.0	95.2%, 100%, 96%, 88.8%, 57.1%*, 62.5%*, 33.3%*, 100%
		2.5	96.8%, 100%, 100%, 89.3%, 33.3%, 28.1%, 24%, 100%
	7×10^{-13}	1.5	100%, 100%, 100%, 91.7%*, 100%, 0%*, 95.5%, 100%
		2.0	100%, 100%, 100%, 100%, 50%, 35.3%, 16.7%, 100%
		2.5	100%, 100%, 97.5%, 42.9%, 27%, 25%, 31.8%, 84%
	4×10^{-13}	1.5	100%, 100%, 100%*, 100%*, 100%*, 0%*, 100%*, 100%
		2.0	97.5%, 97.1%, 85.7%, 25%, 18.5%, 38.2%, 13.5%, 27%
		2.5	35.2%, 11.4%, 5.1%, 1.2%, 0%, 2.1%, 1.1%, 5.7%
	1×10^{-13}	1.5	4.8%, 3%, 1.2%, 0%, 0%, 0%, 1.1%, 9.7%
		2.0	0%, 0%, 0%, 0%, 0%, 0%, 0%, 1.1%
		2.5	0%, 0%, 0%, 0%, 0%, 0%, 2.1%, 3.1%

Table A.2: Characterization of the second source spectral features as a pointlike with a given σ of the Gaussian at the center of the pointing direction, with flux normalization and photon index equal to $1 \times 10^{-12}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$ and 2.0, respectively. For all the cases, the distance between the sources has been varied from 0.05° to 0.40° with a step of 0.05°. The percentage refers to those eight steps of distance; values reported with " * " refer to cases in which the number of fits terminated successfully are $\lesssim 20$ over 100 performed.

Pointlike and Gaussian sources

Sigma σ [deg]	Flux normalization $\left[\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}\right]$	Photon index	Two sources detection rate (for every distance step)
0.5°	1×10^{-12}	1.5	97%, 87.5%, 92.3%*, 94.7%, 62.5%*, 100%, 100%*, 100%
		2.0	100%, 95.5%, 81%, 33.3%*, 33.3%*, 100%, 100%*, 100%
		2.5	100%, 92.3%, 73.9%, 80%*, 75%*, 96.6%, 95%, 95.2%
	7×10^{-13}	1.5	100%, 90%, 80%*, 100%, 100%, 100%, 100%*, 98.2%
		2.0	97.4%, 100%, 87.5%, 60%*, 0%*, 100%*, 100%*, 100%*
		2.5	100%, 93.8%, 71.4%*, 42.9%*, 50%*, 92.9%*, 100%, 100%
	4×10^{-13}	1.5	94.9%, 93.3%*, 100%*, 100%*, 66.7%*, 100%, 100%*, 100%
		2.0	100%, 93.8%, 27.3%*, 33.3%*, 44.4%*, 50%*, 100%*, 95.7%
		2.5	34.5%, 20%, 4.1%, 11.9%, 4.7%, 13.2%, 57.1%, 80%
	1×10^{-13}	1.5	0%*, 7.1%*, 7.1%*, 6.9%, 6.25%, 12.2%, 29.2%, 90.5%
		2.0	0%, 0%, 0%, 2.2%, 0%, 12.2%, 12.9%, 15.5%
		2.5	0%, 0%, 0%, 1.1%, 6.3%, 2.6%, 3.9%, 7.1%

Table A.3: Characterization of the second source spectral features as a pointlike with a given σ of the Gaussian at the center of the pointing direction with flux normalization and photon index equal to $1 \times 10^{-12}\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$ and 2.0, respectively. For all the cases, the distance between the sources has been varied from 0.1° to 0.8° with a step of 0.1°. The percentage refers to those eight steps of distance; values reported with " * " refer to cases in which the number of fits terminated successfully are $\lesssim 20$ over 100 performed.

Pointlike and Gaussian sources with flux normalization

Sigma σ [deg]	Flux normalization $\left[\mathrm{cm^{-2}s^{-1}TeV^{-1}}\right]$	Photon index	Two sources detection rate (for every distance step)
1.0°	1×10^{-12}	1.5	42.9%, 37.5%, 50%, 11.1%, 33.3%, 80%, 100%, 0%
		2.0	27.3%†, 22.2%, 41.7%†, 60%, 25%, 50%, 66.7%, 60%
		2.5	0%, 0%†, 0%, 40%, 20%, 33.3%, 0%, 60%,
	7×10^{-13}	1.5	33.3%, 20%†, 20%, 50%, 100%, 0%, 50%, 50%
		2.0	42.9%, 14.3%, 33.3%, 60%, 50%, 100%, 33.3%, 57.1%
		2.5	57.1%, 12.5%, 20%, 18.2%†, 33.3%, 20%†, 0%, 14.3%
3.0°	1×10^{-12}	1.5	0%, 100%, 0%, 0%, 0%, 0%, 100%, 100%
		2.0	50%, 100%, 33.3%, 100%, 0%, 100%, 100%, 0%
		2.5	25%, 40%, 60%, 50%, 60%, 40%, 33.3%, 25%
	7×10^{-13}	1.5	0%, 50%, 100%, 100%, 0%, 0%, 100%, 0%
		2.0	71.4%, 25%, 100%, 50%, 0%, 66.7%, 0%, 0%
		2.5	66.7%, 33.3%†, 0%, 0%, 50%, 0%, 50%, 50%

Table A.4: Characterization of the second source spectral features as a pointlike with a given σ of the Gaussian at the center of the pointing direction with flux normalization and photon index equal to $1 \times 10^{-13}\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and 2.0, respectively. For all the cases, the distance between the sources has been varied from 0.1° to 0.8° with a step of 0.1°; the percentage refers to these eight steps of distance. All the scientific cases have a number of fits terminated successfully that is < 20 over 100 performed; however, some cases are reported with the † symbol to indicate that at least 10 fits were successfull and that some kind of trend can still be derived from those data.

Two Gaussian sources

	First source with $\sigma = 0.2^\circ$		
Sigma σ [deg]	Flux normalization $[\mathrm{cm^{-2}s^{-1}TeV^{-1}}]$	Photon index	Two sources detection rate (for every distance step)
0.2°	1×10^{-12}	1.5	24.7%, 100%, 100%, 100%, 98.5%, 100%,
		2.0	0%, 95%, 100%, 100%, 100%, 100%,
		2.5	1.17%, 87%, 100%, 100%, 100%, 100%,
	7×10^{-13}	1.5	20.8%, 98.3%, 100%, 100%, 100%, 100%,
		2.0	0%, 59%, 100%, 100%, 100%, 100%,
		2.5	0%, 39%, 99%, 100%, 100%, 100%,
	4×10^{-13}	1.5	4.2%, 100%, 100%, 100%, 100%, 100%
		2.0	0%, 13%, 100%, 100%, 100%, 100%
		2.5	0%, 3.2%, 36.4%, 64%, 65%, 60%
	1×10^{-13}	1.5	0%, 11%, 81%, 98%, 96%, 94%
		2.5	2.8%, 0%, 0%, 2.4%, 1.1%, 2.4%
		2.5	0%, 0%, 1.4%, 0%, 1.4%, 0%

	First source with $\sigma = 1.0^\circ$		
Sigma σ [deg]	Flux normalization $[\mathrm{cm^{-2}s^{-1}TeV^{-1}}]$	Spectral index	Two sources detection rate (for every distance step)
0.2°	1×10^{-12}	1.5	0%, 3.7%, 8.8%, 0%, 7.1%*, 0%,
		2.0	0%, 2.6%, 4.2%, 0%, 0%, 4.9%
		2.5	0%, 0%, 4%, 4.2%, 8.7%, 8.6%
	7×10^{-13}	1.5	4.9%, 0%, 0%, 0%, 10%*, 3.3%
		2.0	0%, 0%, 0%*, 18.8%, 10.5%, 0%
		2.5	0%, 6.3%, 8.1%, 0%*, 8.3%*, 4.5%
	4×10^{-13}	1.5	0%, 0%, 3.2%, 0%*, 0%*, 3.3%
		2.0	0%, 0%, 0%, 5%, 9.1%*, 3.3%
		2.5	0%, 0%, 0%*, 0%, 10%, 11.8%

Table A.5: Characterization of the second source spectral features as a Gaussian with $\sigma = 0.2$ along with a given σ of the Gaussian at the center of the pointing direction. The photon index of the primary source is 2.0 while the flux normalization is equal to $1 \times 10^{-12}\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ when has $\sigma = 0.2^\circ$ and $1 \times 10^{-13}\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ when has $\sigma = 1.0^\circ$. For all the cases, the distance between the sources has been varied from 0.2° to 1.0° with a step of 0.2°, adding an additional value equal to 1.3°; the percentage refers to these six steps of distance. Values reported with “ * ” refer to cases in which the number of fits terminated successfully are $\lesssim 20$ over 100 performed

Appendix B

The lowest flux for a pointlike source

During the simulation process, it is necessary to perform a preliminary analysis of the extreme cases in terms of flux normalization: the lowest detectable flux for pointlike sources can be established producing 2-dimensional significance maps and running a peak finder to detect source candidates, as shown in Code Listing B.1.

The tool `gammapy.estimators.utils.find_peaks` analyses the significance image, i.e. the TS map, to get a list of point-source candidates (positions and peak significance values). The `find_peaks` function performs a local maximun search in a sliding window, the argument `min_distance` is the minimum pixel distance between peaks (smallest possible value, and default is 1 pixel).

Code Listing B.1: Given a dataset with a simulated source, it is possible to convert it to a 2-dimensional map and compute the TS map in a fixed energy range. Then, using `gammapy.estimators.utils.find_peaks`, it is possible to verify if any source candidate can be detected above a given threshold. If not, the code below will end with a *KeyError* because no `source_ts["ra"]` has been created.

```
dataset_image=dataset.to_image() # Convert to a 2-D map
estimator = TSMapEstimator(model1,energy_edges=[0.5, 200] * u.TeV,
    kernel_width=0.05*u.deg)
images_ts = estimator.run(dataset_image) # Compute TS map
sources_ts = find_peaks(images_ts["sqrt_ts"].get_image_by_idx((0,)),
    threshold=5, min_distance=10) # Find any detectable source
print(sources_ts)

# Plotting the significance map
plt.figure(figsize=(15, 5))
```

```
_, ax, _ = images_ts['sqrt_ts'].plot(add_cbar=True)

# Plotting a circle around any detectable source to check the position on the
    significance map
ax.scatter(
sources_ts["ra"],
sources_ts["dec"],
transform=plt.gca().get_transform("icrs"),
color="none",
edgecolor="b",
marker="o",
s=200,
lw=1.,
)
```

By performing this preliminary analysis, the flux normalization lower limit for pointlike sources was set to $4 \times 10^{-14}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$; actually, for a pointlike with a steep spectrum ($\gamma = 2.5$) the true lower limit is $5 \times 10^{-14}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$, but for completeness with the cases considered, the lower limit has been applied equal for all the photon index.

Appendix C

Confidence ellipse code

Related to the confidence ellipses shown in Section 6.2, here the full code implemented to compute those plots is displayed. As already stated, this can be found in the online *Matplotlib* documentation.

Code Listing C.1: Full code implemented to compute confidence ellipses for a given dataset.

```
from matplotlib.patches import Ellipse
import matplotlib.transforms as transforms
from scipy.stats import norm

def confidence_ellipse(x, y, ax, n_std=3.0, facecolor=’none’, **kwargs):
"""
Create a plot of the covariance confidence ellipse of *x* and *y*.

Parameters
----------
x, y : array-like, shape (n, )
Input data.

ax : matplotlib.axes.Axes
The axes object to draw the ellipse into.

n_std : float
The number of standard deviations to determine the ellipse’s radiuses.

**kwargs
Forwarded to ‘˜matplotlib.patches.Ellipse‘
```

```

Returns
-------
matplotlib.patches.Ellipse
"""
if x.size != y.size:
raise ValueError("x and y must be the same size")

cov = np.cov(x, y)
pearson = cov[0, 1]/np.sqrt(cov[0, 0] * cov[1, 1])
# Using a special case to obtain the eigenvalues of this
# two-dimensionl dataset.
ell_radius_x = np.sqrt(1 + pearson)
ell_radius_y = np.sqrt(1 - pearson)
ellipse = Ellipse((0, 0), width=ell_radius_x * 2, height=ell_radius_y * 2,
facecolor=facecolor, **kwargs)

# Calculating the stdandard deviation of x from
# the squareroot of the variance and multiplying
# with the given number of standard deviations.
scale_x = np.sqrt(cov[0, 0]) * n_std
mean_x = norm.fit(x)[0]  #np.mean(x)

# calculating the stdandard deviation of y ...
scale_y = np.sqrt(cov[1, 1]) * n_std
mean_y = norm.fit(y)[0]  #np.mean(y)

transf = transforms.Affine2D() \
.rotate_deg(45) \
.scale(scale_x, scale_y) \
.translate(mean_x, mean_y)

ellipse.set_transform(transf + ax.transData)
return ax.add_patch(ellipse)
```

Appendix D

Summarizing plots

Regarding the summarizing plots, some of them are shown here instead to be presented in Chapter 6 in order to avoid heavy reading.

D.1 Pointlike sources

Regarding the confused region in which two pointlikes sources have been simulated, 90 scientific cases are considered (Table A.1). The results from the fits with the two models, single source and double sources, are displayed in some plots like those in Figure D.1: photon index and flux normalization at 1 TeV as a function of distance between the sources and photon index as a function of flux normalization. Two sources are considered distinguished where the percentage for which this happened among the fit terminated successfully is larger than 68%: the values of the parameter are then reported in those plots as red and orange data, referring to first and second source, respectively. Otherwise, when the two sources have not been distinguished, the reported data are the blue ones, referring to the fits as a single source.

It is very clear that when the second source is simulated with lower flux normalization ($< 4 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$), it is more difficult to distiguish the two sources, especially with a steeper spectrum (> 1.5). Even when they are both detected, the errorbars are larger than those at higher flux normalization. Moreover, when a single source is detected, the flux normalization at 1 TeV is equal to that of the primary source or is slightly overestimated, while the photon index can result over- or underestimated. It is worth noting that, because the lower limit of the flux normalization at

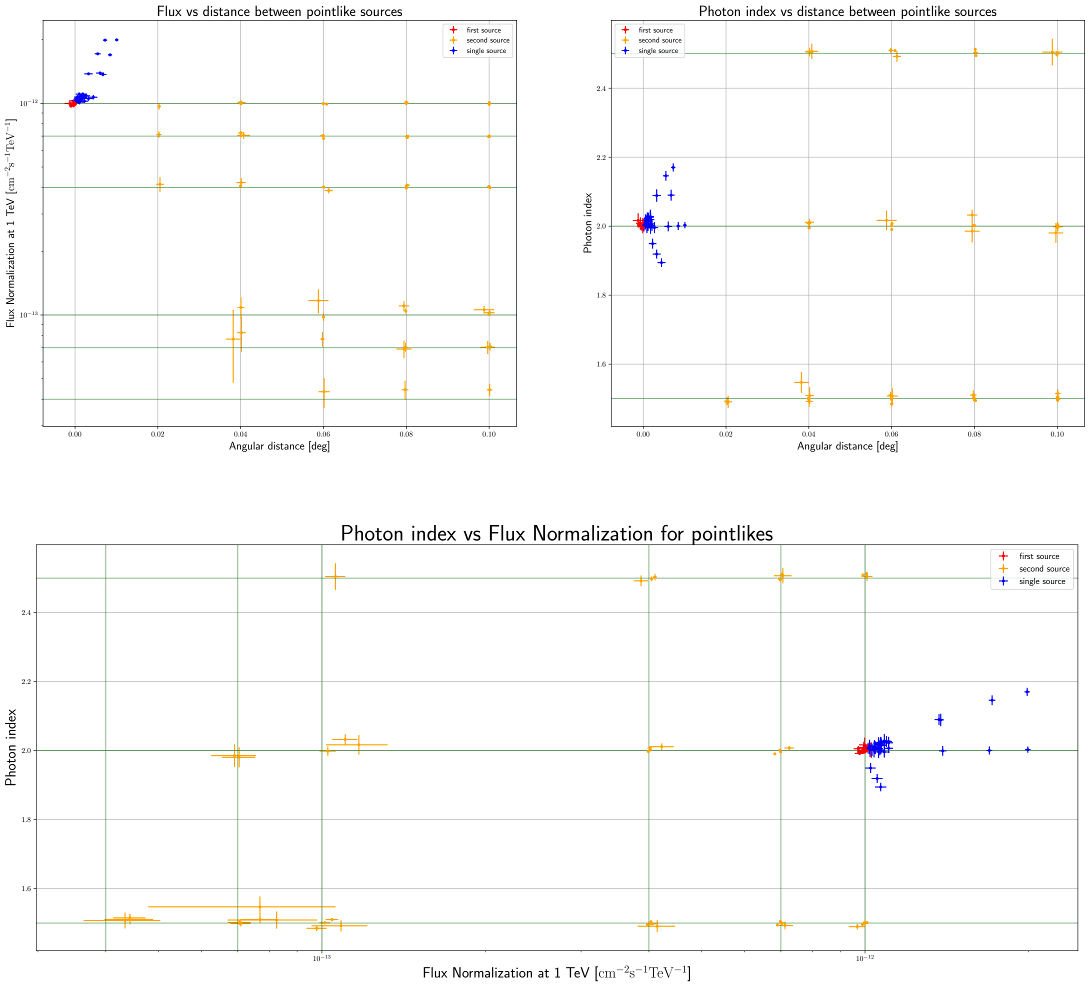


Figure D.1: Flux normalization and photon index as a function of distance between the sources (top) and photon index as a function of flux normalization (bottom). These results are obtained from the 90 scientific cases including two pointlike sources each time. Blue data refer to the fit in which a single source has been detected; red and orange data refer to the fit in which the two sources are distinguished. The values are reported with their errorbar, even if they are not always distinguishable, and blue errorbars are multiplied by a factor of 20. The parameter values are displayed when the rate of distinguished sources is higher or lower than 68%, depending on the detection of the two sources or not. Finally, the green lines refer to the simulated values of the spectral parameters to show the agreement between those values and the fitted ones.

1 TeV is $5 \times 10^{-14}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ for the steeper spectrum considered in this work, the two sources are not distinguished at any distance because the second source will never be detected.

A comparison between the parameter values of the two sources detected as a single one and of the primary source is shown in Figure D.2: the errorbar referring to the detection of the primary target, when the two sources are distinguished, seems larger only because the vertical scales are smaller and the spectral parameter values are more concentrated around the original values with which the source have been simulated. In these plots, one can see how the two sources appear as a single one when they are not distinguished; in particular, the resulting spectrum seems to be that of the primary source somehow distorted, as the flux normalization is $\geq 1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and the photon index value is always not so different from 2.

D.2 Pointlike and extended sources

Here, an extended source has been considered as the primary source at the center of the pointing direction with a pointlike of which spectral parameters are varied, along with its distance from the first. For the extended sources with $\sigma \geq 1°$, the flux normalization of the primary source at 1 TeV has been changed to $1 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, as already stated in Section 4.3. These are the scientific cases shown in Tables A.2, A.3 and A.4.

From the Tables it appears evident that there were many cases in which the failed fits have exceeded the threshold set (these are indicated with an "*"). Therefore, it is still possible to obtain a distiguished rate, but this is no longer statistically valid, as well as the values of the parameters fitted. This has been taken into account when working on the summarizing plots so that those "failed cases" are not considered; it also results in a number of limited tested case caused by the failed fits and a consequent larger errorbars estimate. Moreover, there is much more dispersion regarding the parameter values, both considering one or two sources. Summarizing plots are produced for each sigma value.

In Figure D.3 it is shown the $\sigma = 0.2°$ case. Here, ignoring the cases in which the successfully terminated fits were $\lesssim 20$, 87 out of 96 scientific cases are considered.

When a single source is detected, the resulting spectrum seems to be that of the primary source somehow distorted, as in Section D.1, with flux normalization around

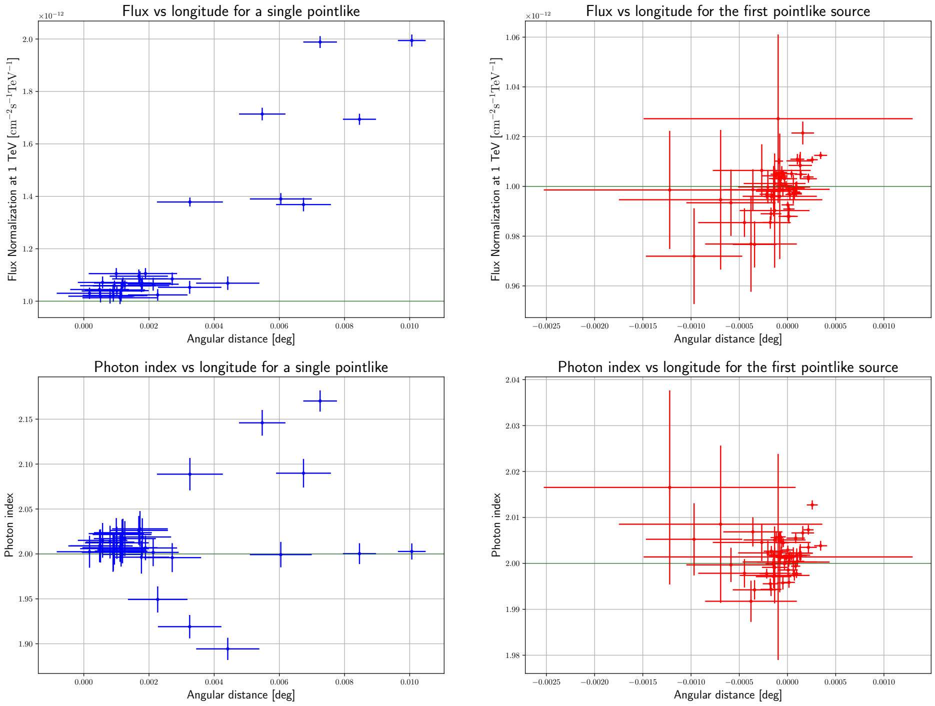


Figure D.2: Comparison between the parameter values of the two sources detected as a single one (left panels) and of the primary source (right panels) for the pointlike cases. The parameter values are displayed when the rate of distinguished sources is higher or lower than 68%, depending on the detection of the two sources or not. Note that errorbars displayed in the left panels are multiplied by a factor of 20, otherwise they would not have been visible.

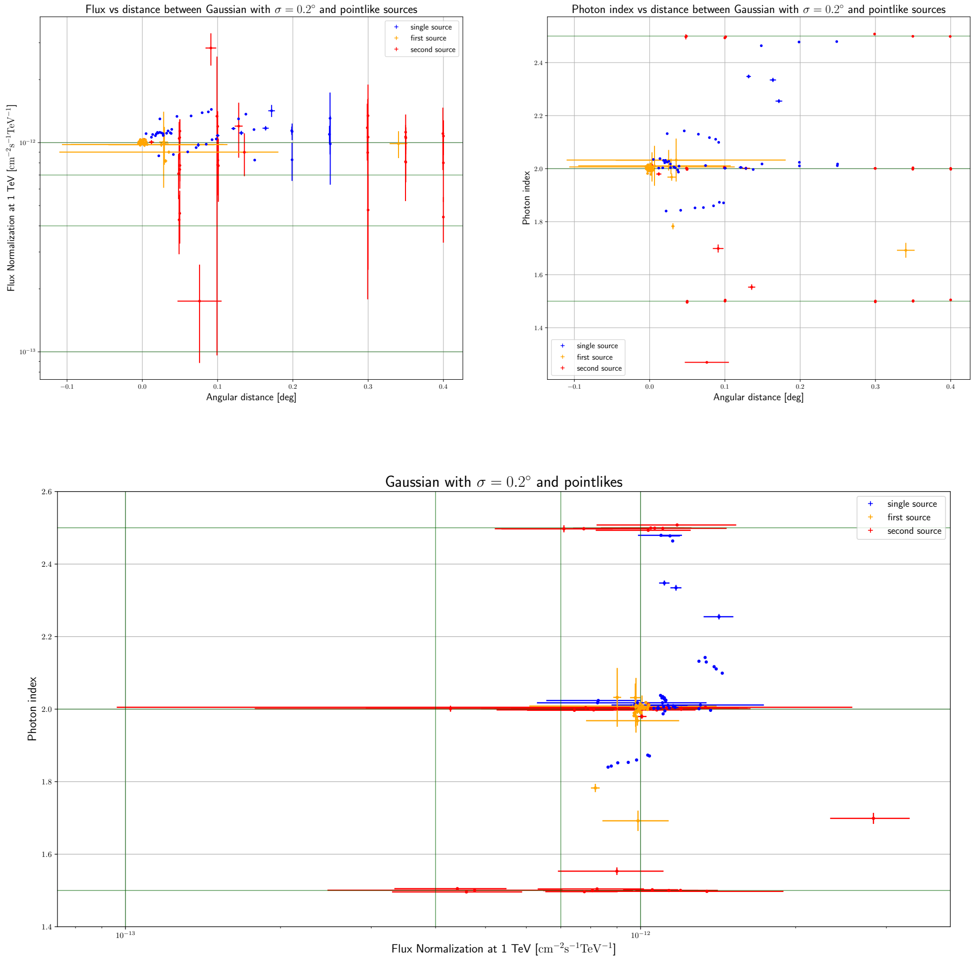


Figure D.3: Flux normalization and photon index as a function of distance between the sources (top) and photon index as a function of flux normalization (bottom). Those are results of 87 out of 96 scientific cases including a Gaussian source in the center of the pointing direction with a pointlike sources each time. The excluded cases are that with $\lesssim 20$ of successfully terminated fits. Blue data refer to the fit in which a single source has been detected; red and orange data refer to the fit in which the two sources are distinguished. The parameter values are displayed when the rate of distinguished sources is higher or lower than 68%, depending on the detection of the two sources or not. Finally, the green lines refer to the simulated values of the spectral parameters to show the agreement between those values and the fitted ones.

$1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and a photon index value that varies in a range from 1.8 to 2.5. Otherwise, when two sources are detected, the primary source has not always spectral values compatible with the original simulated ones, as well as for the second source. Flux normalization errorbars are badly estimated: this means that, even if both the pointlike and the extended source have a strong flux and they are distincted, there are significant uncertainties about their spectrum features. Another possible explanation is that, even if we consider more than 20 good fits, these can not be sufficient in some cases since errorbars are estimated from Gaussian distributions.

It is worth noting that, when the simulated flux normalization of the second source is $1 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, this value is almost never reached by the estimated parameters.

Similar conclusion can be drawn from the cases with $\sigma = 0.3°$ (Figure D.4) and $\sigma = 0.5°$ (Figure D.5), for which 83 and 63 among the 96 scientific cases simulated, respectively, are considered in the analysis.

Results regarding the scientific cases in which an extended source has been considered at the center of the pointing direction with a lower flux normalization are more problematic. Here, the fits fail in a too high number of cases and, for the few cases in which at least 10 over 100 are successful, the Gaussian and the pointlike sources are not distinguished. These results are presented only in the above Table A.4.

D.3 Extended sources

Scientific cases in which two extended sources has been considered are reported here. At first, the results related to two Gaussian sources both with $\sigma = 0.2°$ are displayed. From the plots in Figure D.6 it is possible to see how those cases are very similar to that in which both sources are pointlike. There is less variability in the estimate of the parameters with respect to the gaussian-pointlike cases, therefore, when two sources are distinguished, the values are consistent with the simulated ones.

Moreover, it is possible to make a comparison between the parameter values of the two sources detected as a single one and of the primary source is shown in Figure D.7: the errorbar referring to the detection of the primary target, when the two sources are distinguished, seems larger only because the vertical scales are smaller and the spectral parameter values are more concentrated around the original values with which the source have been simulated. In these plots the focus is on how the two sources appear

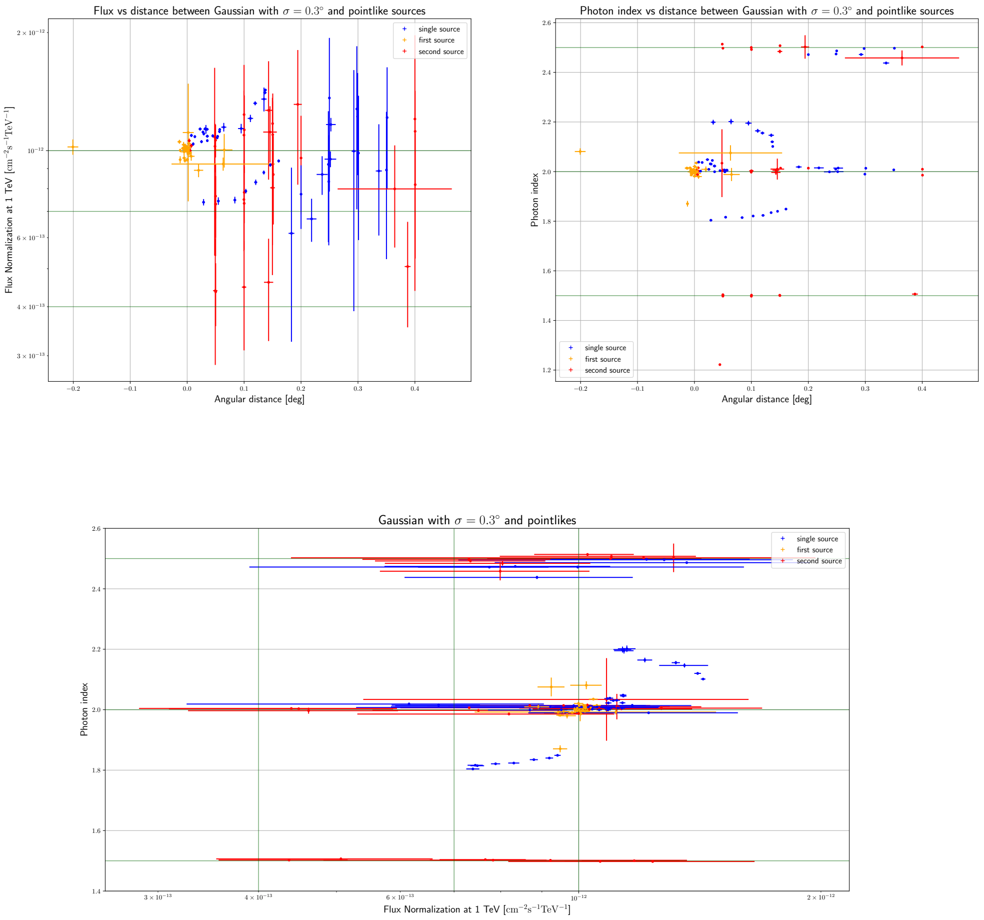


Figure D.4: Flux normalization and photon index as a function of distance between the sources (top) and photon index as a function of flux normalization (bottom). Those are results of 83 out of 96 scientific cases including a Gaussian source with $\sigma = 0.3°$ in the center of the pointing direction with a pointlike sources each time. The excluded cases are that with $\lesssim 20$ of successfully terminated fits. Blue data refer to the fit in which a single source has been detected; orange and red data refer to the fit in which the two sources are distinguished. The parameter values are displayed when the rate of distinguished sources is higher or lower than 68%, depending on the detection of the two sources or not. Finally, the green lines refer to the simulated values of the spectral parameters to show the agreement between those values and the fitted ones. Note that when the second source is simulated with the flux normalization value equal to $1 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$, this is never reached by the estimated parameters because only a single source is detected. Moreover, the estimated flux for the primary source is even more distorted compared to that of the Gaussian with $\sigma = 0.2°$.

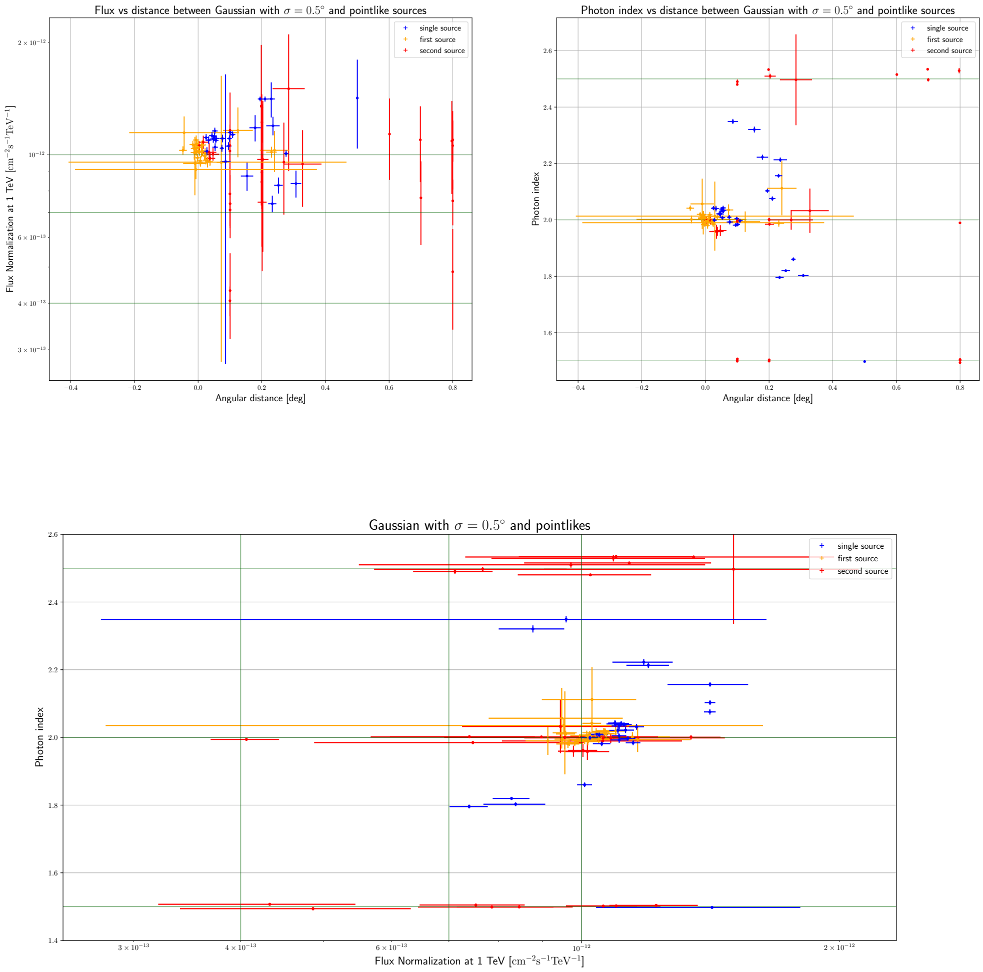


Figure D.5: Flux normalization and photon index as a function of distance between the sources (top) and photon index as a function of flux normalization (bottom). Those are results of 63 out of 96 scientific cases including a Gaussian source with $\sigma = 0.5^\circ$ in the center of the pointing direction with a pointlike sources each time. The excluded cases are that with $\lesssim 20$ of successfully terminated fits. Blue data refer to the fit in which a single source has been detected; orange and red data refer to the fit in which the two sources are distinguished. The parameter values are displayed when the rate of distinguished sources is higher or lower than 68%, depending on the detection of the two sources or not. Finally, the green lines refer to the simulated values of the spectral parameters to show the agreement between those values and the fitted ones.

The estimated flux for the primary source is even more distorted compared to that of the Gaussian with $\sigma = 0.2^\circ$ and $\sigma = 0.3^\circ$.

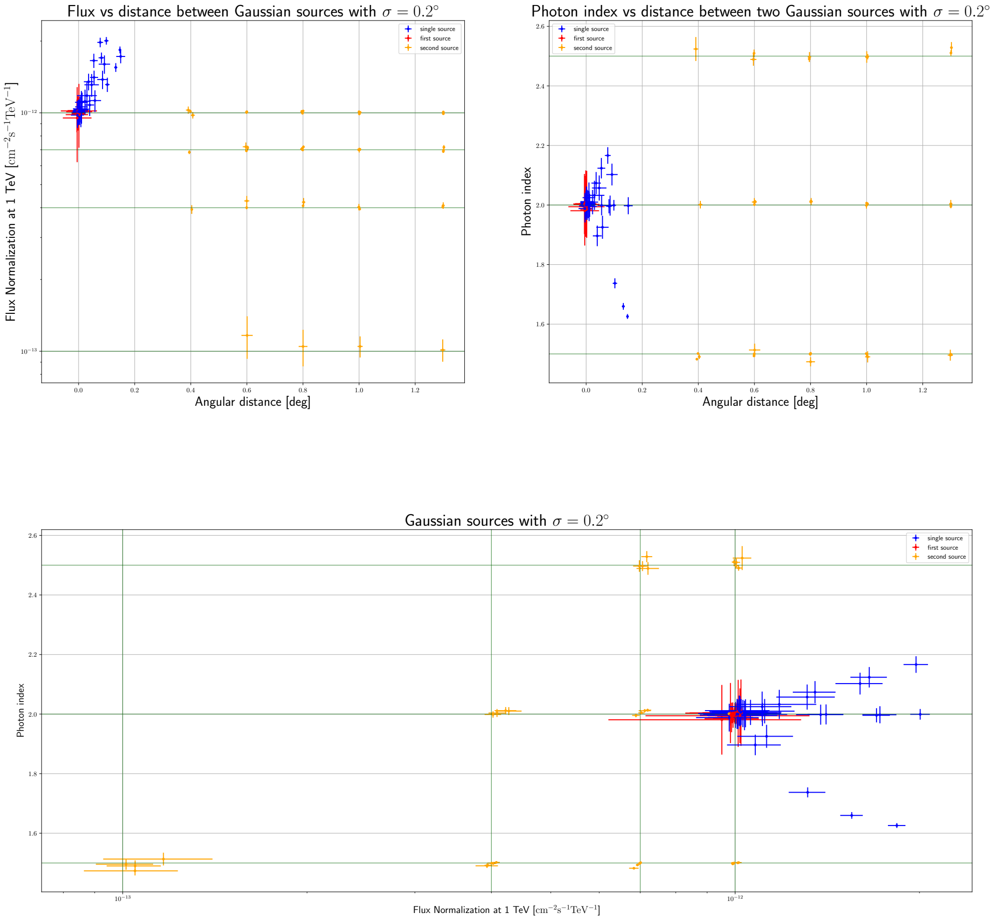


Figure D.6: Flux normalization and photon index as a function of distance between the sources (top) and photon index as a function of flux normalization (bottom). Those are results of 72 scientific cases including two Gaussian sources with $\sigma = 0.2°$ each time. Blue data refer to the fit in which a single source has been detected; red and orange data refer to the fit in which the two sources are distinguished. The values are reported with their errorbars, but blue and red errorbars are multiplied by a factor of 20 and 10, respectively, to make them more visible. The parameter values are displayed when the rate of distinguished sources is higher or lower than 68%, depending on the detection of the two sources or not. Finally, the green lines refer to the simulated values of the spectral parameters to show the agreement between those values and the fitted ones.

as a single one when they are not distinguished; in particular, the resulting spectrum seems to be that of the primary source somehow distorted, as the flux normalization is $\gtrsim 1 \times 10^{-12}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and the photon index value is almost always not so different from 2. Furthermore, the sigma value results distorted in some cases but reach at most values of 0.25°. On the contrary, when two sources are detected, the first has spectral values consistent with the simulated ones, but the sigma can result very distorted, reaching values of 0.4°.

When the extended source at the center of the pointing direction has $\sigma = 1.0°$ with a flux normalization equal to $1 \times 10^{-13}\,\mathrm{cm^{-2}s^{-1}TeV^{-1}}$ and the second has $\sigma = 0.2°$, the two targets are never distinguished. However, the fits converged successfully 45 times over 54 scientific cases simulated and the single source detected has characteristics similar to that of the second source. Therefore, the first source actually plays as a background, so that it is not recognized like a target, and influence the spectral features of the second source. This is the perfect example of how the source confusion problem can affect the detection of a target. Summarizing plots of those cases have been already reported in Figure 6.12

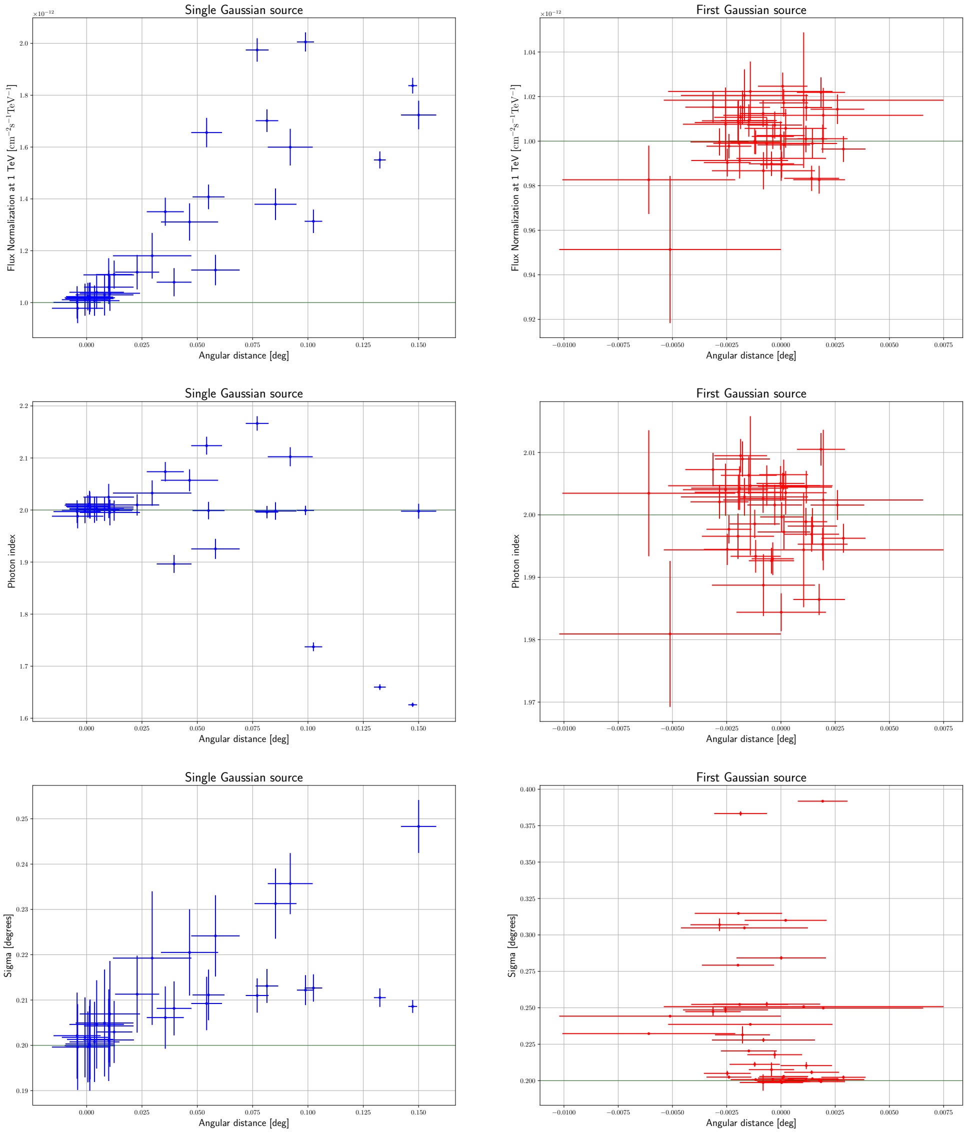


Figure D.7: Comparison between the parameter values of the two sources detected as a single one (upper panles) and of the primary source (bottom panels) for the scientific cases in which both target are Gaussian with $\sigma = 0.2°$. The parameter values are displayed when the rate of distinguished sources is higher or lower than 68%, depending on the detection of the two sources or not. Note that errorbar displayed in left panels are multiplied by a factor of 10.

Acronyms

AGN	Active Galactic Nucleus
CR	Cosmic ray
CTAO	Cherenkov Telescope Array Observatory
EAS	Extensive Air Shower
EBL	Extragalactic background light
FoV	Field of view
GC	Galactic Centre
GW	Gravitational Wave
IACT	Imaging Air Cherenkov Technique
IEM	Interstellar emission model
IR	Infra-red
IRF	Instrument Response Function
KSPs	Key Science Projects
LMC	Large Magellanic Cloud
LST	Large-Sizes Telescope
MM	Multi-messenger
MSC	Massive star cluster
MST	Medium-Sized Telescope
MWL	Multi-wavelength
PL	Power law
PMT	Photomultiplier tube
PSF	Point spread function
PWN	Pulsar wind nebula
SCT	Schwarzschild-Couder Telescope
SED	Spectral energy distribution
SNR	Supernova remnant
SiPM	Silicon Photomultiplier
SSC	Synchrotron self-Compton
SST	Small-Sized Telescope
VHE	Very high energy
UV	Ultra-violet

Bibliography

[1] J. Aasi et al. "Prospects for Localization of Gravitational Wave Transients by the Advanced LIGO and Advanced Virgo Observatories." In: (2013).

[2] J. Abadie et al. "Predictions for the rates of compact binary coalescences observable by ground-based gravitational-wave detectors". In: *Classical and Quantum Gravity* 27 (2010).

[3] H. Abdalla et al. "The H.E.S.S. Galactic plane survey". In: *Astronomy & Astrophysics* 612 (2018).

[4] S. Abdollahi et al. "The *Fermi*-LAT fourth source catalog (4FGL)". In: *ApJS* 247 (June 2020), p. 33.

[5] A. U. Abeysekara et al. "The 2HWC HAWC Observatory Gamma-Ray Catalog". In: *ApJ* 843.1 (2017), p. 40.

[6] A. Abramowski, F. Aharonian, and et al. "Acceleration of petaelectronvolt protons in the Galactic Centre." In: *Nature* 531 (2016), pp. 476–479.

[7] A. Abramowski et al. "Discovery of extended VHE gamma-ray emission from the vicinity of the young massive stellar cluster Westerlund 1". In: *Astronomy & Astrophysics* 537 (2012), A114.

[8] F. Acero and et al. "Development of the model of galactic interstellar emission for standard point-source analysis of *Fermi* large area telescope data". In: *ApJS* 223 (2016), p. 26.

[9] M. Ackermann et al. "A Cocoon of Freshly Accelerated Cosmic Rays Detected by Fermi in the Cygnus Superbubble". In: *Science* 334 (2011).

[10] F. Aharonian, R. Yang, and E. de Oña Wilhelmi. *Massive Stars as Major Factories of Galactic Cosmic Rays*. 2019.

[11] M. Ajello et al. "3FHL: The Third Catalog of Hard Fermi-LAT Sources". In: *ApJS* 232 (2017), p. 23.

[12] B. Bartoli et al. “Identification of the Tev Gamma-ray source ARGO J2031+4157 with the Cygnus Cocoon”. In: *The Astrophysical Journal* 790 (2014).

[13] B.M. Baughman. “Status and updates from the High Altitude Water Cherenkov (HAWC) Observatory”. In: *Nuclear Physics B - Proceedings Supplements* 239-240 (2013), pp. 220–224.

[14] A. R. Bell et al. “Cosmic-ray acceleration and escape from supernova remnants”. In: *Monthly Notices of the Royal Astronomical Society* 431.1 (2013), pp. 415–429.

[15] K. Bernlöhr et al. “Monte Carlo design studies for the Cherenkov Telescope Array”. In: *Astroparticle Physics* 43 (2013), pp. 171–188.

[16] R. Brose et al. “Cosmic-ray acceleration and escape from post-adiabatic supernova remnants”. In: *Astronomy & Astrophysics* 634 (2020), A59.

[17] A.M. Bykov et al. “High-Energy Particles and Radiation in Star-Forming Regions.” In: *Space Sci Rev* 216 (2020), p. 42.

[18] Z. Cao et al. “Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 gamma-ray Galactic sources.” In: *Nature* 594 (2021), pp. 33–36.

[19] S. Carrigan et al. *The H.E.S.S. Galactic Plane Survey - maps, source catalog and source population*. 2013.

[20] LHAASO collaboration. *Performance of LHAASO-WCDA and Observation of Crab Nebula as a Standard Candle*. 2021.

[21] The CTA Consortium. “Science with the Cherenkov Telescope Array”. In: *World Scientific* (Feb. 2018).

[22] gamma-cat database. URL: `https://gamma-cat.readthedocs.io`.

[23] TeVCat database. URL: `http://tevcat.uchicago.edu/`.

[24] A. De Angelis and M. Pimenta. *Introduction to Particle and Astroparticle Physics*. (2nd ed.) Springer, 2018.

[25] T. Deckers et al. “Observation of TeV Gamma Rays from Markarian 501 with the First HEGRA Cherenkov Telescope”. In: *ICRC1997*. Vol. 3. 1997, p. 245.

[26] C. Deil et al. “Gammapy - A prototype for the CTA science tools”. In: *35th International Cosmic Ray Conference (ICRC2017)*. Vol. 301. Jan. 2017, p. 766.

[27] G. Di Sciascio. “Ground-based Gamma-Ray Astronomy: an Introduction”. In: *Journal of Physics Conference Series*. Vol. 1263. June 2019.

[28] S. Funk et al. *CTA first science data challenge 1 (DC-1) close-out document*. 2019.

[29] S. Gabici and F. A. Aharonian. "Searching for Galactic Cosmic-Ray Pevatrons with Multi-TeV Gamma Rays and Neutrinos". In: *The Astrophysical Journal* 665.2 (2007), L131–L134.

[30] Gammapy Developers. *Gammapy Documentation*. Version 0.18.2. Nov. 19, 2020. URL: `https://docs.gammapy.org/0.18.2`.

[31] T. P. Li and Y. Q. Ma. "Analysis methods for results in gamma-ray astronomy". In: *ApJ* 272 (1983), pp. 317–324.

[32] M.S. Longair. *High Energy Astrophysics*. (3rd ed.) Cambridge University Press, 2011, pp. 493–496.

[33] A. Mitchell et al. "Predictions for gamma-rays from clouds associated with supernova remnant PeVatrons". In: *PoS* ICRC2021 (2021).

[34] Q. Remy et al. "Survey of the Galactic Plane with the Cherenkov Telescope Array". In: *PoS* ICRC2021 (2021).

[35] Frank M. Rieger, Emma de Ona-Wilhelmi, and Felix A. Aharonian. *TeV Astronomy*. 2013.

[36] M. Spurio. *Probes of Multimessenger Astrophysics*. (2nd ed.) Springer, Dec. 2018.

[37] I. M. Stewart. "Maximum-likelihood detection of sources among Poissonian noise". In: *Astronomy & Astrophysics* 495 (2009), pp. 989–1003.

[38] CTA webpage. URL: `https://www.cta-observatory.org`.

[39] H.E.S.S. webpage. URL: `https://www.mpi-hd.mpg.de/hfm/HESS`.

[40] MAGIC webpage. URL: `http://magic.mppmu.mpg.de`.

[41] VERITAS webpage. URL: `https://veritas.sao.arizona.edu`.

[42] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Ann. Math. Statist.* 9 (1938), pp. 60–62.

[43] R. Zanin and J. Holder. "Observing the Galactic Plane with the Cherenkov Telescope Array". In: *PoS* ICRC2017 (2017), p. 740.

[44] P.A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020).

Ringraziamenti

Ringrazio il mio relatore, prof. Vignali, e la mia correlatrice, dott.ssa Zanin, per la possibilità che mi hanno dato di realizzare questa tesi e la pazienza spesa nel seguirmi in questo percorso. Ringrazio anche l'INAF di Bologna per avermi permesso di accedere alle macchine in remoto e velocizzare quindi il processo di simulazione e analisi dati.

Ringrazio i miei genitori, che, nonostante l'incomprensione verso le materie di mio interesse, mi hanno sempre motivata ed è grazie a loro che ho potuto portare avanti il mio sogno di diventare Astrofisica. Grazie alle mie sorelle, sulle quali posso sempre contare, e a nonni, cugini, zie, zii e parenti che mi hanno accompagnato in un modo o nell'altro in questo percorso.

Grazie a Benedetta, l'amica di una vita. Anche se spesso lontana, "tu ci sei sempre stata".

A Luca, che nel tempo di elaborazione di questa tesi mi ha visto sbalzare dall'euforia più grande alla delusione più profonda e che mi ha sostenuto sempre aiutandomi a superare ostacoli mentali che mi costruivo da sola, grazie.

Grazie ad Asia e Laura, che anche se conoscendomi soltanto dal liceo, mi hanno sempre aiutata a credere nel mio sogno.

Ringrazio anche la mia maestra delle elementari, che mi ha fatto scoprire l'astronomia da piccola, e il mio professore di Matematica e Fisica del liceo, che mi ha inizialmente indirizzato e motivato per la scelta della triennale in Fisica, senza la quale non sarei arrivata fin qua.

Ricordare qui tutti i nomi di chi più, chi meno mi ha affiancato in questo percorso da quando ne fantasticavo da piccola renderebbe questa pagina troppo lunga, ma anche a tutti loro va il mio pensiero e ringraziamento.