Exploring active galactic nuclei at extreme energies : analysis and modeling of multi-wavelength flares and preparation of CTA

Anton Dmytriiev

Université de Paris
Ecole Doctorale “Astronomie et Astrophysique d’Île-de-France” (N° 127)

Observatoire de Paris
Laboratoire Univers et Théories (LUTH)

Exploring Active Galactic Nuclei at extreme energies: analysis and modeling of multi-wavelength flares and preparation of CTA

par

Anton Dmytriiev

Thèse de doctorat en astronomie et astrophysique

Dirigée par **Hélène SOL** et par **Andreas ZECH**

Présentée et soutenue publiquement le **27 novembre 2020** devant un jury composé de:

Stéphane CORBEL, Professeur, Université de Paris, *Président*

Paula CHADWICK, Professeur, Université de Durham (Royaume-Uni), *Rapporteur*

Gilles HENRI, Professeur, Université Grenoble Alpes, *Rapporteur*

Markus BÖTTCHER, Professeur, Université du Nord-Ouest (Afrique du Sud), *Examinateur*

Bohdan HNATYK, Professeur, Université nationale Taras-Chevtchenko de Kiev (Ukraine), *Examinateur*

Hélène SOL, Professeur, Observatoire de Paris, *Directrice de thèse*

Andreas ZECH, Maître de conférence, Observatoire de Paris, *Co-Directeur de thèse*

Abstract

Titre: Exploration des Noyaux Actifs de Galaxies aux énergies extrêmes: analyse et modélisation des sursauts multi-longueurs d'onde et préparation du CTA

Résumé: De nombreuses questions liées à la physique des jets des Noyaux Actifs de Galaxies restent ouvertes. Une classe particulière d'AGN, les blazars, a un jet pointant vers la Terre. Une telle orientation du jet nous permet de sonder une riche variété de phénomènes physiques mal compris sur les écoulements relativistes. Les blazars montrent une émission non thermique, provenant du jet, qui est très variable sur tout le spectre électromagnétique, des radiofréquences aux rayons gamma du TeV. Le flux d'énergie peut augmenter d'un ordre de grandeur sur des échelles de temps aussi courtes que quelques minutes, un phénomène appelé "sursaut" (flare), et aussi longues que des mois ou même des années. Malgré la quantité croissante de données disponibles sur plusieurs longueurs d'onde (multi-wavelength, MWL), l'origine et les mécanismes physiques derrière les sursauts fréquemment observés dans les blazars ne sont toujours pas bien compris. De nombreuses tentatives ont été faites pour décrire les flares avec différents modèles d'émission, mais les propriétés détaillées de l'évolution temporelle des flux dans différentes bandes spectrales restent difficiles à reproduire. Afin d'identifier les processus physiques impliqués lors des sursauts de blazars, j'ai développé un code radiatif polyvalent, basé sur un traitement dépendant du temps de l'accélération des particules, de l'échappement et du refroidissement radiatif. Le code calcule l'évolution dans le temps de la fonction de distribution des électrons dans la zone d'émission du blazar et le spectre de l'émission Synchrotron Self-Compton (SSC) par ces électrons. J'ai appliqué le code à un sursaut multi-lambda géant du blazar Mrk 421, représentant de la classe des BL Lacertae, qui est le sursaut le plus brillant détecté jusqu'ici en provenance de cette source. Dans notre approche, nous considérons le sursaut comme une perturbation modérée de l'état de flux stationnaire et recherchons des interprétations avec un nombre minimum de paramètres libres. En conséquence, j'ai développé un nouveau scénario physique de l'activité observé pendant le sursaut, qui décrit l'ensemble des données, comprenant des spectres à l'état haut de la source dans différentes gammes d'énergie, et des courbes de lumière multi-lambda du domaine optique aux rayons gamma VHE. Dans ce scénario, le processus déclenchant le sursaut est l'accélération des particules par un processus de type Fermi du second ordre, dû à la turbulence qui emerge au voisinage de la région d'émission stationnaire du blazar.

Dans cette thèse, j'ai également effectué une analyse des données du High Energy

Stereoscopic System (H.E.S.S.) de deux sursauts géants du blazar 3C 279, représentant de la classe des Flat Spectrum Radio Quasars (FSRQ).

Enfin, j'ai contribué à la préparation du Cherenkov Telescope Array (CTA), qui est un observatoire de rayons gamma au sol de nouvelle génération, dont l'entrée en service est prévue à partir de 2022. L'instrument, qui est actuellement en cours de développement, aura des performances considérablement améliorées par rapport aux Imaging Atmospheric Cherenkov Telescopes (IACTs) qui sont actuellement en fonctionnement, y compris une couverture spectrale sans précédent de quelques dizaines de GeV à ∼300 TeV. Dans le cadre du CTA, j'ai effectué des simulations de performances optiques du Gamma-Ray Cherenkov Telescope (GCT), l'un des trois modèles proposés de télescopes de petite taille (SST) pour CTA. De plus, en utilisant les observations d'étoiles brillantes effectuées par le prototype de télescope installé sur le site de l'Observatoire de Paris à Meudon, j'ai étudié l'effet de la micro-rugosité des miroirs du télescope sur la fonction d'étalement du point (PSF) et calculé le niveau de qualité de polissage des miroirs requis pour optimiser les performances.

Mots clefs: Noyaux Actifs de Galaxies ; Cherenkov Telescope Array ; sursauts des blazars ; modélisation d'émission ; accélération des particules ; performances optiques ; analyse des données de rayons gamma ; High Energy Stereoscopic System; Mrk 421 ; 3C 279

Title: Exploring Active Galactic Nuclei at extreme energies: analysis and modeling of multi-wavelength flares and preparation of CTA

Abstract: Many questions related to the physics of jets of Active Galactic Nuclei remain open. A particular subclass of AGN, blazars, have a jet pointing towards the Earth. Such suitable orientation of the jet allows us to probe a rich variety of poorly understood physical phenomena related to relativistic outflows. Blazars show non-thermal emission, originating from the jet, which is highly variable across the entire electromagnetic spectrum, from radio frequencies to TeV γ-rays. The energy flux can enhance by an order of magnitude on time-scales as short as minutes, a phenomenon referred to as a "flare", and as long as months or even years. Despite the growing amount of available multi-wavelength (MWL) data, the origin and the physical mechanisms behind the frequently observed flaring events in blazars are still not well understood. Many attempts have been made to describe the flares with different emission models, but detailed properties of flux variation patterns (light curves) in different wavebands remain difficult to reproduce. In order to identify physical processes that are involved during blazar outbursts, I have developed a versatile radiative code, based on a time-dependent treatment of particle acceleration, escape and radiative cooling. The code computes time evolution of the distribution function of electrons in the blazar emitting zone and the spectrum of the Synchrotron Self-Compton (SSC) emission by these electrons. I applied the code to a giant MWL flare of the blazar Mrk 421, a representative of the BL Lacertae class, which is the brightest VHE flare ever detected from this source. In our approach, we consider the flare as a moderate perturbation of the quiescent state and search for interpretations with a minimum number of free parameters. As a result, I developed a novel physical scenario of the flaring activity that describes the data set, comprising spectra in the high state of the source in different energy ranges, and MWL light curves from the optical domain to the VHE γ-ray band. In this scenario, the process initiating the outburst is the second-order Fermi acceleration of particles due to turbulence arising in the vicinity of the blazar stationary emission region.

In this thesis, I also performed analysis of High Energy Stereoscopic System (H.E.S.S.) data of two giant flares of the blazar 3C 279, a representative of the Flat Spectrum Radio Quasars (FSRQ) class.

Finally, I contributed to preparation of Cherenkov Telescope Array (CTA), which is a new-generation ground-based γ-ray observatory, expected to start operations in 2022. The instrument, which is presently under development, will have greatly improved performance compared to currently operating Imaging Atmospheric Cherenkov Telescopes (IACTs), including unprecedented spectral coverage from a few tens of

GeV to ∼300 TeV. In the context of CTA, I performed simulations of the optical performance of the Gamma-Ray Cherenkov Telescope (GCT), one of the three proposed designs of Small-Size Telescopes (SST) for CTA. Also, using the observations of bright stars done by the telescope prototype installed on the site of Paris Observatory in Meudon, I studied the effect of micro-roughness of the telescope mirrors on the point spread function (PSF) and calculated the level of the mirror polishing quality required to optimize the performances.

Keywords: Active Galactic Nuclei ; Cherenkov Telescope Array ; blazar flares ; emission modeling ; acceleration of particles ; optical performance ; gamma-ray data analysis ; High Energy Stereoscopic System ; Mrk 421 ; 3C 279

Contents

List of Figures

List of Tables

Acronyms

ADAF Advection-Dominated Accretion Flow

AGN Active Galactic Nuclei

ALP Axion-Like Particle

ASM All-Sky Monitor

ATOM Automatic Telescope for Optical Monitoring

BAT Burst Alert Telescope

BH Black Hole

BLR Broad Line Region

BRDF Bidirectional Reflectance Distribution Function

CAT Cherenkov Array at Thémis

CHEC Compact High Energy Camera

CMB Cosmic Microwave Background

CR Cosmic Ray

CTA Cherenkov Telescope Array

DM Dark Matter

EBL Extragalactic Background Light

EC External Compton

EHT Event Horizon Telescope

EMBLEM Evolutionary Modeling of BLob EMission

FACT First Geiger-mode Avalanche photodiode Cherenkov Telescope

FR I/II Fanaroff-Riley type I/II

FSRQ Flat Spectrum Radio Quasar

GCT Gamma-ray Cherenkov Telescope

GRB Gamma-Ray Burst

GRMHD General Relativistic MagnetoHydroDynamic

HAGAR High Altitude GAmma Ray telescope

HAP H.E.S.S. Analysis Package

HBL High-frequency peaked BL Lac

HEGRA High Energy Gamma Ray Array

HESS High Energy Stereoscopic System

IACT Imaging Atmospheric Cherenkov Telescope

IBL Intermediate BL Lac

IC Inverse Compton

IGMF Intergalactic Magnetic Field

IR Infrared

IRF Instrument Response Function

LAT Large Area Telescope

LBL Low-frequency peaked BL Lac

LC Light Curve

LHC Large Hadron Collider

LIV Lorentz Invariance Violation

LST Large-Sized Telescope

MAGIC Major Atmospheric Gamma-ray Imaging Cherenkov telescopes

MAXI Monitor of All-sky X-ray Image

MC Monte Carlo

MJD Modified Julian Date

MST Medium-Sized Telescope

MWL Multi-WaveLength

NLR Narrow Line Region

NRAO National Radio Astronomy Observatory

NSB Night Sky Background

OVRO Owens Valley Radio Observatory

PCA Proportional Counter Array

PIC Particle-In-Cell

PMT PhotoMultiplier Tube

PSF Point Spread Function

QED Quantum ElectroDynamics

QSO Quasi-Stellar Object

ROBAST ROOT-BAsed Simulator for ray Tracing

RXTE Rossi X-ray Timing Explorer

S-C Schwarzschild-Couder

SED Spectral Energy Distribution

SM Standard Model

SMBH Super-Massive Black Hole

SiPM Silicon PhotoMultiplier

SSC Synchrotron Self-Compton

SSRQ Steep Spectrum Radio Quasar

SST Small-Sized Telescope

TACTIC TeV Atmospheric Cherenkov Telescope with Imaging Camera

ToO Target of Opportunity

UV UltraViolet

VERITAS Very Energetic Radiation Imaging Telescope Array System

VHE Very High Energy

VLA Very Large Array

VLBA Very Long Baseline Array

VLBI Very-Long-Baseline Interferometry

WIMP Weakly Interacting Massive Particle

XRT X-Ray Telescope

Acknowledgments

So, here comes the end of my incredible three-year PhD adventure. I would like to thank all those people who made this dissertation possible, and all those who were making history and participated in writing of a fascinating chapter "On The Astrophysics PhD Secret Service" in the book of my life.

First of all, I wish to express my deepest gratitude to my thesis advisors Hélène Sol and Andreas Zech, for guiding me along the winding path of my PhD, for their patience, continuous encouragement and countless valuable words of advice. Hélène and Andreas gave me a profound belief in my abilities, and they were not merely scientific supervisors, but also, in a way, mentors, and taught me to be independent, flexible and deal with difficulties and challenges. I truly appreciate their exceptional involvement and dedication, which allowed me to always feel confident and motivated on my way to the PhD degree, even when the road was tough. Special thanks also to Hélène for giving me a ride to downtown Meudon from time to time and interesting and thought-provoking discussions on the way. These rides allowed me to distract and remain in good humor during the intense months of PhD manuscript writing.

I would like also to pay my special regards to the LUTH for having hosted me for these three wonderful years and for a very pleasant working and social atmosphere. This really nice vibe would have been impossible without all the great people working at the laboratory. In particular, I wish to thank my colleagues Catherine Boisson, Mathieu Servillat, Zakaria Meliani and Michael Zacharias for interesting exchanges, your useful tips, and for many insightful discussions around a cup of coffee or tea. Many thanks also to Nicolas Renault-Tinacci for engaging conversations during the times when we were sharing the office in 2018-2019. I am also grateful to PhD students Christelle with whom I shared my office (on average) one day per week, as well as Gaëtan and Jordan, for enriching exchanges of ideas, as well as entertaining chats. Having all these interactions very much inspired and motivated me, and I was lacking them a lot during the lockdown in spring 2020. In addition, I would like to express my gratitude to the coffee machine in the LUTH kitchen, that day after day filled me with energy and powered my work, vitalized me, boosted my mood and productivity, and helped to get me through the day. I also wish good luck to the new PhD student Anna, my successor, who is taking over for the blazar flares modeling and will continue my work, extending and improving my developments.

I would also like to thank Gabriel Emery, who helped me to get started and learn the basics of H.E.S.S. data analysis software. For my work on CTA preparation, particularly helpful were colleagues from GEPI laboratory Oriane Le Blanc, Jean-Michel Huet, Philippe Laporte, Lucie Dangeon, Johann Gironnet, Fatima De Frondat and Gilles Fasola. I very much appreciate also numerous helpful pieces of advice and

practical suggestions provided by Martin Lemoine in the context of my modeling of blazar flares, which gave me invaluable insight into particle acceleration processes.

During my fascinating trip to Namibia in July 2018 for the shift on the H.E.S.S. site, I met some truly wonderful people, thanks to whom this one-month adventure remains a vivid memory. I wish to thank Volker, Albert and Frikkie for their guidance and assistance during the array operation. I am also sincerely grateful to my co-shifter Cornelia for her satirical remarks and numerous teachings that have helped to keep me alert and more organized. Thanks also Cornelia for giving me rides to the supermarket in Windhoek and for the exciting trip to Solitaire.

I am also deeply indebted to the committee members Markus Böttcher, Stéphane Corbel and Bohdan Hnatyk for having accepted this task, and especially to the referees Paula Chadwick and Gilles Henri who brilliantly coped with the challenging task of reading in detail my dissertation of more than 200 pages in just one month.

I cannot leave Paris without mentioning my friends Julien, William, Yanomi, Bryan, Alexandre and Lola (and many others!) that I met in this city at an amazing event named "Bla-Bla language exchange". I am grateful to all of you for organizing and being part of this event, and for lots of unforgettable moments we had together. Thanks to you, I was always looking forward to Thursday night to come to this event and have a great time with you. Many thanks also to Saad and Valentine, my former roommates in Geneva, who fortunately also came to Paris for their studies or work, so that our friendship continued and we were able to share many delightful moments.

Special thanks go also to my Ukrainian friends Pavlo, Andrew, Serhii, Anastasia, Yaroslav, George, Maria, Alina, Artem and Valentyna. I am deeply grateful to all of you for sharing joyful moments with me and supporting and cheering me up when I was facing difficulties. Thanks to your genuine dedication, I never felt lonely, and was never afraid of difficulties, because I knew that I can always rely on you.

I would also like to extend my sincere thanks to Victoria and Héloïse, the daughters of my landlord Stéphane, for many amusing live conversations in the shared garden during the lockdown in spring 2020. This, the only face-to-face communication for me in those days, helped me to not feel alone and remain in good spirits. I gratefully acknowledge also Jacqueline and her husband Roland for having hosted me at their house during the last 4 months of my PhD.

Finally, I am eternally grateful to my parents Oksana and Oleksandr for their constant and invaluable support, encouragement and guidance, as well as their unwavering belief in my success. Thanks to you, I always had a sense of inexhaustible optimism and self-confidence.

To be honest, I am certainly feeling a little sad about leaving Paris, but there are still lots of exciting opportunities and adventures in the future, many uncharted territories to explore, and all the most interesting experiences are still ahead.

I would like to dedicate this work to my parents, Oleksandr and Oksana

– *"Nature uses only the longest threads to weave her patterns, so that each small piece of her fabric reveals the organization of the entire tapestry."*

Richard Feynman

Chapter 1

Introduction

Many astrophysical objects work as powerful particle accelerators. This fact is established, first of all, from the detection of non-thermal emission, in particular, of γ-rays arriving from distant sources. One of the most fascinating and extreme objects of this type are Active Galactic Nuclei (AGN) – compact and highly luminous regions in the cores of certain types of galaxies, characterized by a range of phenomena caused by the activity of a central supermassive black hole. AGN are the most luminous persistent phenomena in the Universe, observed up to the highest photon energies achievable with current instruments (several tens of TeV). Some AGN eject highly collimated relativistic outflows of plasma – *jets*, extending from the central core for distances up to tens, hundreds or even thousands of kiloparsecs. AGN jets are enthralling phenomena that manifest themselves throughout the entire electromagnetic spectrum, from radio frequencies to Very High Energy (VHE) γ-ray range ($E_\gamma > 100$ GeV), indicating acceleration of particles to *at least* TeV energies. Furthermore, AGN are considered as one of the candidates of sources producing Ultra High Energy Cosmic Rays (UHECR): these objects are suspected to boost protons to energies up to 10^{20} eV, which is some seven orders of magnitude higher than achieved at the Large Hadron Collider (LHC). Extreme environments of AGN are therefore natural laboratories of plasma and high-energy physics, allowing to explore regimes unreachable at Earth-based facilities.

A lot of questions related to the physics of AGN jets remain unanswered, e.g. matter content (purely leptonic or lepto-hadronic?), particle acceleration and emission mechanisms at work, jet formation, etc. Particularly suitable for studies of the poorly understood physics of relativistic jets, are *blazars* – AGN with a jet, which happens to be very closely aligned with the line of sight. Their emission is dominated by the non-thermal radiation of the jet. Blazars show strong variability in all frequency bands

from the radio domain up to TeV γ-ray regime. The energy flux can experience an increase by a factor of $\sim$ 10, or even more, over an impressively wide range of variability time-scales: from high flux states lasting a few months or even years, down to dramatic flux variations over as short as $\sim$1 minute. These spectacular phenomena are referred to as "*flares*" for the variability proceeding at a time-scale of less than $\sim$ 1 week, and "high activity/flux states" for longer time-scales. Despite the fact that more and more observational data of blazar flares is collected in different spectral bands, the nature of the flaring activity and physical processes triggering it remain obscure. Especially puzzling is the origin of the most rapid variability proceeding at $\sim$ 1 minute time-scale, as it implies a size of the γ-ray emitting zone which is smaller than the radius of the event horizon of the central black hole. Various scenarios are proposed to explain blazar flaring behavior: shock waves passing through the jet, spontaneous generation of strong turbulence, jet bending and even stars crossing the jet, etc.

The key method to get an insight into violent processes in the jets that are responsible for launching flares, is *physical modeling* of the observed behavior of the blazar emission during the outburst. In order to identify the underlying physical processes as precisely and unambiguously as possible, one needs to maximize the number of observational constraints. This implies measurement of spectral and timing properties of the source's emission during the flare in different energy bands, i.e. spectra in different flux states in different energy ranges, and multi-wavelength (MWL) light curves, which requires to coordinate quite a large number of instruments to organize MWL campaigns. Of a special interest are VHE γ-ray flares, as they carry information about poorly known processes involving particles with the highest energies, and allow to probe phenomena occurring on the shortest time-scales and the smallest spatial scales. Flares at TeV energies are typically accompanied by their counterparts at lower energies. Self-consistent *time-dependent* modeling of the observed varying MWL emission is the crucial approach in order to test various scenarios of flaring activity.

A significant progress in our understanding of various high-energy phenomena in the Universe, and in particular of the blazar flaring behavior, is expected with the advent of the Cherenkov Telescope Array (CTA). This future telescope system, currently under development, will have substantially improved performances compared to present-day instruments sensitive in the VHE γ-ray band, including order of magnitude higher flux sensitivity and an extended spectral range from $\sim$ 30 GeV to $\sim$ 300 TeV. In order to achieve ambitious scientific goals set by CTA, it is crucial that the performance of telescopes of the array is well characterized and optimized.

In this thesis, we focus on analysis, interpretation and physical modeling of MWL

data of flares of two blazars, 3C 279 and Mrk 421, as well as perform preparatory studies for CTA. The manuscript is organized as following. In Chapter 2 a general introduction to AGN with a focus on blazars is made. In Chapter 3 we present our analysis and interpretation of H.E.S.S. data of two giant outbursts of the blazar 3C 279. Chapter 4 and Chapter 5 are devoted to blazar emission models. In Chapter 4 we focus on emission mechanisms and in Chapter 5 we first cover various physical processes that are thought to operate during blazar flares, then present the general time-dependent flare model and the associated numerical code "EMBLEM" that we developed for the modeling task. In Chapter 6 we apply our code to a MWL data set of the brightest VHE flare of the blazar Mrk 421 up to now, and develop a novel physical scenario to describe the variability characteristics during the outburst. Next, Chapter 7 is dedicated to preparation and development of CTA, in which we present characterization of the performance of one of the optical designs proposed for Small-Sized Telescopes sub-array of CTA. Finally, in Chapter 8 concluding remarks are made and perspectives are discussed.

Chapter 2

Active Galactic Nuclei

Active Galactic Nuclei (AGN) are the cores of galaxies which generate much higher amounts of energy than observed in normal galaxies, which cannot be explained by the activity of stars. These objects appear as extremely bright regions in the centers of certain types of galaxies, and are characterized by very high luminosity, fast and violent variability, in some cases, intense radio emission, strong and broad emission lines in the optical spectra, and broad-band continuous spectra stretching over a much wider domain than the one of normal galaxies. AGN also display a range of spectacular phenomena, not seen in normal galaxies, e.g. accretion disks, large-scale jets, etc.

AGN are one of the most remarkable astronomical objects in the Universe. First of all, because of their enormous luminosity: AGN are the most energetic non-transient, sustained phenomena in the Universe. Secondly, AGN are the most efficient machines for conversion of mass to energy in Nature: e.g. combustion of natural gas leads to a release of only $\sim 10^{-8}$ % of the rest energy of the fuel, nuclear power plants performing fission of Uranium-235 release only $\sim 0.09\%$ of the fuel rest energy, thermonuclear fusion $4p \rightarrow He$ in the Sun core has a $\sim 0.7\%$ yield, whereas AGN are able to convert to energy up to 42% of the matter rest mass! Even more astonishing is that, as we will see later, such huge energy release occurs thanks to gravity, which is the weakest force among the four fundamental ones in Nature, at the same time the reactions involving much stronger forces, such as electromagnetic interaction (combustion and chemical reactions in general), and strong interaction (fusion and fission) results in a way more modest efficiency of the energy output. But what exactly powers an AGN, how these objects extract the energy and what defines their observed properties?

While the very first observations of AGN date to over a century ago, only in the

1960s it became clear that the exceptionally high luminosities of these sources are not produced by thermonuclear fusion. Moreover, the observed very short variability time-scales, implied according to the causality condition that the energy of an AGN is derived from a very compact region. Finally, very broad spectral energy distributions pointed to the non-thermal origin of the observed emission. Only later, after collecting different observational pieces of evidences, it was understood that AGN are powered by a supermassive black hole. The observed peculiar features and phenomena in these objects are explained by the activity of the central black hole, with the activity driven by accretion of matter on the black hole.

Apart from very interesting physics of AGN and related phenomena, these objects, being extremely bright, can serve as beacons and carry highly valuable information from very distant locations in the Universe, as well as allow to probe the medium in-between. Overall, the complex nature of AGN and extreme physical conditions in these sources, make them very attractive targets for studies of a whole wealth of various open questions.

In this chapter, we provide a general introduction into Active Galactic Nuclei (AGN), with a special focus on AGN jets. We next consider in detail one of AGN classes, *blazars*, of which two representatives are studied in this thesis (3C 279 in Chapter 3 and Mrk 421 in Chapter 6). Further emphasis is put on γ-ray emission of blazars, together with an overview of broad-band emission models.

2.1 The AGN zoo

Over the XXth century, astronomers discovered a sizable number of galaxies showing noticeable activity, which however manifested in a variety of ways. All the objects shared a number of peculiar AGN-inherent properties (abnormally high luminosities, variability, unusual broad-band spectra, etc.), but at the same time featured important differences in terms of the luminosity level, presence/absence of spectral lines, radio emission, jet, etc. This lead to a division of AGN into a number of different classes. It took a few decades to understand that all these apparently different objects had the same underlying nature, with a global view represented by a so-called *unification scenario*. The two major categories of AGN are **radio-quiet** and **radio-loud**, with the division based on the level of the radio flux. Only around 10% of AGN are radio-loud. The radio-loud and radio-quiet AGN are also quite diverse in terms of their characteristics other than radio-loudness, and are in turn sub-divided into narrower classes, based on various properties identified in early observations. In the resulting AGN zoo, the following classes are distinguished (see the classification

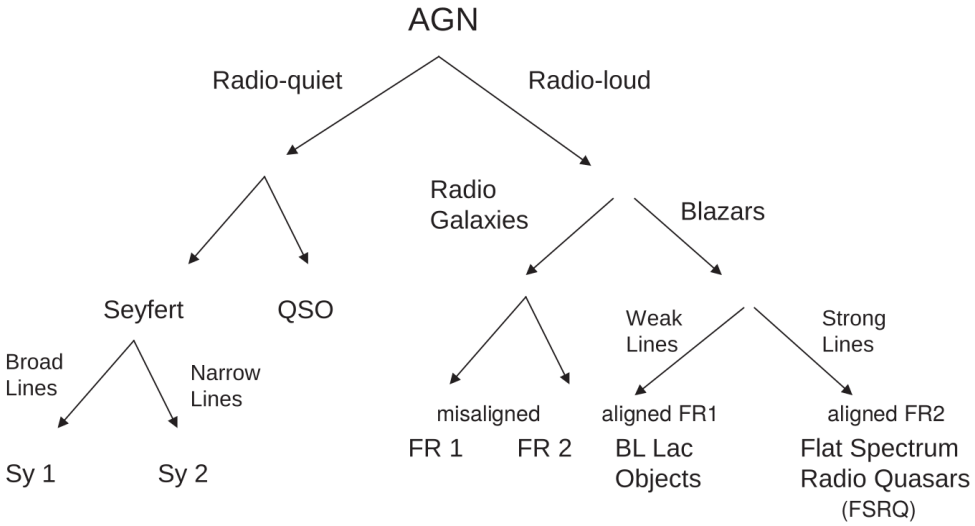


Figure 2.1: Observational classification of AGN. The division is based on the properties, such as the level of radio flux and presence of optical lines in the spectra. (adapted from Dermer & Giebels (2016))

scheme in Fig. 2.1):

- *Seyfert galaxies*: the (arguably) first observed AGN[1]. Discovered in the 1940s by C. Seyfert, who observed a number of apparently normal spiral galaxies, which displayed a nucleus resembling to a stellar-like object, with a very unusual spectrum showing strong and surprisingly broad emission (rather than absorption) lines. Interpretation of the broadening with a Doppler effect indicated that the emitting matter moves with velocities in the 10^3 km/s range, which is much faster than typical rotational velocities. Seyfert galaxies are divided into Seyfert 1 and Seyfert 2, depending on the width of emission lines, with Seyfert 1 showing broad and narrow emission lines and Seyfert 2 – only narrow ones. Seyfert 1 galaxies are also bright in X-ray band, while Seyfert 2 show faint X-ray emission.

- *Radio galaxies*: discovered at the dawn of the radio astronomy era (1950s). This name was attached to AGN, which were found to be very prominent in the radio band, and the host galaxy of which was spatially resolved. Radio galaxies show overall smaller energetics than that of radio-loud quasars. The map of

[1]The first observations of AGN in fact date back to 1909, done by E. Fath in California (Fath 1909), however he did not realize that the observed objects were galaxies. Another early observations belong to H. Curtis who discovered a jet in M87 (Curtis 1918). But a strong gravitational field as an explanation for the observed activity was suggested only after Seyfert galaxies were discovered.

these objects at radio frequencies is characterized by a notable extended linear structure, a *jet*. The outflows can extend up to several kiloparsecs in distance. The radio galaxies are divided into two sub-classes, based on their observational appearance in the radio domain. "Fanaroff-Riley Type I" (FR I) show radio emission mostly coming from the compact core region (see top panel of Fig. 2.2), while "Fanaroff-Riley Type II" (FR II) display distant from the core large-scale radio lobes with bright hot-spots, forming at the termination shock between the relativistic jet and the intergalactic medium (see bottom panel of Fig. 2.2).

- *Radio-loud quasars*: also detected in the early times of radio astronomy (1960s). Stellar-like objects were found in the optical range in the small-angular-size regions on the sky from where the strong radio emission emanated. In addition, the optical spectra of these sources defied interpretation, until M. Schmidt in 1963 understood that they were highly redshifted. The redshift Schmidt (1963) deduced for 3C 273 appeared to be $z = 0.158$. These measurements indicated very large distances to these objects, and allowed to estimate quasar luminosities, appearing in the range of 10^{45} – 10^{48} erg/s, several orders of magnitude higher than typical luminosities of normal galaxies. Later on, higher-quality optical observations of these objects revealed that quasars are (in most cases) associated with elliptical galaxies, with the central bright core by far outshining the host galaxy. Radio-loud quasars are in turn sub-divided into *Flat Spectrum Radio Quasars* (FSRQ) and *Steep Spectrum Radio Quasars* (SSRQ), with the former having spectral index $\alpha_r < 0.5$ of the radio-band spectrum $F_\nu \propto \nu^{-\alpha_r}$.

- *Radio-quiet quasars*: or QSO (quasi-stellar objects) are nearly as luminous as radio-loud quasars, however showing quite dim radio emission. Spectra of these objects display strong emission lines. Overall, AGN with an unresolved host galaxy are referred to as quasars.

- *BL Lac objects*: named after a prototypical object BL Lacertae. These sources first appeared as rapidly varying peculiar "stars" with very weak or no spectral lines and partially polarized emission. They also show strong radio emission. The spectrum of BL Lac objects spans from radio domain to γ-ray band.

This distinction is rather synthetic and emerged due to exclusively historical reasons, and thus does not necessarily reflect entirely different origin of the activity in members of different classes. With subsequent multi-band observations and extensive studies of various AGN, some clues on the connections between these different classes were found. For example, radio-quiet quasars and nuclei of Seyfert galaxies have very similar spectral characteristics, with the only major difference being their luminosity

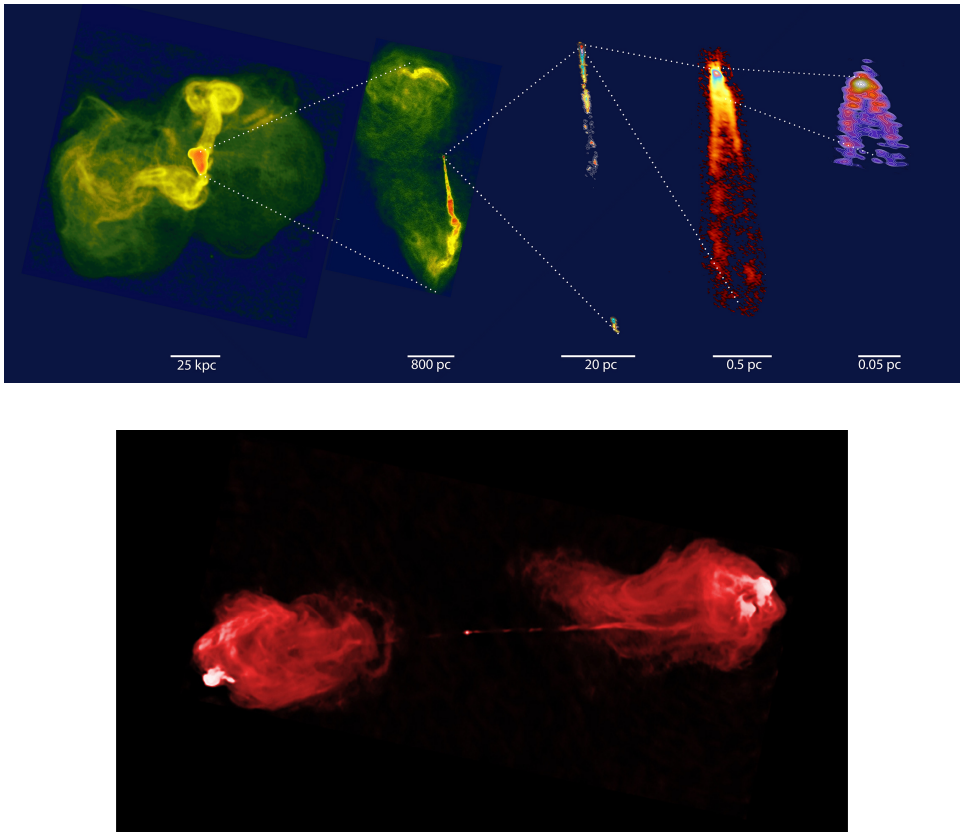


Figure 2.2: Illustration of the difference between morphology of the radio emission of FR I and FR II galaxies. *Top*: a radio image of an example FR I radio galaxy M87, on different spatial scales from the outer radio lobes to the jet launching region in the vicinity of the black hole. One can see that the core region dominates the observed radio emission. (Credit: NRAO, 90 and 20 cm VLA, 20 cm and 7 mm VLBA, and 3mm global VLBI ; image source: Blandford et al. (2019)). *Bottom*: radio image of an example FR II galaxy Cygnus A. One can see the jets emanating from the core region, which after a certain distance dissipate into giant radio lobes featuring conspicuous hot-spots at their extreme ends. (Credit: NSF/NRAO/AUI/VLA ; image source: chandra.harvard.edu)

(quasars are much brighter). FR I and FR II, as well as BL Lac and FSRQ, seemed to differ according to the same aspect, with FR II and FSRQ being brighter. The two sub-categories, BL Lacs and FSRQs, are collectively referred to as "*blazars*" (from merging of words "BL Lac" and "quasar", combined in a way to integrate into the term "blazing" characteristic of these objects). Observational appearance of the jet for radio-loud AGN, as well as radio galaxy / blazar division, is naturally explained if one observes the same object from a different viewing angle. Similar situation can be considered for the case of Seyfert 1 / Seyfert 2 division to interpret different width of emission lines and X-ray brightness between the sub-class representatives, assuming presence of obscuring material in the vicinity of the core region. Thus, all the great diversity of the observed differences between various AGN classes can be described as due to only several factors. Combination of the information on the links between different AGN classes into one picture led to a unified scheme of the AGN phenomenon.

2.2 Unified scheme of AGN phenomenon

The unification model of AGN, developed by Urry & Padovani (1995), describes in a self-consistent framework the variety of the observed AGN properties as different manifestations of the same type of object.

The exact structure of this underlying object was established based on multiple observations of AGN of different classes, and represents the key ingredient of the unification scenario. In this view, being now the most commonly accepted, an AGN is composed of a central supermassive black hole (SMBH), an **accretion disk**, a **dusty torus**, **clouds**, and a **jet** in the case of a radio-loud AGN (see Fig. 2.3). The group of fast-moving small clouds of gas close to the SMBH represent a so-called "**broad line region**" (BLR), and the collection of slowly-moving more distant clouds is named "**narrow line region**" (NLR). The material in the BLR and NLR emits a spectrum with broad and narrow lines respectively, due to different speeds of matter (Doppler broadening).

In the unified scheme by Urry & Padovani (1995) and several further developments, the observational appearance of an AGN having the structure described above, is defined typically by only three factors, namely the **orientation** effects, the **accretion rate** and the **black hole spin**. The latter two are in turn thought to be determined by environmental and evolutionary factors. The unification scenario of the AGN phenomenon is schematically illustrated in Fig. 2.3.

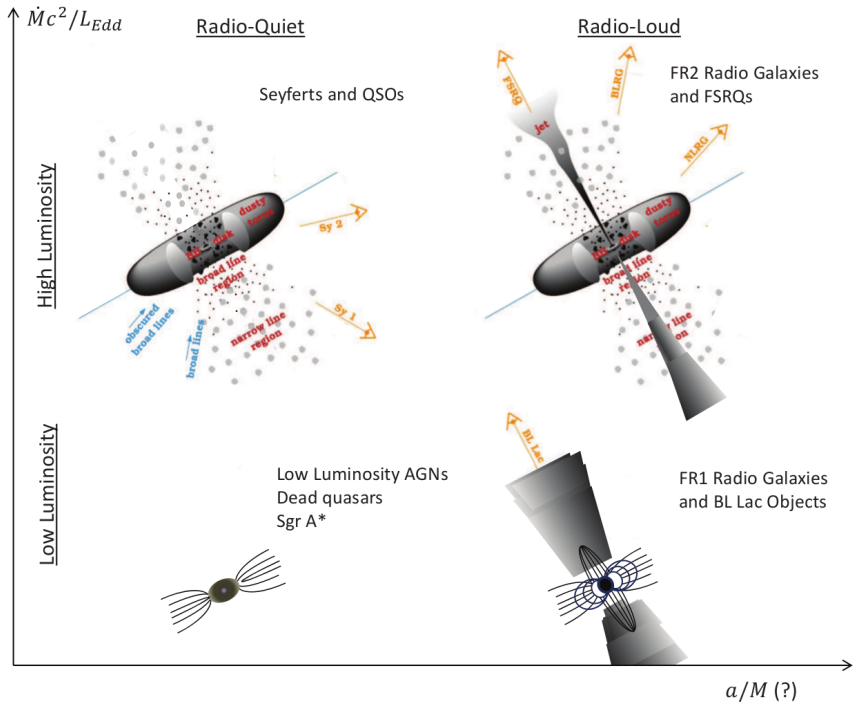


Figure 2.3: Illustration of a typical unification scenario for the AGN phenomenon. In this model, the activity of a galactic nucleus is controlled by only two parameters: mass-accretion rate regulating the luminosity of the source, and the central black hole spin responsible (supposedly) for a presence of a jet and therefore prominent radio emission. Finally, orientation of the observer's line of sight with respect to the symmetry axis defines the observational properties such as width of emission lines in spectra, and observational appearance of radio-loud AGN, as well as the intensity of their γ-ray emission. Blazars, comprising BL Lacs and FSRQs, are FR I and FR II galaxies respectively with the jet aligned with the line of sight. (adapted from Dermer & Giebels (2016))

2.2.1 What defines the observed luminosity?

Probably the first fundamental problem to address is the nature of AGN activity, and what physical parameter regulates the energy output. Since the discovery of Seyfert galaxies and quasars, various works endeavored to explain the origin of their abnormally high luminosities. After some time it was understood that such enormous energy release cannot be generated by thermonuclear fusion. It was concluded that the only reasonable option of the energy reservoir powering an AGN, that was in agreement with the observed luminosities, is gravitational energy, provided by very large masses of the central core region. The lower limit on the mass of the central object can be estimated assuming that the observed luminosity saturates the Eddington limit, the highest possible luminosity of an object before the radiation pressure starts blowing away the constituent material. The Eddington luminosity is

$$L_{\rm edd} = \frac{4\pi G M_{\rm co} m_{\rm p}\, c}{\sigma_{\rm T}} \simeq 1.3 \times 10^{38} \frac{M}{M_\odot}\,{\rm erg/s} \tag{2.1}$$

where $m_{\rm p}$ is the proton rest mass, G is the gravitational constant, $M_{\rm co}$ is mass of the central object, and $\sigma_{\rm T}$ is the Thomson cross-section.

For the most powerful quasars with luminosities of the order of 10^{48} erg/s, one obtains the minimal mass of central object about $10^{10}\,M_\odot$.

Next, rapid variability of AGN on time-scales of months and less implied that this giant mass is concentrated in a very compact region of the size smaller than 1 light month, which is much shorter than typical distances between stars in a galaxy. These constraints on the mass and size of the core are consistent with the central object being a black hole. Indeed, the radius of the event horizon (or the Schwarzschild radius) is

$$r_{\rm s} = \frac{2GM_{\rm co}}{c^2} \tag{2.2}$$

For the central object mass $M_{\rm co} \sim 10^{10}\,M_\odot$, we obtain $r_{\rm s} \sim 1$ light day, well in agreement with the causality argument. It was however not entirely clear which exact mechanism leads to a release of the gravitational energy stored in the central supermassive black hole. The only plausible process of conversion of the gravitational potential energy into heat and radiation is when a body falls into a deep gravitational well. A now commonly accepted scenario of AGN activity, proposed for the first time independently by Salpeter (1964) and Zel'dovich (1964), explains the observed high luminosities of AGN by accretion of matter on a massive object (a black hole).

If the material rotating around the BH in its vicinity is dense enough, the friction forces due to viscosity are important. The viscosity is most likely related to the turbulence in the magnetized plasma. The viscous friction evacuates the angular momentum outwards and dissipates the mechanical energy of the material, which causes it to proceed to a lower orbit, and so on, so that the matter spirals down towards the central BH. As a result, an *accretion disk* around the BH is formed. The matter in the differentially rotating disk heats up due to viscous friction between the annual layers and, as the plasma is optically thick, emits *thermal emission*. Unlike in the accretion on a neutron star, a BH does not possess any solid surface on which the matter can fall. The accreting plasma radiates only until reaching the innermost stable circular orbit around the BH, which for Schwarzschild (non-rotating) BH is $r_{\mathrm{isco}} = 3r_{\mathrm{s}}$. After crossing this boundary, as stable orbits are no more possible, the matter falls on the BH directly and does not radiate. Consequently, the radius r_{isco} defines the inner radius of the accretion disk.

Let us estimate the accretion rate $\dot{m}$ required to produce the luminosities observed in quasars. According to the virial theorem, during the accretion process, one half of gravitational potential energy is released in a form of kinetic energy (the matter speeds up when proceeding to a lower orbit) and one half in heat energy that is radiated away. Based on that, one obtains

$$L = \frac{1}{2}\frac{GM_{\mathrm{BH}}\,\dot{m}}{r_{\mathrm{isco}}} = \frac{1}{12}\,\dot{m}c^2 \tag{2.3}$$

This is a remarkable result implying that accretion on a non-rotating BH leads to a release of $\sim 8\%$ of the rest energy of the accreting matter (more precisely, $\sim 6\%$ due to the gravitational redshift correction). Even higher conversion efficiency is achieved for a Kerr (rotating) black hole, which for the fastest possible rotation (maximum possible angular momentum $J_{\mathrm{BH,max}} = \frac{GM_{\mathrm{BH}}^2}{c}$), is $\sim$42%. Such high percentage of energy release is explained by a strong gravitational field of the BH: the matter moves in the vicinity of the BH with speeds close to the speed of light, so that the kinetic energy of the material becomes comparable with its rest energy. The accretion rate necessary to yield the observed luminosity, is $\dot{m} \sim 100\,M_{\odot} \cdot \frac{L}{10^{48}\,\mathrm{erg/s}}$ year^{-1}.

The main result expressed by the Eq. 2.3 is that the AGN luminosity is proportional to the accretion rate (up to the Eddington limit). *Within the context of the unification scheme*, the difference in the luminosity of Seyfert galaxies and radio-quiet quasars, as well as of FR I and FR II radio galaxies, is explained by a difference in the accretion rate (see Fig. 2.3). The activity of high-luminosity AGN (quasars and FR II radio galaxies) is powered by accretion of matter on the SMBH in sub-Eddington regime. Because of that, for quasars, the central core vastly dominates the emission.

Distant quasars in the early observations appeared as stellar-like objects, since it was very difficult to resolve the faint host galaxy against the highly luminous core. Nowadays, for some quasars located at moderate distances, advanced instruments with high angular resolution and signal-to-noise ratio (e.g. Hubble Space Telescope) often allow to detect the parent galaxy in the wings of the point spread function. In contrast to quasars, in the lower-luminosity Seyfert nuclei the accretion is weaker, and because of that these objects show only a relatively bright core.

FR I radio galaxies have even lower-luminosity due to very weak accretion, possibly related to a different environment of the SMBH. In these objects, the accretion proceeds in a different regime. The viscous friction mostly heats heavy ions, and since the density of plasma around the BH is very low, the energy exchange between ions and electrons is not efficient enough to equalize their temperatures, so that the electrons remain much cooler. A fraction of the energy dissipated due to viscosity is advected inwards. This type of accretion is named "*advection-dominated accretion flow*" (ADAF). The plasma emits Bremsstrahlung radiation, rather than thermal emission, due to low optical depth. Due to high ion-electron temperature imbalance, the Bremsstrahlung emission time-scale appears to be much longer than the free-fall time-scale. Because of that, ADAFs are *radiatively inefficient*. Also, due to high ion temperature, ADAFs are geometrically thick, contrary to the radiatively-efficient accretion disks, which are geometrically thin.

FR I galaxies do not seem to have significant BLR and dusty torus, while the "medium"-luminosity Seyfert galaxies do posses rather important BLR and dusty torus. As a result, FR I galaxies do not have any important radiation fields internal to the source, with except for the synchrotron emission of the relativistic jet.

The difference between the FR I and FR II galaxies might be related not only to the accretion regime, but also to the particle content in the relativistic jet (e.g. electron-positron vs. electron-proton plasma). Another possibility may be a different environment around the two types of galaxies, i.e. different properties of intergalactic medium, affecting the propagation of the jet. Finally, FR I and FR II may simply be two different stages of an activity of the same object: in the scenario proposed by Böttcher & Dermer (2002), the rate of the accretion on the SMBH may gradually reduce with time, causing the transition FR II $\rightarrow$ FR I.

2.2.2 Radio-loud or radio-quiet?

Another crucial question to answer is what governs whether an AGN is radio-loud or radio-quiet. It is well established that the radio-loudness of an AGN is linked to

the presence of **jets**, which are highly collimated plasma outflows ejected from the central core and moving at a speed often close to the speed of light. Charged plasma particles move relativistically in a magnetic field in the jet, and emit synchrotron radiation in the radio band (the synchrotron mechanism is discussed in more detail in sub-section 4.1.2). The synchrotron origin of the emission also explains the high degree of polarization observed for BL Lac objects.

It is however not entirely clear what effect causes emergence of relativistic jets. The *launching of a jet* is presumably related to the *SMBH spin*[2] and the *rotating accretion disk*. The currently most accepted jet launching mechanism from the BH magnetosphere is that proposed by Blandford & Znajek (1977), which have shown that the rotational energy of a black hole may be extracted by electromagnetic stress and converted into the Poynting flux, which powers an outflow. The Blandford-Znajek process can be described within the framework of general-relativistic magnetohydrodynamics, i.e. hydrodynamics combined with Maxwell equations and with Einstein equations of general relativity (motions of plasma in a curved space-time around the black hole). Qualitatively, a fast-rotating black hole (BH) warps the space-time around it, with the distortion described by the Kerr metrics (Kerr 1963) featuring a peculiar property named rotational frame-dragging (Lense & Thirring 1918), which causes the precession of tilted orbits of test particles orbiting around the BH, and forces the particles close to the BH to participate in its rotation. In case the BH is spinning very fast, the orbits can precess at the speed close to the speed of light. The magnetic field lines embedded in the accretion disk follow the motion of the plasma (frozen-in condition), and close to the BH the frame-dragging effect drags magnetic field lines in the direction of the BH rotation and they start to wind the BH event horizon. As a consequence, the magnetic field lines threading the horizon coil up, resulting in the outgoing angular momentum flux along the direction of the rotational axis of the BH. As a result, the rotational energy of the BH is extracted electromagnetically, i.e. the BH rotational energy is converted into Poynting flux, which launches the plasma outflow (a jet). A schematic view of this process is depicted in the top panel of Fig. 2.4. The rate of extraction is proportional to the square of the magnetic field strength at the BH horizon, and grows with an increasing BH spin.

Another important process that may contribute to jet launching, is the Blandford & Payne (1982) process, in which the outflow is launched by the accretion disk (and not the central BH). Magnetic field lines anchored in the differentially rotating

[2]This view is very convincing on the basis of theoretical arguments. However it is challenged by certain observations: some radio-quiet Seyfert galaxies appear to host a fast rotating SMBH, while do not have a jet. Therefore, it is possible that the fast rotation of the BH is not sufficient to launch the jet, and so the question of the radio quiet/loud dichotomy remains open.

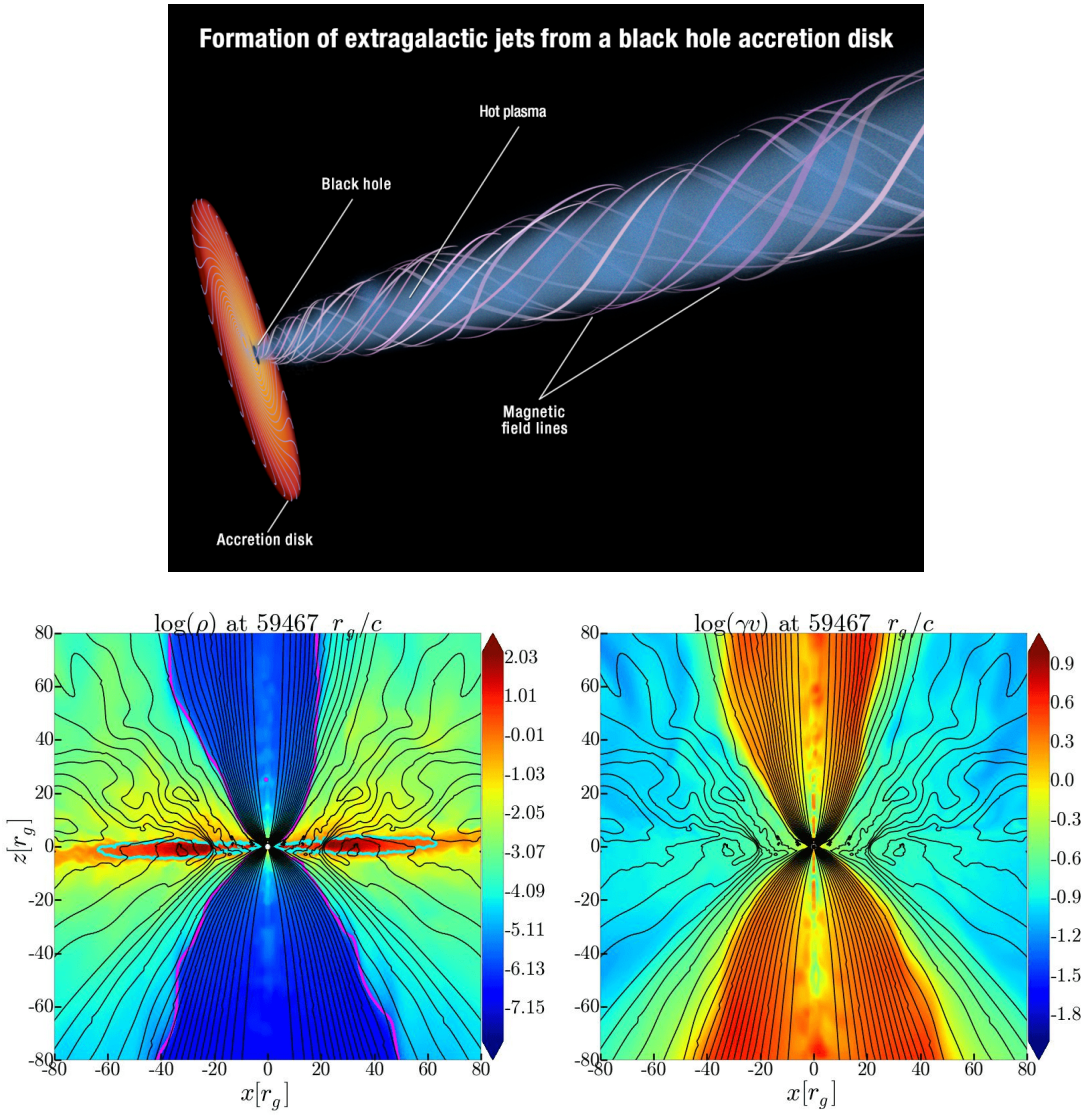


Figure 2.4: Illustration of the jet launching process. *Top*: schematic representation of the Blandford & Znajek (1977) and Blandford & Payne (1982) mechanisms. The jet is driven by magnetic field lines twisted by the black hole frame-dragging and/or the differentially rotating accretion disk. The jet is attached via magnetic field lines to the BH event horizon and to the accretion disk and pumps out their rotational energy (Credit: NASA, ESA, and A. Feild (STScI)). *Bottom*: example of general-relativistic magnetohydrodynamic simulations of jet launching. The left panel displays transverse slices of the logarithm of medium density, the right panel – same for the logarithm of the proper velocity of the medium γv. As one can see, the magnetic field lines (indicated in black) that are connected to the event horizon, are responsible for the Poynting flux dominated jet launching, while those connected to the accretion disk drive a matter-dominated outflow. (adapted from Liska (2019))

accretion disk and leaving its surface, drive angular momentum and energy flux in the direction perpendicular to the disk. This mechanism propels a matter-dominated jet, contrary to the Poynting flux dominated jet in the case of Blandford-Znajek process. Typically, Blandford & Payne (1982) jets are less relativistic than Blandford & Znajek (1977) ones.

General relativistic magnetohydrodynamic (GRMHD) simulations of the jet launching overall confirm the Blandford & Znajek (1977) and Blandford & Payne (1982) mechanisms viability (McKinney & Narayan (2007) ; Penna et al. (2013) ; Liska (2019)). An example of such simulations is shown in the bottom panel of Fig. 2.4. However a variety of open questions on the jet launching remain, e.g. it is not fully clear why jets are highly collimated (due to magnetic tension?) and confined, especially at large distances from the BH. Further observations of jets in various objects, as well as further GRMHD simulations are required to elucidate the unresolved problems.

Therefore, *within the unification scenario* that we adopted here, the radio-loudness is controlled by the value of the black hole spin (see Fig. 2.3). This is a rather common view at the moment, however it is important to stress that there is no consensus yet on what causes the observed bi-modality of AGN (radio-loud / radio-quiet).

It is possible that this dichotomy is related to environmental factors. Radio-quiet AGN, comprising cores of Seyfert galaxies and their higher luminosity counterparts radio-quiet quasars are predominantly observed in spiral star-forming galaxies. In contrast, radio-loud AGN reside mostly in elliptical galaxies having low rates of star formation, and which are very often found in galaxy clusters.

Finally, the radio loudness/quiescence is, in general, related to suitable / non-suitable physical conditions for efficient particle acceleration in the AGN. Radio-quiet AGN do not show any signatures of high-energy particles (in particular, synchrotron emission which is a tracer of relativistic particles), suggesting very low efficiency of particle acceleration, whereas the non-thermal emission of radio-loud AGN indicates highly efficient particle acceleration in these objects. Various observations of radio-loud AGN show that the most favorable conditions for particle acceleration are found in jets.

2.2.3 Orientation effects

The viewing angle, i.e. an angle between the line of sight an the axis of symmetry, has a great effect on the observational appearance of an AGN. As we will see in the next section, the emission produced in the relativistic jet has a highly anisotropic pattern

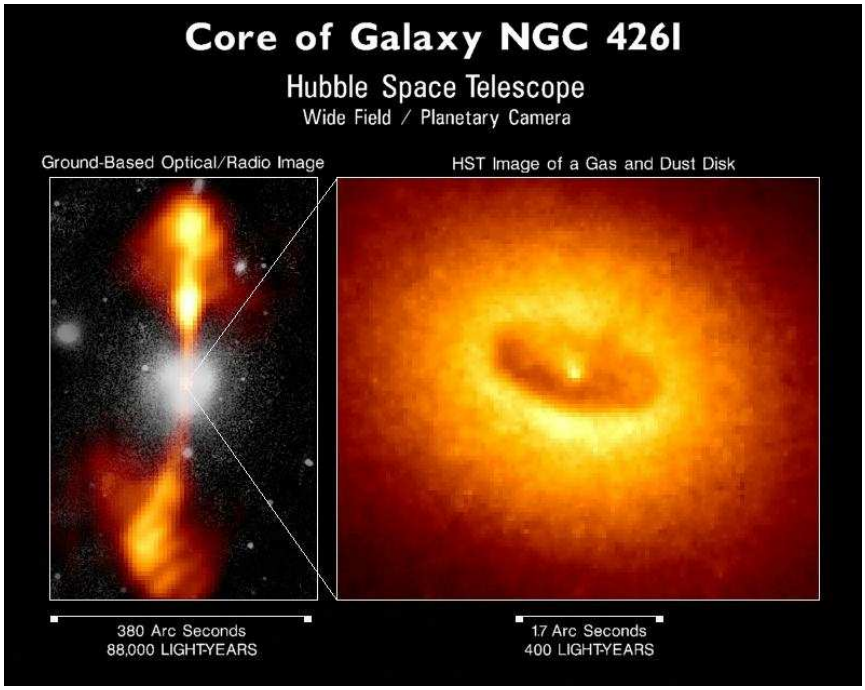


Figure 2.5: The active galactic nucleus NGC 4261. *Left*: superposition of images in the optical and radio band, showing the central core and a pair of relativistic jets emanating from it. *Right*: a zoom into the central region, showing obscuring dusty torus. (Credit: HST/NASA/ESA, adapted from Jaffe et al. (1993))

due to special relativity effects (beaming effect). Because of that, one would expect significant difference in observational properties of radio-loud AGN depending on the *orientation of the jet* with respect to the observer. In the "standard" unification scenario, if an FR I radio galaxy is observed in way that its jet is pointing towards the observer, it will appear as a BL Lac object, whereas an FR II galaxy in this situation will be seen as an FSRQ (see right part of the Fig. 2.3). An FR II observed at intermediate angles will appear as an SSRQ. Due to strong Doppler boosting of the jet emission, the host galaxy of highly distant blazars is often very difficult to resolve. The absence or weakness of lines in spectra of BL Lac objects is due to the absence or insignificance of BLR in FR I galaxies. As in case of blazar observation, we have a direct look inside the relativistic jet, we conclude that strong and rapid variability of these objects is therefore originating from the jet and arises due to certain violent processes taking place in it (discussed in more detail in sub-section 5.3.1).

Another reason for dependence of the appearance of AGN on the viewing angle is the presence of a *dust torus*. The presence of obscuring material was assumed in Seyfert galaxies to explain the Seyfert 1 / Seyfert 2 dichotomy. Galaxies of these two types share a lot of similar properties, with the only essential difference being width

of emission lines and X-ray brightness, which alluded to the idea that a Seyfert 1 and a Seyfert 2 nucleus is, in fact, the same underlying object observed as obscured or non-obscured by putative opaque material. In this view, the low-luminosity radio-quiet AGN appears as Seyfert 1 when one directly observes the BLR (and so the broad lines), as well as the central core, so that the X-ray emission emanating from it is visible. The X-ray emission comprises two components, with the soft X-ray one believed to be the high-energy tail of the accretion disk radiation, and the hard X-ray one being produced in the "hot corona" above the accretion disk via inverse Compton scattering (see sub-section 4.1.2 and Fig. 4.1) of soft disk photons. In contrast, the same object is seen as Seyfert 2 when the observer is located at lower latitudes[3] and the BLR and the central core are obscured by the dusty material, so that only the NRL is observed directly (and hence the narrow lines), and the X-ray emission is significatly attenuated by the dust due to photoelectric absorption. The attenuation is especially strong at soft X-ray energies, due to a higher interaction cross-section. This orientation effect for Seyfert galaxies is illustrated in the top-left part of Fig. 2.3.

This interpretation of the Seyfert 1 / Seyfert 2 divide within the unification scheme has several observational confirmations. Firstly, broad lines were observed in polarized light from Seyfert 2 galaxies, which indicates scattering of BLR emission towards the observer in this case (Antonucci & Miller 1985). Secondly, the dusty torus was observed directly first in high-resolution infrared, and then in optical observations. An example image (however not for a Seyfert galaxy) is shown in Fig. 2.5. The obscuring material has a form of a torus and is composed of optically thick clouds of molecular dusty material. The presence of such torus in high- and medium-luminosity AGN (quasars and Seyfert galaxies), and its insignificance/absence in low-luminosity AGN (FR I) is naturally explained by (*i*) higher density of dust in the vicinity of the BH in more luminous AGN (due to higher accretion rate), and (*ii*) higher radiation pressure in more powerful AGN, which sweeps dust away.

2.3 AGN jets

2.3.1 Imaging of jets

As already presented before, some AGN launch jets, presumably due to fast rotation of the central black hole, or the accretion disk. The first observation of jets was done by Curtis (1918), who noticed a straight "ray" emanating from the M87 nucleus.

[3]Latitude, in this case, is the angle between the plane of the accretion disk and the line of sight.

Present-day high-resolution imaging techniques, together with the proximity of the object (distance about 20 Mpc), enable to obtain high-quality images of the M87 jet, with the examples shown in the top panel of Fig. 2.2 and the top panel of Fig. 2.6.

Deeper zooms obtained in the radio band with the interferometry technique allow to reveal the jet launching region very close to the BH. The top panel of Fig. 2.2 represents a collage of M87 images at radio frequencies at different spatial scales, with the second (from the left) image showing the galaxy jet with the inner lobes, and further images showing step-by-step zoom towards the jet launching region near the central BH (rightmost image), which is resolved down to an impressive spatial scale of $7\,r_{\rm s}$ (Kim et al. 2018) thanks to global mm VLBI network. Finally, in 2019, the Event Horizon Telescope (EHT) collaboration for the first time detected and performed imaging of the shadow of the M87 black hole. The shadow image is shown in the bottom panel of Fig. 2.6.

2.3.2 Kinematics and structure

Various observations indicate, that the jet is not uniform, and is rather composed of two components: a *faster inner spine* and a *slower outer sheath*. Early observations (in radio band) of such configuration were reported by e.g. Giovannini et al. (1999), and confirmed later, based on imaging with more advanced instruments. For example, M87 43 GHz VLBI observations by Mertens et al. (2016) showed a faster spine moving with bulk Lorentz factor $\Gamma = 2.5$, surrounded by mildly relativistic outer sheath layer. In addition, VLBI mm observations of the base of the jet by Kim et al. (2018) (rightmost image in the top panel of Fig. 2.2) revealed limb brightening morphology, implying that the jet features spine-sheath configuration even at this spatial scale.

Besides direct observations, the spine-sheath structure for the jet has been also suggested on the basis of theoretical considerations (e.g. Sol et al. (1989) ; Henri & Pelletier (1991)). In particular, Sol et al. (1989) propose a two-flow model of a relativistic jet, in which the faster inner spine is produced by Blandford & Znajek (1977) mechanism, while the outer mildly relativistic envelope by Blandford & Payne (1982) process. It was shown by the authors that such jet configuration is stable with respect to various plasma instabilities that may arise at the interface between the two components.

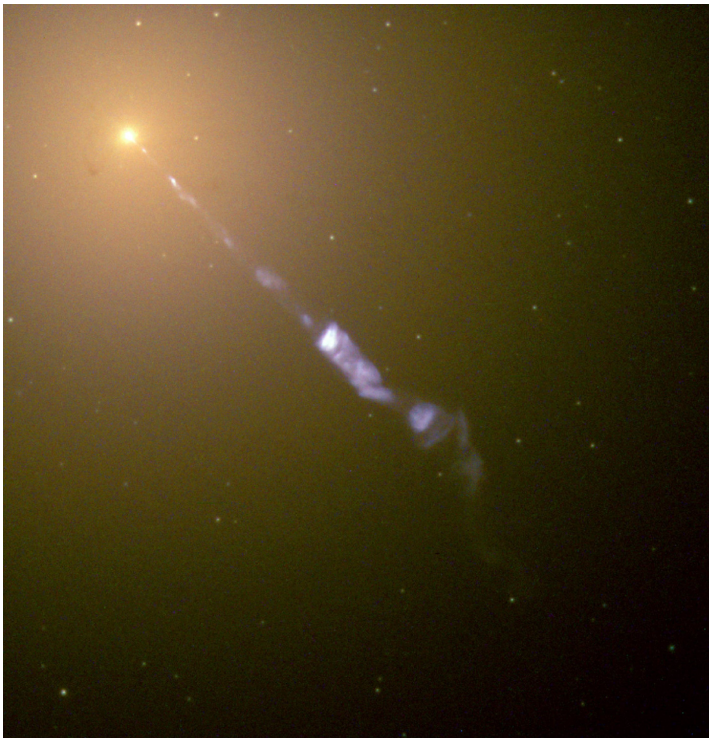


Figure 2.6: *Top*: high-resolution image of the M87 jet obtained by Hubble Space Telescope (Credit: NASA). *Bottom*: image of the shadow of the M87 black hole (Credit: EHT collaboration ; image source: eventhorizontelescope.org).

2.3.3 High-energy particles in jets

A substantial progress in our understanding of matter content and particle processes in jets was made by observations at higher energies, from radio to X-ray domain. These observations revealed brighter spots in jets, called "*knots*", especially prominent in X-ray band (an example is shown in Fig. 2.7 for the jet of M87). The knots show non-thermal emission, with the spectral energy distribution (SED) having typically a broken power law shape. For the case of M87, Marshall et al. (2002) find that the SED of the knot emission is well represented by a broken power law model, with the spectral index in radio-to-optical band of $\alpha_{\rm kn,r\text{-}o} \simeq 0.7$, and in X-ray band of $\alpha_{\rm kn,x} \simeq 1.5$, and with the break frequency $\nu_{\rm kn,br} \sim 10^{16}$ Hz. Authors also conclude that the SED is well described with a *synchrotron emission* model, and estimate the Lorentz factors of particles responsible for the observed emission to be $\gamma_{\rm kn} \sim 10^7$. The most plausible option for the particles that produce the radiation is electrons, as protons are inefficient at emitting synchrotron emission due to their higher mass. Therefore, knots are filled with either electron-proton, or electron-positron plasma, with maximum energy of electrons around $E_{\rm e,kn,max} \sim 10$ TeV.

This result indicates that jets contain *high-energy particles*, and hence that *particle acceleration processes* operate inside jets. Let us estimate the synchrotron cooling time-scale (see sub-section 5.1.2 and the Eq. 5.8) for the 10 TeV electrons in the jet. Assuming that the magnetic field in the knots is slightly stronger than in the interstellar medium of a galaxy, $B_{\rm kn} \sim 10^{-5}$ G, one finds $t_{\rm e,cool,kn} \sim 10^{11}$ s. During this time, highest-energy electrons would propagate for distances of ~ 1 kpc along the jet while losing their energy. However, observations show that knots are situated at much further distances of about a few tens of kiloparsecs from the central engine. This contradiction represents one of the long-standing problem of jet physics: why jets extend to distances as long as a few tens, hundreds, or in some cases even thousands of kiloparsecs. The standard explanation is that particles experience continuous acceleration along the path of the jet (e.g. by shocks or turbulence) and are injected locally. An alternative scenario, proposed by Neronov et al. (2002) assumes injection at the base of the jet of a powerful beam of γ-rays with energies $\gtrsim 10^{15}$ eV, producing electron-positron pairs when encountering cosmic microwave background (CMB) photons while propagating through the jet. Such mechanism allows to generate and supply high-energy particles along the entire length of the jet up to distances of hundreds of kiloparsecs.

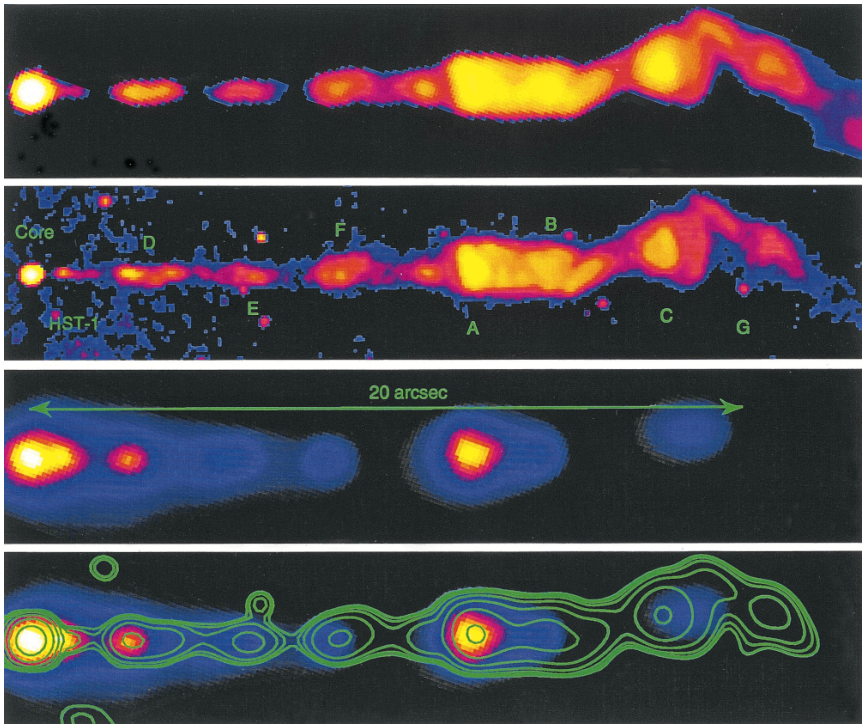


Figure 2.7: Multi-band view of the jet of M87. *Top panel*: VLA image at 14.435 GHz. *Second panel*: optical image (in the red part of the optical spectrum) obtained by Hubble Space Telescope. *Third panel*: Chandra X-ray Observatory image. *Fourth panel*: same as third panel, but with superimposed contours of smoothed optical image. The brightness level is displayed with a logarithmic scale for radio and optical images, and with linear scaling for the X-ray image. (adapted from Marshall et al. (2002))

2.3.4 Energy dissipation

Jets of FR I and FR II galaxies terminate in large-scale lobes (see Fig. 2.2), where particles lose much of their energy through radiation and acceleration of particles of the ambient medium, contributing to the AGN feedback on their environment. The energy carried in the jet flow, in fact, comprises two components: *kinetic energy* (contained in matter) and *electromagnetic energy* (Poynting flux). The picture of dissipation of the total energy budget by jets of FR I and FR II galaxies is different and not yet fully understood. For jets of FR I galaxies, a significant fraction of their kinetic energy is dissipated at the base of the jet in the vicinity of the BH, however the electromagnetic energy is transported much farther, powering the outflow for large distances. To the contrary, jets of FR II galaxies, lose most of their electromagnetic energy close to the BH, but transferring kinetic energy flux to far distances all the way

to the lobes (e.g. Blandford et al. (2019)). This is suggested by the difference in the morphology of radio maps of FR I and FR II galaxies (demonstrated in Fig. 2.2), with the FR I objects having bright core region, and FR II showing prominent emission from the lobes.

The more powerful FR II jets also produce hot-spots, clearly visible in the lobes (an example is shown in the bottom panel of Fig. 2.2). The hot-spots correspond to termination shocks due to the interaction with the ambient medium. Particles are accelerated at the front of these shocks and emit synchrotron radiation, therefore making hot-spots appear very bright.

2.3.5 Relativistic motions in jets

Apart from high-energy particles, observations show that the matter in jets moves, in many cases, with relativistic speeds. The strongest evidence is provided by observation of knots motion with *apparent superluminal speed*. An example of such observation is presented in the top panel of Fig. 2.8. Knowing the distance to the source, more precisely, angular size distance calculated from the redshift, and angular displacement of the knots over a given time interval, one obtains the linear displacement and hence the velocity. The velocities of knots in jets measured in this way were found to be superluminal: for example, for PKS 1510-089 $v_{\mathrm{kn},1510} \simeq 22c$, for M87 $v_{\mathrm{kn,M87}} \simeq 5c$.

Such superluminal motion stems from relativistic velocities of matter in the jet and geometrical effects. Fig. 2.8 illustrates a geometrical scheme explaining the emergence of the apparent superluminal motion phenomenon. A jet is observed at a small angle θ with respect to the line of sight. A knot moves with a speed v close to the speed of light along the jet, and an electron inside the knot emits a photon at the moment of time t_0, and another photon after some time, at the moment t_1. The physical displacement of the knot over the time interval $\Delta t = t_1 - t_0$ as seen by the observer, is a projection of the traveled distance on the direction perpendicular to the line of sight, i.e. $\Delta x = v\Delta t \sin\theta$. The time interval between the detection by an observer of a photon emitted at point "0" and at point "1", is the path difference between these two photons along the line of sight divided by the speed of light: $\Delta t_{\mathrm{det},1-0} = (c\Delta t - v\Delta t\cos\theta)/c = \Delta t\,(1 - \beta\cos\theta)$, where $\beta = v/c$. The apparent velocity of the knot is then

$$v_{\mathrm{app}} = \Delta x/\Delta t_{\mathrm{det},1-0} = \frac{\beta\sin\theta}{1 - \beta\cos\theta}\, c \tag{2.4}$$

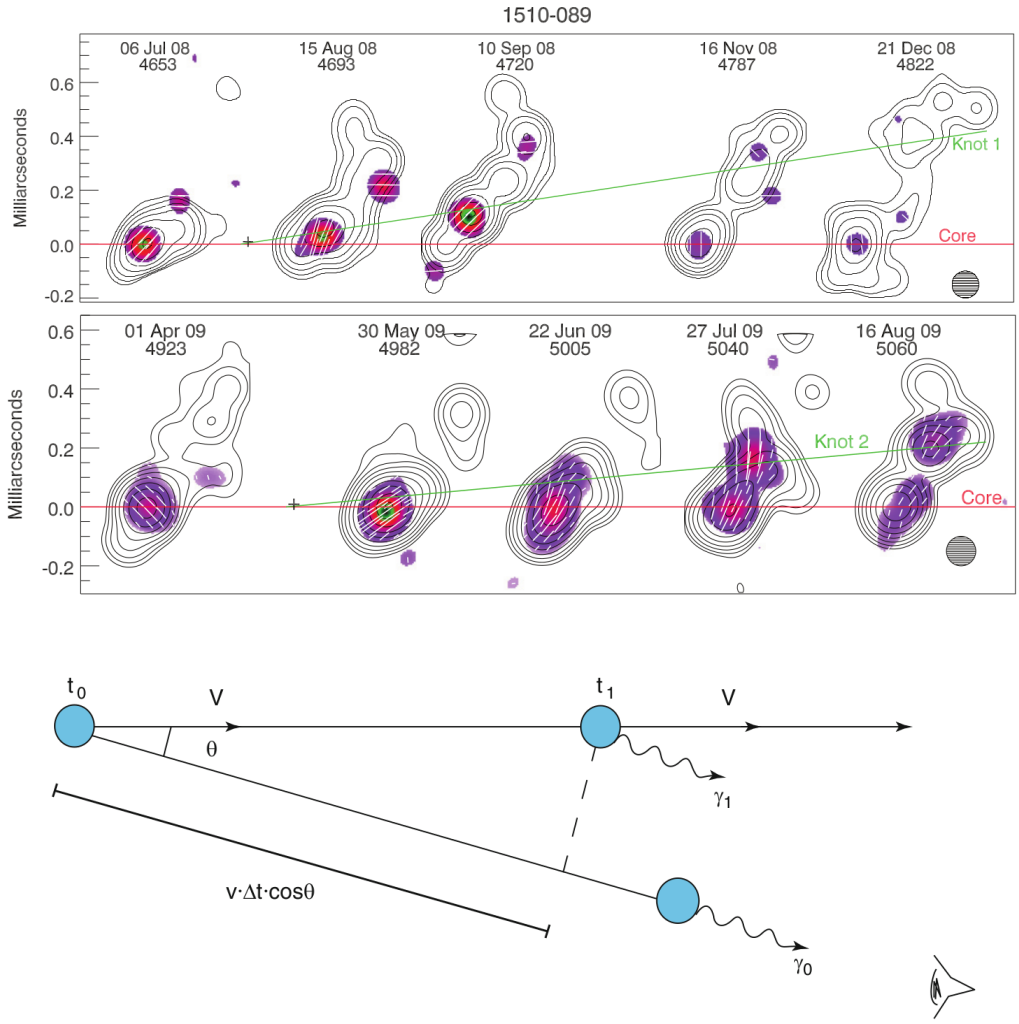


Figure 2.8: Illustration of apparent superluminal motions in relativistic jets. *Top*: Observations of superluminal motion of two knots (top and bottom panel) in PKS 1510-089 (z=0.361) performed in radio band by VLBA at 43 GHz. Contours display the intensity level of the total flux, and the color – of the polarized flux. White linear segments indicate the direction of the linear polarization. The first knot has an apparent velocity of $24 \pm 2\,c$, and the second one of $21.6 \pm 0.6\,c$. The scale of the y-axis is 0.5 pc / 0.1 mas. (adapted from Marscher et al. (2010)). *Bottom*: a scheme explaining the origin of apparent superluminal motions. (adapted from Courvoisier (2013))

If a relativistic ($\beta \to 1$) jet is oriented at a sufficiently small angle with respect to the line of sight, one measures an apparent superluminal velocity. The maximal velocity is achieved when $\beta = \cos\theta$, and has a value of

$$v_{\rm app,max} = \Gamma v \tag{2.5}$$

where $\Gamma = (1-\beta^2)^{-1/2}$ is the bulk Lorentz factor of the knot. From this equation for the case of PKS 1510-089 one infers $\Gamma_{1510} \simeq 22$, and for M87 $\Gamma_{\rm M87} \simeq 5$, with the angles between the jet axis and the line of sight $\theta_{1510} \simeq 3^\circ$, and for M87 $\theta_{\rm M87} \simeq 11^\circ$. Typical apparent velocities observed in jets of various other AGN are in the $v_{\rm app} \sim 10c$ range. Therefore, observation of superluminal apparent velocities is a direct indication of *relativistic motions of matter in jets* with typical bulk Lorentz factors $\Gamma \sim 10$.

2.4 Blazars

As already mentioned, blazars are FR I and FR II galaxies with the jet very closely aligned with the line of sight. Accidental orientation of the jet (nearly) towards the observer, as well as relativistic motions of plasma in it, produce not only apparent superluminal speed phenomenon, but also a strong emission boosting, discussed just below. We also focus in this section on different properties of blazars, including their broad-band SED and emission at high energies.

2.4.1 Doppler boosting

Due to bulk relativistic motion of plasma in the jet, a phenomenon named "Doppler boosting" emerges as a consequence of special relativity effects, specifically, relativistic time dilation.

For an emitter moving relativistically with the Lorentz factor Γ, at a small angle with respect to the line of sight, the relativistic transformations of physical quantities involve a characteristic *Doppler factor* δ, which is

$$\delta = \frac{1}{\Gamma\,(1-\beta\cos\theta)} \tag{2.6}$$

so that the time interval is transformed as $\Delta t_{\rm obs} = \Delta t_{\rm em}\,\delta^{-1}$, and the frequency as $\nu_{\rm obs} = \delta \cdot \nu_{\rm em}$, with the sub-script "*em*" indicating quantities in the emitter's frame, and "*obs*" in the observer's frame.

Emission intensity transformation involves four contributions: (*i*) relativistic time dilation for the time interval between arrival of two photons, (*ii*) relativistic decrease of emission solid angle (aberration), (*iii*) relativistic Doppler effect (frequency shift) for each individual photon, and (*iv*) relativistic transformation of the frequency interval. This results in the following transformation:

$$I_{\rm obs}(\nu_{\rm obs}) = \delta^3 \, I_{\rm em}(\nu_{\rm em}) \tag{2.7}$$

Same result can be obtained from the invariance of the photon distribution function in the phase space, expressed as

$$\frac{I_{\rm obs}(\nu_{\rm obs})}{\nu_{\rm obs}^3} = \frac{I_{\rm em}(\nu_{\rm em})}{\nu_{\rm em}^3} \tag{2.8}$$

The total energy flux integrated over the entire electromagnetic spectrum, $F = \int_0^\infty I(\nu)\, d\nu$, will be further boosted by a factor of δ:

$$F_{\rm obs} = \delta^4 \, F_{\rm em} \tag{2.9}$$

Therefore, the Doppler boosting effect dramatically enhances the flux from the approaching jet, while heavily suppresses the one from the receding jet. This explains why in most cases, the opposite jet is not visible. Also, the non-thermal jet emission exceedingly dominates the observed blazar emission due to the Doppler boosting effect.

Another relevant effect of special relativity is the so-called "beaming" (headlight) effect. The radiation emitted by the relativistically approaching source isotropically in its own frame, is observed as concentrated in a cone with an opening angle of $\theta_{\rm cone} \sim \Gamma^{-1}$.

The beaming effect in particular explains why blazars emit the strongest γ-ray signal, while radio galaxies are not very bright γ-ray-emitters. In the above-mentioned spine-sheath jet structure, typically the γ-ray emission originates from the inner spine, while the most of the radio emission is generated in the outer sheath (e.g. Blandford et al. (2019)). As the spine is faster than the sheath, the observed γ-ray emission will be much more beamed and will be concentrated in a much narrower cone, than that of the radio flux. As a result, only radio galaxies oriented very closely to the line of sight are brightest γ-ray-emitters, in this case they are named blazars. Consequently, the SED of blazars spans from radio domain up to γ-ray band.

Also, bulk motions of emitting matter in blazar jets and the associated relativistic time dilation are responsible for shortening the observed variability time-scale: $\Delta t_{\rm var,obs} = \Delta t_{\rm var,source}/\delta$. In some blazars, it can be as short as a few minutes (for more details, see sub-section 5.3.1).

2.4.2 Blazar emission

As already mentioned, BL Lac objects correspond to aligned FR I galaxies, whereas FSRQs – to aligned FR II galaxies. The main difference between BL Lac and FSRQ is in their luminosity, with FSRQ being much brighter (typically by a factor of ~ 100).

Short variability of blazars (especially in the γ-ray band) indicates that the region in which the γ-ray emission is produced, should be very compact. According to the causality arguments, medium variability time-scale of $t_{\rm var,obs} \sim 1$ day, implies a small size of the emitting zone, $R_{\rm ez} \leq c t_{\rm var,obs}\,\delta/(1+z) \sim 10^{16}$ cm, with typical value of $\delta \sim 20$. Even shorter variability time-scales of a few minutes, correspond to a size of the emitting region commensurable with the radius of the horizon of a BH with mass $10^8\,M_\odot$. It is believed that the blazar γ-ray emission is produced in a compact plasma region, a "*blob*" filled with relativistic particles, and having higher density and magnetic field stronger than on average in the large-scale jet. Such a blob is moving at a speed close to the speed of light along the jet axis and presumably represents an ejecta produced in the inner spine of the jet due to magnetohydrodynamical instabilities and other effects. The exact location of the γ-ray-emitting site in the jet is not immediately obvious. For the case of an FSRQ, as we will see just below, the observed γ-ray spectrum can depend on the blob position with respect to the central engine. One of the ways to determine the distance of the γ-ray emission region from the SMBH is to study the correlation between radio and γ-ray emission. Such studies find that γ-rays are generated in a zone inside the jet, outlying typically up to a few parsecs away from the SMBH (e.g. Marscher et al. (2008)).

The very detection of γ-ray emission from blazars, and in particular, TeV γ-ray emission from BL Lac objects, imposes a lower limit on the value of the Doppler factor of the blob, $\delta_b \gtrsim 10$, otherwise the jet becomes opaque to the γ-ray photons due to γ-γ pair-production effect (the effect is explained in sub-section 4.1.5). The "blob-in-jet" model for blazar broad-band emission and relevant emission processes are described more in detail in Section 4.1.

A typical SED of blazar broad-band emission represents a *two-bump structure*, spanning over 15 or more decades in energy (see an example in Fig. 2.9). The low-energy bump extends from radio band up to UV or soft/hard X-rays, and peaks in

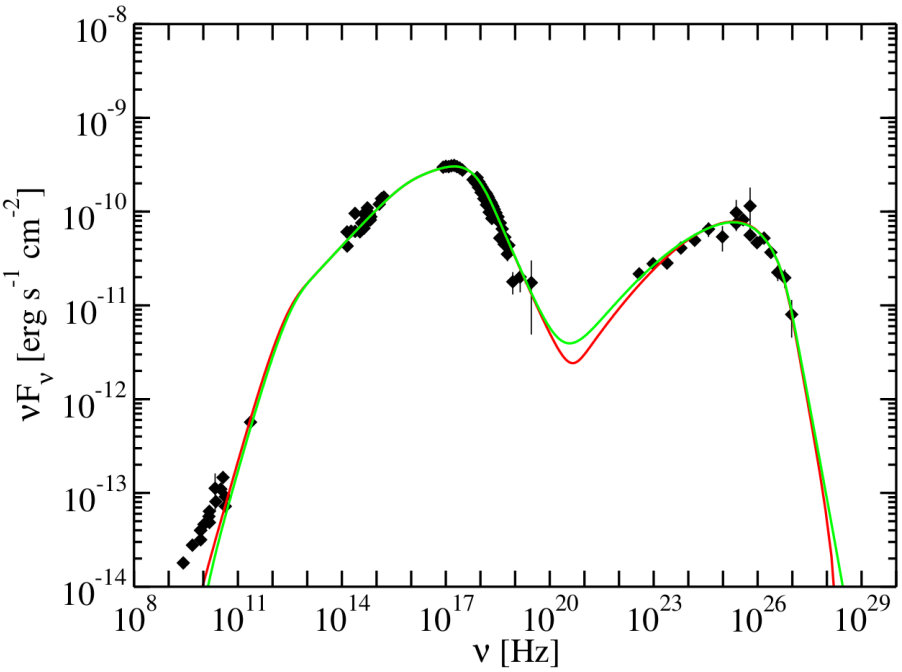


Figure 2.9: An example of blazar SED. Black points indicate MWL data set (host galaxy subtracted) of HBL BL Lac object Mrk 421 (analyzed in this thesis). Green and red curves represent the SSC models assuming a different variability time-scale. (adapted from Abdo et al. (2011))

the IR, optical, UV or X-ray band, and the higher-energy bump stretches up to GeV or TeV γ-rays, showing a maximum around MeV to sub-TeV, or even at TeV energies. The very detection of high-energy γ-rays from blazars provides an additional evidence of the presence of highly energetic particles inside the source. Analogically to the emission of the knots observed in jets of misaligned objects, the origin of the low-energy SED component of blazars is attributed to synchrotron emission generated by high-energy electrons moving in a magnetic field. A strong evidence in favor of this emission mechanism is the observation of emission polarization in blazars. While there is no doubt about the origin of the low-energy bump, the high-energy component still remains a subject of debate. The most common view is that it is produced by inverse Compton scattering mechanism of soft photons (see sub-section 4.1.4), however, alternative interpretations of the high-energy bump invoking hadronic processes exist (more discussed in sub-section 4.1.1). The dominant photon field for the inverse Compton process depends on the object type. For the case of a BL Lac with no important soft photon radiation fields except for the synchrotron emission, relativistic electrons upscatter the same synchrotron photons that they themselves produce, boosting them in energy up to γ-ray energies (so-called Synchrotron Self-Compton

scenario, SSC). For the case of an FSRQ, due to strong accretion, the dominant target photon field may be the thermal emission of the accretion disk (e.g. Dermer et al. (1992)), optical-to-UV radiation of the broad line region (BLR) (e.g. Sikora et al. (1994)), or IR radiation of the torus (e.g. Błażejowski et al. (2000)). This scenario is called "external Compton" since the target photons are external to the γ-ray emitting zone. The target radiation field for the inverse Compton upscattering process depends on the position of the γ-ray emitting zone inside the jet, specifically its distance from the central engine. For the emission region located at distances closer than $\sim$0.01 pc from the central black hole, the accretion disk photons dominate the external photon field, for distances in the domain $\sim$ 0.01 – 0.1 pc, the BLR radiation serves the main target for upscattering, and for distances $\sim$ 0.1 – a few pc, upscattering of the emission of dust torus prevails.

In the radio band, the blazar SED is typically dominated by the synchrotron *emission of the large-scale jet*, rather than that produced in the blob. Also, in the optical domain, the *host galaxy contribution* to the observed flux might be quite important.

2.4.3 Blazar sequence

BL Lac objects are classified into several types, depending on the peak frequency of the synchrotron bump. The currently standard classification distinguishes the following types: Low-frequency peaked BL Lac objects (LBLs), show synchrotron peak in the IR band, in the frequency range 10^{13} – 10^{14} Hz, Intermediate BL Lac objects (IBLs) display synchrotron peak at optical/UV wavelengths in the range 10^{14} – 10^{15} Hz, and High frequency-peaked BL Lac objects (HBL), exhibit their synchrotron peak at in the X-ray regime, at frequencies above 10^{15} Hz. Conversely, FSRQs have the lowest position of the synchrotron peak, in the 10^{12} – 10^{13} Hz range.

Statistical study of a large selection of blazars (BL Lacs of different types, and FSRQs; 126 objects in total) by Fossati et al. (1998) shows that the frequency of the synchrotron peak anti-correlates with the source power, more precisely, with the observed radio flux: more powerful objects have their synchrotron peak at lower frequencies. This trend is named "*blazar sequence*", illustrated in the right panel of Fig. 2.10.

The study by Fossati et al. (1998) was done in the first decade of γ-ray astronomy, as well as multi-wavelength astronomy, when the γ-ray instruments had quite modest sensitivity and the MWL coverage was rather incomplete. Therefore, the original blazar sequence was mostly based on the radio data. However, new instruments

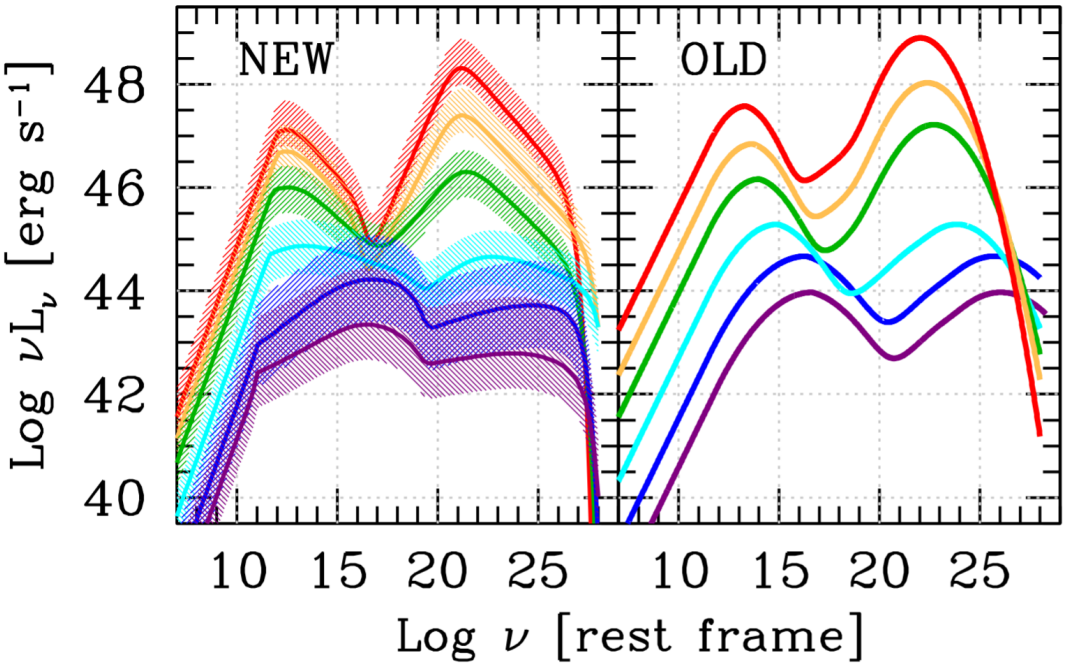


Figure 2.10: Updated (*left*) and original (*right*) blazar sequence. The initial classification was performed based on the radio flux, while the revised version uses γ-ray flux. The curves display SEDs of blazars, with color sequence from red to violet corresponding to the sequence FSRQ – LBL – IBL – HBL. (adapted from Ghisellini et al. (2017))

appeared in the past two decades, in particular, *Fermi* Gamma-ray Space Telescope (Atwood et al. 2009), launched in 2008, as well as ground-based instruments sensitive in the TeV γ-ray regime (for more details, see Chapter 3). X-ray instrumentation has also substantially evolved. Recently, an *updated version* of the blazar sequence was published by Ghisellini et al. (2017). Thanks to the advance in detection capabilities across the entire electromagnetic spectrum, it became possible, firstly, to increase the sample of blazars to 747 objects, and secondly, to access spectral properties in energy bands other than radio, in particular, γ-ray domain. The new study is based on the γ-ray flux measured by *Fermi* Telescope, instead of the radio flux as in the previous study. This revisited blazar sequence is shown in the left panel of Fig. 2.10.

One can see that the updated blazar sequence, rather similarly to the old one, exhibits the same trend of shift of the synchrotron peak to lower frequencies with an increasing radio flux, as well as the overall luminosity. The same situation concerns the high-energy SED bump: its peak migrates to lower energies with higher γ-ray flux and overall luminosity. In addition, one can clearly see that the high-energy component by far dominates the emission power for the most luminous blazars (FSRQs),

while the least luminous objects (HBL BL Lacs) show less power in high-energy bump than in the synchrotron one. Dominance of the high-energy component in the case of FSRQs can be readily explained (within the inverse Compton mechanism) by the presence of important external photon fields (from the accretion disk / BLR / torus) in addition to the synchrotron radiation field, so that the target photon density is higher and the high-energy component rises up.

The appearance of the blazar sequence is further explained using qualitative arguments in sub-section 4.1.1. The alternative view is that the blazar sequence might originate from a selection bias.

2.4.4 Probing the Universe with blazars

Using distant blazars as beacons and studying the effect of propagation of γ-rays over long distances, one could probe a variety of phenomena of cosmology and new fundamental physics, including those beyond the Standard Model of particle physics. This topic is covered in more detail in sub-section 7.1.2.

Chapter 3

Cherenkov gamma-ray astronomy and H.E.S.S. data analysis

Gamma-ray astronomy studies astrophysical sources emitting photons in the γ-ray part of the electromagnetic spectrum, i.e. with energies above $\sim$ 100 keV. The technique used to detect γ-rays is defined by the dominant process of their interaction with matter. For γ-rays with energies below $\sim$ 10 MeV, the interaction cross-section is dominated by Compton scattering. This is the underlying process of the detection principle used by Compton telescopes, observing in this energy range. The only such telescope which was operational up to now is the COMPTEL instrument of the *Compton* Gamma-Ray Observatory. The instrument was operational in the 1990s, and detected only around 10 sources because of its low sensitivity caused by peculiarities of the detection technique. Poor sensitivity of telescopes in this energy domain, known as "MeV sensitivity gap", precludes astronomical observations in the MeV regime. For γ-rays with energies above $\sim$ 10 MeV interacting with matter, electron-positron pair creation (Bethe-Heitler pair production, see the next section) prevails over the Compton scattering. The pair production effect is the basic ingredient in the detection technique used by e.g. Large Area Telescope (LAT) instrument of the presently operating *Fermi* Gamma-Ray Space Telescope (Atwood et al. 2009), a space-based instrument launched in 2008. Its detector (the "tracker") is made of a material with high atomic number Z favoring higher cross-section of the pair creation. A γ-ray entering the detector experiences conversion into an electron-positron pair. As the energy of the γ-ray is much higher than the electron rest energy, the resulting electron and positron are highly relativistic and they move inside the tracker in a direction at a small angle with respect to the direction of the initial γ-ray. The multilayer structure of the tracker allows to determine the positions of the electron and positron at several reference points and therefore to track the paths of the charged

particles as they traverse the detector. Information about their trajectories is used to reconstruct the arrival direction of the incident γ-ray. Another component of the telescope, the "calorimeter", located at the bottom of the tracker, measures the energies of the electron and positron; the sum of their energies is equal to the energy of the initial γ-ray (minus small energy lost by ionization). The *Fermi*-LAT instrument operates in the energy range from ~ 50 MeV up to a few hundreds of GeV.

Detection of γ-rays with higher energies with space-based pair-conversion telescopes is impeded by two limitations. The first one is related to "leakage" of particle showers from the bottom of the calorimeter, so that the energy reconstruction becomes very inaccurate. The second, and much more important obstacle, is insufficiently large collection area of spaced-based pair-converting instruments for observations in this energy regime. γ-ray spectra of various astrophysical objects usually show a fall of photon flux from a source with increasing γ-ray energy, in most cases, in a power law manner. Typical TeV fluxes from the brightest γ-ray sources on the sky, are such that a telescope with an effective area of ~ 1 m $\times$ 1 m $= 10^4$ cm^2 will detect only about one TeV photon per month. Any detailed spectral analysis is barely possible with such low signal statistics; setups with much larger collection areas are needed for this task. Clearly, it is very difficult to imagine to place an instrument with a collecting area several orders of magnitude larger than 10^4 cm^2 up in space, which means that a different detection technique should be used. Observations in the VHE γ-ray band are generally[1] performed with ground-based Imaging Atmospheric Cherenkov Telescopes (IACTs), which use the Earth's atmosphere as a detection medium and offer collection areas of the order of 10^8 cm^2. These instruments detect γ-rays indirectly, by recording Cherenkov light emitted by particles of the extensive air shower initiated when a highly energetic γ-ray hits the Earth's atmosphere.

In this chapter, we discuss in detail the IACT observational technique and present some of the currently operating instruments, including the High Energy Stereoscopic System (H.E.S.S.). As a next step, we focus on the analysis of H.E.S.S. data on two bright flares of the FSRQ object 3C 279.

[1]Another approach to detect VHE γ-ray photons is water Cherenkov technique in which charged particles of γ-ray-induced air showers are sampled on the ground using water tanks. This technique is used by e.g. the High-Altitude Water Cherenkov Gamma-ray Observatory (HAWC) located in Mexico. While the IACT approach is the most advanced one for studies of individual sources as it provides the best angular and spectral resolution, water Cherenkov technique remains highly complementary due to higher duty cycle and larger field of view.

3.1 VHE gamma-ray astronomy

3.1.1 Detection of VHE gamma-rays with an IACT

Particle cascades in the Earth's atmosphere

The Earth's atmosphere is not transparent to γ-ray photons. A VHE γ-ray arriving from an astrophysical source, after entering the Earth's atmosphere, will experience an inevitable interaction with the Coulomb field of the nuclei of the atmospheric atoms, which leads to creation of an electron-positron pair. This process is called *Bethe-Heitler pair production* (see left panel of the Fig. 3.1). Let us estimate the mean free path of the γ-ray in the air: $\lambda_{\gamma,\text{B-H}} = (\sigma_{\text{B-H}}\, n_{\text{air}})^{-1}$, where $\sigma_{\text{B-H}}$ is the Bethe-Heitler pair production cross-section and $n_{\text{air}} \sim 10^{21}$ cm^{-3} is the average number density of the atmospheric nuclei. A rough estimate of the cross-section can be done based on the corresponding leading Feynman diagram. Since the Feynman diagram of the Bethe-Heitler pair creation process (left panel of the Fig. 3.1) has three vertexes, according to the quantum electrodynamics (QED) perturbation theory, the cross-section of this process is proportional to the third order of the coupling constant of the QED (fine structure constant α), $\sigma_{\text{B-H}} \propto \alpha^3 \sim \alpha\, \sigma_{\text{T}}$, where σ_{T} is the Thomson cross-section. As the contribution of the nucleus Coulomb field to the scattering amplitude is $\propto Z\, e$, the cross-section will be $\sigma_{\text{B-H}} \sim Z^2 \alpha\, \sigma_{\text{T}}$, where Z is the nucleus atomic number. With this order of magnitude estimate of the cross-section, we find that the mean free path of a γ-ray in the Earth's atmosphere is

$$\lambda_{\gamma,\text{B-H}} \sim 1\,\text{km} \tag{3.1}$$

Indeed, the mean free path of the TeV γ-ray in the atmosphere is much shorter than the characteristic atmospheric height scale ($h_{\text{atm}} \sim 10$ km), meaning that the pair production will occur already in the high-altitude atmosphere layers. In reality, the atmosphere has a non-homogeneous density with altitude, and a more accurate description of the γ-ray propagation through the atmospheric medium is achieved by using the quantity named atmospheric depth $X_{\text{a}} = \int \rho(h)\, dh$, where $\rho(h)$ is the density of the atmosphere depending on the altitude h. The atmospheric depth can be understood as a quantity proportional to the total mass of the atmosphere encountered on the way of the particle. Then, in terms of the atmospheric depth, accounting for the non-homogeneity of the atmospheric density, γ-ray pair-production in the atmosphere will occur at an average atmospheric depth $X_{\text{a,B-H}}$

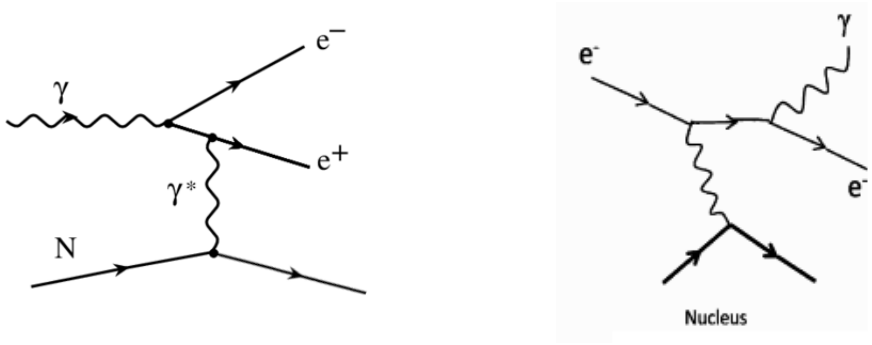


Figure 3.1: Feynman diagram of the Bethe-Heitler pair production process (*left*), and of the Bremsstrahlung process (*right*).

$$X_{\text{a,B-H}} = \frac{m_{\text{a}}}{\sigma_{\text{B-H}}} \tag{3.2}$$

where $m_{\text{a}} \approx A\, m_{\text{p}}$ is the average mass of atmospheric nucleus, with A being an atomic mass number (number of nucleons in a nucleus), and m_{p} proton rest mass. Typically, the first interaction of the primary γ-ray leading to pair production happens at altitudes around 20 – 30 km.

The resulting electron and positron, traveling down the atmosphere, interact with the same Coulomb field of atmospheric nuclei, and emit *Bremsstrahlung radiation* (see right panel of the Fig. 3.1), which is the main energy loss channel for charged particles having energy above a critical value of $E_{\text{cr}} \simeq 84$ MeV. This radiation is emitted by high-energy charged particles passing by heavier nuclei in a medium, the Coulomb field of which deflects and decelerates the passing charge, causing it to emit electromagnetic radiation (Larmor formula). The Bremsstrahlung radiation of the ultrarelativistic electron and positron is emitted in the γ-ray band, with photon energies up to those of the electron/positron (so that only one or a few photons are emitted). The charged particles continue emitting Bremsstrahlung γ-ray photons until their energy drops down to a critical one, E_{cr}, below which the ionization (non-radiative) losses become dominant. As a result, electron and positron dissipate their energy in Bremsstrahlung γ-rays. One could notice that the Feynman diagram of the Bremsstrahlung process (right panel of the Fig. 3.1) shows three vertexes as well, implying that its cross-section and hence the mean free path of charged particles before losing their energy on Bremsstrahlung emission are commensurate to the ones for the Bethe-Heitler pair production, $\lambda_{\text{e,bs}} \simeq \lambda_{\gamma,\text{B-H}}$. These secondary (Bremsstrahlung) γ-rays undergo the same pair production process, yielding electron-positron pairs, which again emit Bremsstrahlung γ-ray photons, and so on. The process of energy conversion from γ-rays to $e^{\pm}$ and vice versa is repeated many times. As a result of this

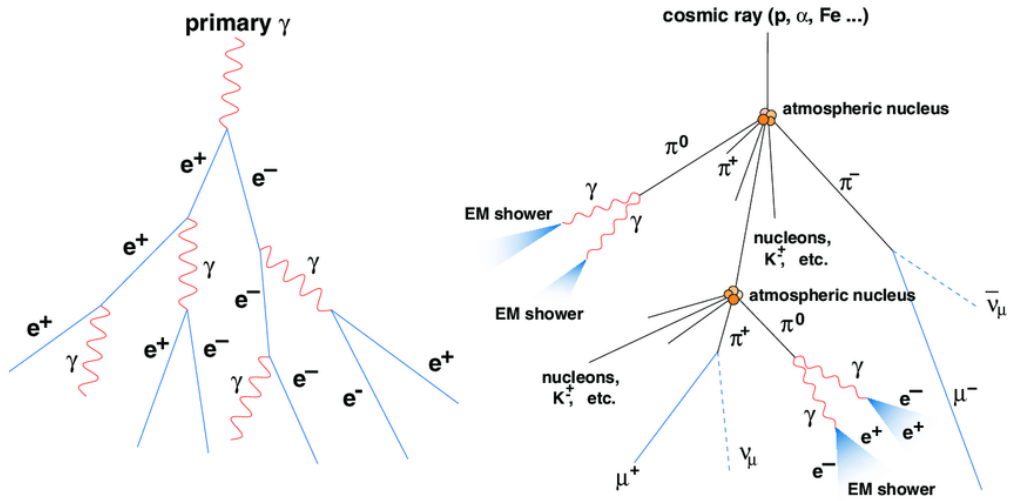


Figure 3.2: A schematic representation of an electromagnetic (*left*) and a hadronic (*right*) air shower. (adapted from Wagner (2006))

multiplicative process, a particle cascade develops in the Earth's atmosphere, called an **air shower**. The scheme of the shower initiated by a VHE γ-ray is depicted in the left panel of Fig. 3.2. After each cascading level, the number of particles (very roughly) doubles compared to the previous generation, and the energies of the particles become lower, due to distribution of the total energy over an increased number of particles. In addition, the mean free path of particles $\lambda_{\gamma,\text{B-H/bs}}$ will shorten for each next generation, since the atmospheric density increases for each next cascade level. The whole cascade develops until energies of particles (γ-rays and $e^{\pm}$) reach the critical value E_{cr}. At this point the further development of the cascade will be ceased, because all the charged particles will lose their energy via non-radiative ionization losses, rather than the Bremsstrahlung, and particle number multiplication will stop. Once the remaining electrons and positrons fully dissipate their energy on ionization, the air shower completely fades. The maximal achieved number of particles in the air shower is $N_{\text{as,max}} \simeq E_0/E_{\text{cr}}$, where E_0 is the energy of the primary γ-ray. Typically, the shower maximum is reached at altitudes of $h_{\text{as,max}} \simeq 10$ km (in the TeV range).

The cascade described above is named electromagnetic or leptonic, because only leptons constitute the air shower and only the electromagnetic interaction is involved (left panel of Fig. 3.2). However, not only γ-rays trigger cascades in the atmosphere of the Earth. *Charged cosmic ray particles* frequently hitting the Earth's atmosphere, initiate air showers with quite different particle content. These showers are named *hadronic*, as the primary particle is a hadron (e.g. a cosmic ray proton, α particle or any other nucleus), as well as a substantial fraction of secondary particles. When

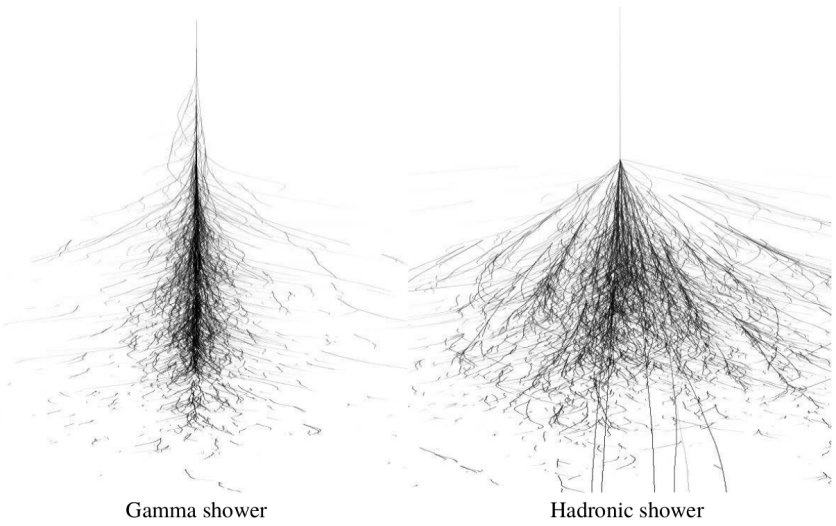


Figure 3.3: Computer simulation of a γ-ray-induced (*left*) and a hadron-induced (*right*) air shower. (adapted from Völk & Bernlöhr (2009))

a hadron penetrates the Earth's atmosphere, it collides with the atmospheric nuclei, leading to production of pions (π-mesons) and other mesons, particles consisting of a quark and an antiquark. This process is governed by the strong interaction, in contrast to leptonic showers, where particle creation is electromagnetic in nature. Three types of pions are produced, π^+, π^- and π^0. The neutral pion almost immediately decays (within $\sim 10^{-16}$ s) into two γ-ray photons, which initiate secondary electromagnetic showers. The lifetime of charged pions is much longer ($\sim 10^{-8}$ s), and they can either strike another atmospheric nucleus and create more pions, giving rise to a secondary hadronic shower, or simply decay into a muon and muon (anti)neutrino: $\pi^+ \rightarrow \mu^+ + \nu_\mu$, $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$. A sketch illustrating an air shower induced by a hadron is shown in the right panel of Fig. 3.2. Due to such "branchy" structure, hadronic showers are much more inhomogeneous and less symmetric when compared to γ-ray-induced showers. Another specificity of hadronic showers is that they are much wider in their extent, as particles produced in hadron-nuclei collisions in most cases get higher momenta perpendicular to the shower axis than in particle creation processes during the development of an electromagnetic cascade. This difference between the stucture of γ-ray-induced and hadron-induced air shower is illustrated in Fig. 3.3.

Muons created in hadronic showers are charged leptons, and interact with the atmospheric medium via Bremsstrahlung process. As they are heavy ($m_\mu \approx 207\, m_\mathrm{e}$), muons have a smaller Bremsstrahlung cross-section by a factor of $(\frac{m_\mu}{m_\mathrm{e}})^2 \approx 4 \times 10^4$ compared to the one for the electrons/positrons, and therefore their mean free path in the atmosphere is increased by the same factor, and is roughly $\lambda_\mu \sim 4 \times 10^4$ km. As

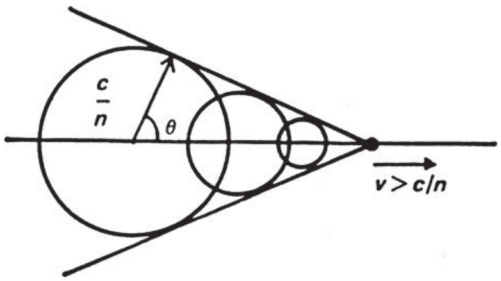


Figure 3.4: Huygens' construction for the wavefront of Cherenkov emission by a relativistic charged particle moving through a medium at a constant velocity exceeding the speed of light in this medium (adapted from Longair (2011))

this value is much longer than the height scale of the atmosphere, muons entering the Earth atmosphere will not produce any air showers, striking the surface of the Earth without interaction on the way. A fraction of them will decay into electron/positron and neutrinos: $\mu^- \to e^- + \bar{\nu_e} + \nu_\mu$, $\mu^+ \to e^+ + \nu_e + \bar{\nu_\mu}$.

Cherenkov radiation

A charged particle, moving in a medium at a constant velocity faster than the speed of light in this medium, emits *Cherenkov radiation*. This emission was experimentally detected for the first time by Pavel Cherenkov in 1934. The Cherenkov effect can be considered as an analog of a sonic shock wave (sonic boom) for the light. In a simple representation with Huygens' construction, the charged particle outruns the wavefronts forming around it, these wavefronts sum up coherently, and the radiation is emitted with the resulting emission wavefront having an angle with respect to the particle velocity vector (see Fig. 3.4). Such geometry leads to emission of the Cherenkov radiation in a cone with an opening angle

$$\cos\theta = \frac{c}{n_{\mathrm{r,m}}\, v} \tag{3.3}$$

where $n_{\mathrm{r,m}}$ is the refractive index of the medium, and v is the particle velocity.

The spectrum of the Cherenkov emission is given by the Frank-Tamm formula

$$\frac{dN}{d\lambda\, dx} = \frac{2\pi\,\alpha}{\lambda^2}\left(1 - \frac{c^2}{v^2\, n^2_{\mathrm{r,m}}(\lambda)}\right) \tag{3.4}$$

with α being the fine structure constant. The spectrum peak is located in the

ultraviolet part of the electromagnetic spectrum.

IACT technique

Charged particles produced in the air showers, move faster than the speed of light in the air, and therefore radiate Cherenkov emission. The refractive index of the air is $n_{\mathrm{r,air}} \approx 1.000293$ and the cascade particles move with velocities $v \sim c$, so according to the Eq. 3.3, a charged particle of the air shower will emit Cherenkov radiation in the direction of the ground in a cone with an angle $\theta_{\mathrm{air}} \approx 1.4°$. As a result, a circular area with a radius $R_{\mathrm{cher}} \simeq 100$ m on the ground will be illuminated, named Cherenkov light pool. The Cherenkov emission from an air shower arrives to the ground level in a very short time interval, from about 5 to 20 ns.

A ground-based IACT detects this very short flash of Cherenkov light. A typical setup includes a large segmented mirror (having mirror areas in the 10 – 100 m^2 domain), reflecting the Cherenkov light on a camera, which records the air shower image pattern. The camera typically employs photomultipliers (PMTs) to register Cherenkov photons, however other light sensors are used as well (see further below). Fast electronics is connected to the camera in order to sample the short-duration Cherenkov signal. Intensity, orientation and shape of the air shower image allow to recover properties of the primary particle, specifically its *energy*, *arrival direction* and *type* (γ-ray or cosmic ray) respectively. As the Cherenkov light from a single air shower gets distributed on the ground over an elliptical surface with a characteristic extent of $R_{\mathrm{cher}} \simeq 100$ m, this implies very large collection areas of the order of 10^8 cm^2 for a single IACT, which is some four orders of magnitude higher than typical collection areas of space-based pair-conversion instruments like *Fermi*-LAT. Such huge effective areas enable IACTs to be sensitive to very low TeV fluxes from distant astrophysical sources. The IACT observational technique is schematically illustrated in Fig. 3.5.

The *event reconstruction* is based on the following methods. The flux of Cherenkov photons at the ground level is proportional to the number of charged particles in the air shower, and hence to the energy of the particle that initiated the cascade. Therefore, the total shower image intensity correlates with the primary particle energy. The IACT technique could operate generally only above a few tens of GeV, since γ-rays of lower energy produce too few Cherenkov photons to be reliably detected above the fluctuations of the night sky background (NSB). The maximum energy accessible by the IACT technique is only limited by the telescope sensitivity, that depends on the area of the ground surface equipped with IACTs. Presently operating IACTs (see further below) are typically able to reach up to a few tens of TeV, exploring thus a window of the electromagnetic spectrum of roughly 3 decades. The orientation of the

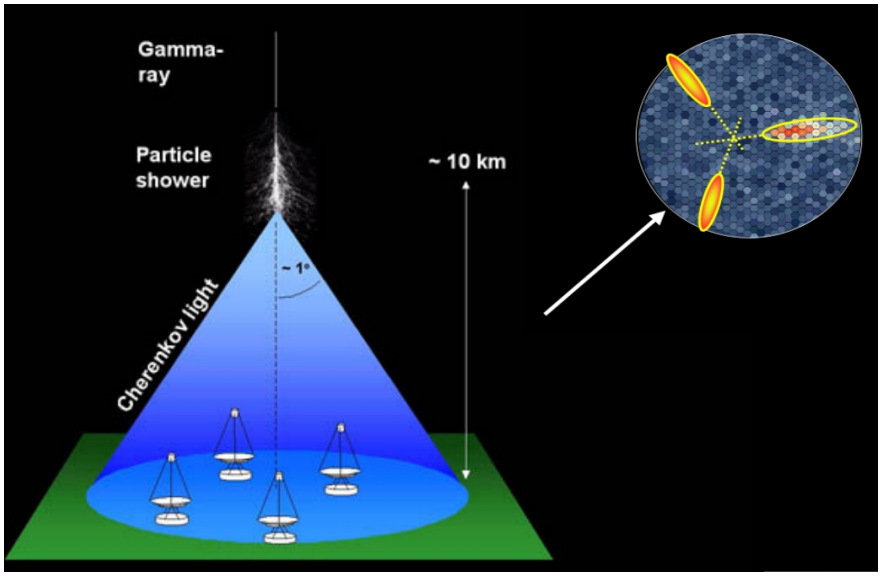


Figure 3.5: Illustration of the IACT observational technique. A VHE γ-ray enters the Earth atmosphere and initiates an electromagnetic cascade. Ultrarelativistic particles of the cascade emit Cherenkov light, which is detected by ground-based telescope(s). (image source: isdc.unige.ch)

air shower image in the camera plane is linked to the arrival direction of the primary particle. However due to projection effects, this reconstruction appears to be only two-dimensional. In order to restore the arrival direction of the particle in 3D, a *stereo-vision technique* is used. This approach is based on simultaneous imaging of an air shower from different viewing points, done by several (at least two) IACTs, placed at a distance comparable to R_{cher} from each other. The configuration in which individual telescopes are arranged on the ground is optimized to yield the most accurate possible reconstruction of the air shower geometry in 3D. Such IACT systems are conventionally referred to as "IACT arrays". The stereoscopic technique allows to achieve much higher accuracy in recovering the position on the sky from where the particle arrived compared to the monoscopic mode. In addition, IACT arrays generally offer larger effective area and therefore flux sensitivity, compared to single telescopes. Another advantage of the stereoscopic approach is a much better cosmic ray and NSB background rejection, discussed further below. Because of these reasons, most of the current-generation IACT setups consist of several telescopes.

The IACTs are usually placed at relatively high altitudes in the range of 1 – 3 km, so that, first of all, flux of Cherenkov photons that arrive from an air shower is increased, as the telescope(s) in this case is closer to the air shower maximum, and secondly, the observations are much less affected by low-altitude optically thick

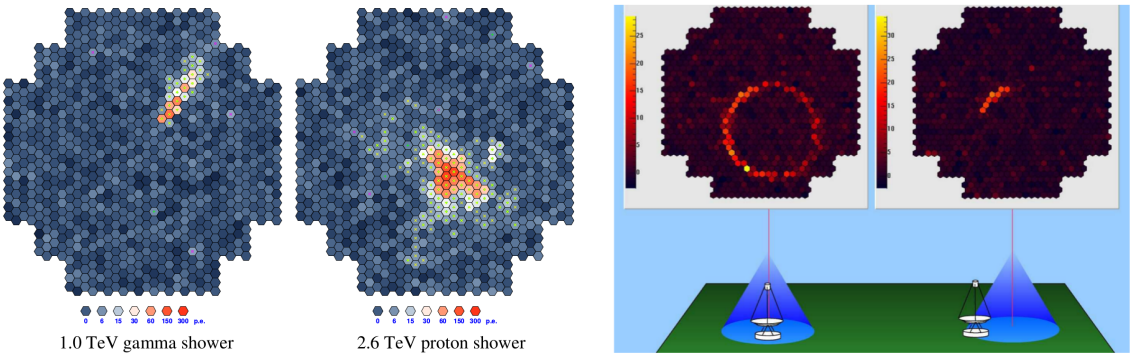


Figure 3.6: Difference between Cherenkov images of events caused by particles of different types (γ-ray, hadron and muon) seen by an IACT camera. *Left panel*: Cherenkov images of γ-ray-induced (left) and hadron-induced (right) air showers. One can see that the γ-ray events have an elliptical-like shape, whereas hadronic events display irregular and inhomogeneous morphology with multiple islands. *Right panel*: Cherenkov images of muon events for a muon hitting the telescope mirror and producing a ring image (left) and hitting the ground close to the periphery and producing an arc image (right). (adapted from Völk & Bernlöhr (2009))

clouds absorbing and scattering Cherenkov emission and perturbing the observations. IACT arrays are usually not installed at altitudes higher than $\sim$ 3 km to avoid the noise due to shower particles hitting the telescope mirrors.

As it was already mentioned, not only γ-rays induce showers in the atmosphere of the Earth, but also charged cosmic ray particles (mostly protons, and a fraction of heavier nuclei). These particles arrive from every direction on the sky nearly isotropically, as being charged they are deflected by the magnetic field of the Milky Way. Therefore, when an IACT observes a particular TeV-emitting source, *hadronic showers* will strongly interfere with the observations. Typically, the rate of hadronic showers is $\sim 10^5$ times higher than that of γ-ray-induced air showers. Hadronic events represent a background during VHE observations, which has to be rejected. The *cosmic-ray background suppression* is performed based on the air shower image shape. Images of γ-ray- and hadron-induced air showers have a very different morphology (see left panel of Fig. 3.6). One could see that γ-rays typically produce an elliptical-like rather uniform shower image, while hadronic events are far less homogeneous and symmetric, have an irregular shape and comprise several islands due to multiple sub-cascades. Muons produced in hadronic showers, as already mentioned, do not trigger air showers, and merely emit Cherenkov light as they move through the atmosphere. An isolated Cherenkov light cone of a muon directed towards the ground surface, projected on the camera plane, is seen by the telescope as a ring,

or as an arc depending on the impact distance (see right panel of Fig. 3.6). The IACT data analysis software uses algorithms that analyze the image morphology and based on that "recognize" the type of the primary particle that caused the event. When applied to an event sequence, this procedure allows to discriminate between γ-ray and cosmic-ray-induced events and therefore remove the cosmic-ray background. Stereo-imaging technique enhances the quality of γ-hadron separation, as more accurate and complete information on air shower development significantly reduces the chance of event misidentification. The relevant algorithms for event reconstruction are described in more detail in the sub-section 3.2.5.

The IACT technique has several limitations. Observations with IACTs can be only performed during the night, in absence of strong moonlight, as it is extremely difficult to detect faint Cherenkov light over bright moonlight. This substantially limits the observational time and duty cycle of IACTs. Another drawback of the IACT technique is a rather small field of view (FoV), typically about only a few degrees, so that simultaneous observation of multiple sources is often impossible, unlike with the *Fermi* Space Telescope, which has a wide FoV of 2.4 sr.

3.1.2 Presently operating IACTs

The first-ever operational IACT was *Whipple* (Cawley et al. 1990), which in 1989 detected its first TeV source, Crab Nebula (Weekes et al. 1989). In 1995, a Cherenkov telescope with a novel triple-dish system, *The University of Durham Mark 6 Telescope* (Armstrong et al. 1999), started its operations in Australia, and has detected multiple VHE γ-ray emitting sources (e.g. Chadwick et al. (1998) ; Chadwick et al. (1999)). The multiple-reflector design allowed to improve the sensitivity and decrease the low-energy threshold. Another instrument of that generation is *CAT* (Cherenkov Array at Thémis) (Barrau et al. 1998), built in France and started observing in autumn 1996. The *HEGRA* (High-Energy Gamma-Ray Astronomy) IACT system (Daum et al. 1997) was the first one consisting of multiple telescopes and employing the stereo-vision technique. The system was constructed on the Canary Islands, Spain. Since then, bigger instruments were constructed, in order to boost the number of detected sources, enhance the performance and improve the quality of data. The experience accumulated throughout operation of those IACTs was subsequently used in operation of their successors. The currently operating major IACTs are:

- H.E.S.S. : *High Energy Stereoscopic System*

  is an array of five telescopes situated in Namibia and operational since 2003 (Aharonian et al. 2004). More information on the instrument can be found in

Figure 3.7: *Top*: VERITAS array in Arizona, USA. *Bottom*: MAGIC system on Canary Islands in Spain. (Credits: MAGIC and VERITAS collaborations)

the section that follows.

- MAGIC: *Major Atmospheric Gamma-ray Imaging Cherenkov telescopes*

  is a system of two different IACTs, built in 2004 on the island of La Palma, Canary Islands (see bottom panel of Fig. 3.7) (Baixeras et al. 2004). The instrument is located on the former site of the *HEGRA* system. The MAGIC array is operated and supported by a collaboration of mainly Italian, Spanish and German institutions.

- VERITAS: *Very Energetic Radiation Imaging Telescope Array System*

  is an array comprising four IACTs, situated in Arizona, USA, and taking data since 2007 (see top panel of Fig. 3.7) (Holder et al. 2006). The instrument is a successor of Whipple, and is maintained by American institutions.

Figure 3.8: The full H.E.S.S. array in Namibia. (Credit: H.E.S.S. Collaboration, Frikkie van Greunen)

- FACT: *First Geiger-mode Avalanche photodiode Cherenkov Telescope*

  is the first Cherenkov telescope that employs Geiger-mode avalanche photodiodes (G-APDs) as photodetectors instead of photomultipliers, used by other instruments (Anderhub et al. 2013). FACT is a single IACT, located next to MAGIC system on La Palma island, and is operating since 2011. The FACT collaboration includes German and Swiss institutions.

3.2 H.E.S.S. experiment

3.2.1 Overview

The H.E.S.S. instrument is based in Khomas region in Namibia at an elevation of 1800 m above the sea level (see Fig. 3.8). The choice of the observing site was motivated, first of all, by gaining access to the (not yet explored at that time) Southern sky. Another important reasons for choosing the Namibian site is the absence of light pollution, dry climate favoring significantly reduced attenuation of the UV Cherenkov light by the atmospheric vapor, excellent atmospheric quality, and weather characterized by mostly clear sky and rare clouds. As of 2020, H.E.S.S. is the only IACT observing in the southern hemisphere, making the instrument especially suitable for

studies of Galactic sources and the Galactic center.

The array comprises four identical 12-meter-telescopes (CT1-4, see top panel of Fig. 3.9) and one very large telescope (CT5, see bottom panel of Fig. 3.9). Small telescopes are arranged in a form of a square with a side length of 120 m with a telescope at each corner, and the large telescope is placed at the center. CT1-4 started observing in December 2003 marking the beginning of the first H.E.S.S. phase (HESS-I), and CT5 was commissioned and started operations in July 2012, opening the second H.E.S.S. phase (HESS-II). CT5, having much larger mirror area, allows to access lower energies, extending the instrument's energy range down to ~ 30 GeV. At the moment, CT5 is the largest IACT on Earth. It also has the largest mirror collection area in the world. The small telescopes perform observations only jointly in the stereoscopic mode, while CT5 can operate in stereo-vision with CT1-4 or by itself in mono-vision regime.

3.2.2 The optical system

H.E.S.S. telescopes are single-dish, with CT1-4 employing a Davies-Cotton optical design. This optical system is commonly adopted for construction of IACTs, however other designs have been proposed for use in Cherenkov astronomy (see Section 7.2).

The mirror of CT1-4 has a diameter of 12 m, and a mirror area of 108 m^2. It is not a monolithic, but a segmented reflector, consisting of 380 spherical sub-mirrors with a diameter 60 cm each, having the same focal length as the entire dish, $f_{\mathrm{L,HESS\text{-}I}} = 15$ m (see the top-right part of Fig. 3.9).

The large telescope, CT5, has a mirror with a diameter of 28 m, and an area of 614 m^2. Contrary to HESS-I telescopes, the global shape of the CT5 mirror is a paraboloid, to ensure that the arrival time of Cherenkov photons in the focal plane is isochronous. The parabolic dish is composed of 876 hexagonal sub-mirrors having 90 cm diameter each (see the bottom panel of Fig. 3.9).

3.2.3 The cameras

HESS-I cameras are made up of 960 PMTs, arranged in 60 sub-units named drawers. Each drawer represents a set of 16 PMTs using the same front-end electronics, independent from that connected to the other drawers. Each PMT is equipped with a Winston cone, which guides Cherenkov photons and concentrates them on the photocathode, as well as rejects photons coming from an angle higher than 30° with respect

Figure 3.9: Different H.E.S.S. telescopes. *Top*: A close view of the 12-meter diameter H.E.S.S. telescopes (CT1-4) seen from different angles, showing the mechanical structure (left), and the segmented telescope mirror (right). Photographs taken during July 2018 shift on the H.E.S.S. site (A. Dmytriiev, 2018). *Bottom*: A close view of the large 28-meter H.E.S.S. telescope (CT5) (Credit: H.E.S.S. Collaboration, Frikkie van Greunen)

to the PMT's normal. Each PMT has an associated FoV of 0.16° on the sky, and the total FoV of the camera is 5°.

The camera of CT5 comprises 2048 PMTs organized in 128 drawers, offering therefore much higher Cherenkov image resolution compared to CT1-4. Each PMT has a FoV of 0.07°, with the total FoV of the detector of 3.2°.

3.2.4 The trigger system

H.E.S.S. system uses a trigger system to avoid recording night sky background fluctuations. This system is organized in three levels in order to decrease the load on the Data Acquisition System, and reduce the dead time during the event readout. At each level there is a condition that has to be satisfied, so in total three conditions have to be fulfilled at the same time for the array to record an event. At the first level, the internal trigger system of each individual telescope requires that a collected charge in at least one pixel exceeds the threshold of 4 photo-electrons (p.e.). At the second level, the correlation between pixel signals is examined: it is required that at least 3 pixels surpassing this threshold are located within a "sector", defined as four neighboring drawers, and that the signals are within a temporal window of 1.3 ns. At the third level, the central trigger system checks signal correlation between the individual telescopes: if at least two telescopes of the array have triggered the event (the first two levels were passed) within a coincidence window of 80 ns, the event is recorded. As a result, a typical trigger rate of the array is ~ 1 kHz.

Because of the third coincidence condition, the stereoscopic observational technique allows much better NSB background rejection than the monoscopic approach. CT5, when observing in a monoscopic regime, triggers events alone without the third coincidence condition, leading to a higher fraction of NSB noise in its data than for CT1-4.

3.2.5 Data analysis and reconstruction

Calibration runs

The raw signal at the output of the camera PMTs has to be reduced and calibrated in order to measure the number of Cherenkov photons which arrived at the photocathode throughout the event. The first required step is signal digitization, i.e. conversion of analog signal to digital, done by Analog-to-Digital Converter (ADC), translating the number of p.e. to ADC counts. Next, the pedestal has to be subtracted, which is a

signal of a PMT in absence of light. The estimation of the pedestal is done with a dedicated **pedestal run**, which represents a measurement of PMTs output with the camera lid closed, in the absence of Cherenkov or any kind of light. After pedestal subtraction, the flat-fielding is done, necessary to take into account uneven quantum efficiency across the PMT array. This is achieved via a **flat-field run**, during which the camera pixels are illuminated homogeneously by light-emitting diods (LEDs) mounted at the telescope dish, and the response is measured in each pixel. Finally, the conversion to number of Cherenkov photons is performed thanks to a **single photo-electron run**, during which LEDs send pulsed light flashes to the camera PMTs. The intensity of the light is adapted to cause emission of, on average, one single p.e. from the photocathode. Measurement of the light signal sent by LEDs allows to find out the pixel response to a single p.e. As a result of the calibration procedure, one obtains the effective number of detected Cherenkov photons integrated over the duration of the event in each pixel.

Event reconstruction

Once the event images are calibrated, the image cleaning is performed to remove the NSB, and then the events are ready for analysis. The goal is to reconstruct the energy and arrival direction of the primary particle on event-by-event basis, as well as to perform the hadronic background suppression. A qualitative approach was discussed in the sub-section 3.1.1, here we present more quantitative methods. There were several algorithms developed for this task, such as: (*i*) the Hillas analysis (Hillas 1985), (*ii*) the Model++ analysis (de Naurois & Rolland 2009), and (*iii*) the Multi-Variate analysis (Becherini et al. 2011). These methods are used in the analysis of H.E.S.S. data, with two different analysis chains being used for event reconstruction: **ParisAnalysis** and **HAP** (H.E.S.S. Analysis Package), for cross-check reasons. The former one uses the Model++ algorithm, while the latter one employs the Hillas and the Multi-Variate technique. In this thesis, the analysis of H.E.S.S. 3C 279 data (section 3.3) was performed with the help of the ParisAnalysis tools, therefore we will cover here only the basic Hillas method and the Model++ technique.

The Hillas analysis

The Hillas algorithm was the first one developed, and is the most standard in Cherenkov γ-ray astronomy. In this approach, the shower image is modeled with a 2D ellipse, which is then parametrized with the following parameters (Hillas parameters): the length of the semi-major and semi-minor axes, the center of gravity, the total number of photons contained within the ellipse, the position of the ellipse with respect to the camera center and the distance from the center of gravity to the expected

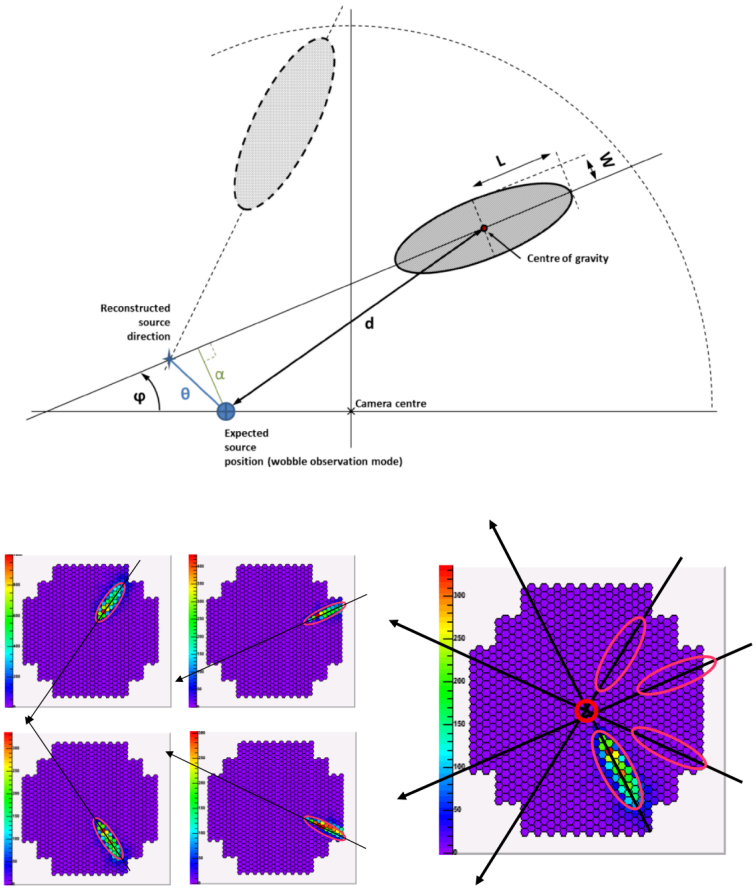


Figure 3.10: Event reconstruction in Hillas analysis. *Top*: scheme representing the Hillas parametrization (adapted from Garrigoux (2015)). *Bottom*: images of an air shower seen by four different telescopes (left) and geometric reconstruction of the true source position in the stereo-vision observational mode (right) (adapted from De Naurois (2012))

source position (see the top panel of Fig. 3.10). The energy of the primary particle is proportional to the total intensity within the ellipse (for a given distance and zenith angle). In the mono-vision mode, the arrival direction of the primary particle is that of the vector along the major axis of the ellipse, and the reconstruction of the true position of the source is degenerate, which is solved by introduction of additional shower image parameters (third-order moments, skewness and kurtosis). In the stereo-vision regime, the true source position is determined by intersection of the major axes of multiple differently oriented images of the same event, provided by the telescopes participating in stereo observations (four or five telescopes in the case of H.E.S.S.). The principle of source position reconstruction in stereo mode is illustrated in the bottom panel of Fig. 3.10. Finally, the discrimination between γ-like events and much

more numerous hadron events is based on the property that hadron events have a much broader radial spread of the Cherenkov signal. The cosmic ray background is suppressed by imposing a threshold value (a *cut*) on the width of the photon angular distribution: events having a signal spread wider than the cut value are classified as hadron ones, and are rejected.

The Model++ analysis

The event reconstruction algorithm, named Model, was initially developed by Le Bohec et al. (1998) and subsequently improved by de Naurois & Rolland (2009) for the analysis of H.E.S.S. data (Model++). This method is based on direct comparison of the event images detected by an IACT with a library of simulated events. This library is generated by (*i*) calculating the Cherenkov emission from an air shower described with a semi-analytical model, for different values of input parameters (γ-ray energy and arrival direction, relative telescope position, etc), and then (*ii*) convolving the resulting Cherenkov photon flux with the instrument response, yielding the predicted air shower image in the telescope's camera. The semi-analytical 3D model of air showers is obtained from Monte-Carlo-simulated cascades by parametrization of the longitudinal, lateral and angular distribution of the cascade particles.

An actual event image is fitted to the entries of the dictionary (γ-ray parameters $\Leftrightarrow$ simulated image), and the best match is searched for using the maximum likelihood approach, that takes into account the Poisson statistics of the real signal. The sought energy and the arrival direction of the γ-ray is given by the corresponding parameters of the best-fit simulated image.

The hadronic events are identified by comparison of the actual and the expected log-likelihood, allowing to estimate the consistency with the hypothesis that the detected event is γ-ray-induced. Further improvement of the background suppression is achieved by exploring this compatibility for two different groups of pixels contributing to the goodness-of-fit separately, rather than computing the log-likelihood for the entire recorded event image. These two groups are the shower pixels and the background pixels, with the contour of the boundary between them delineated by a defined threshold of the signal level in one pixel. The likelihood of the match for shower pixels is characterized by the *ShowerGoodness* parameter, which is sensitive to discrepancies between the real and simulated shower image, and the same measure for the background pixels is expressed via the *BackgroundGoodness* parameter, which is sensitive to islands lying outside the core of the shower, typical for hadronic events, as well as to various other discernible features of hadron-induced showers.

The Model++ analysis is about twice as sensitive as the Hillas reconstruction technique (de Naurois & Rolland 2009). It is used in this thesis for the analysis of

3C 279 data presented in the section that follows.

3.3 Analysis of H.E.S.S. data of 3C 279 flare

The FSRQ 3C 279 underwent two strong flares in January and June 2018 detected by *Fermi*-LAT and H.E.S.S. To analyze the two H.E.S.S. data sets, a dedicated *task force* within the H.E.S.S. collaboration was created, which is a group of scientists working together on the analysis with the responsibilities distributed among the task force members. I joined the task force motivated by a goal of analysis and subsequent physical modeling of the data collected during the two flares.

In this section I present the results of the analysis of the H.E.S.S. data of January and June 2018 flares of 3C 279 I performed, as well as relevant results obtained by the task force[2].

Certain aspects of the analysis of the 3C 279 flares still need to be reviewed by the H.E.S.S. collaboration before their final publication.

3.3.1 The studied source: 3C 279

3C 279 is an FSRQ, located at a redshift $z = 0.536$ (Marziani et al. 1996), highly variable from radio band to VHE γ-ray regime. The source harbors a SMBH with an estimated mass of $(3-8) \times 10^8\, M_\odot$ deduced using two independent methods (Gu et al. (2001) ; Woo & Urry (2002)). 3C 279 was monitored in nearly all spectral bands (e.g. Larionov et al. (2008) ; Hayashida et al. (2012)) and is one of the best studied objects of its class. The broad-band SED of 3C 279 represents a typical double-bump structure with a first bump peaking at the infrared, and the high-energy peak having maximum between 100 MeV and a few GeV. In the VHE γ-ray range, the source is only detectable during flares.

The VLBI observations of the source reported apparent superluminal velocities in the large-scale jet in the 4 – 20 c range, with the jet closely aligned (up to 2°) with the line of sight (Jorstad et al. 2004). The optical emission of the object shows a quite important polarization with a varying level reaching 45.5 per cent in the U-band

[2]Partial results of the presented work have been published in the proceedings of the 36th International Cosmic Ray Conference (ICRC2019) in a contribution "Observations of the FSRQ 3C 279 during the flaring state of 2017 and 2018 with H.E.S.S." by G. Emery, M. Cerruti, A. Dmytriiev, F. Jankowsky, H. Prokoph, C. Romoli and M. Zacharias (July 2019) (Emery et al. 2019)

(Mead et al. 1990).

3C 279 was the first FSRQ detected by the EGRET (Energetic Gamma-Ray Experiment Telescope) instrument aboard the *Compton* Gamma-Ray Observatory in the energy range from 30 MeV to 5 GeV (Hartman et al. 1992), and also the first FSRQ discovered as a VHE γ-ray emitter (MAGIC Collaboration et al. 2008). This source is one of the most distant VHE γ-ray emitting sources detected up to now. The detection of 3C 279 in the VHE band was a rather unexpected result, as the VHE γ-ray flux from the source should be suppressed due to internal absorption on the soft photon fields and due to EBL absorption. The discovery of 3C 279 as a VHE γ-ray emitter posed serious challenges to existing blazar radiative models (discussion in sub-section 4.1.1), as well as had profound implications on EBL models (discussion in Section 4.2), suggesting that the γ-ray opacity of the Universe was overestimated, and even stimulated scenarios involving exotic physics, specifically axion-photon conversion (see sub-section 7.1.2).

Since the launch of *Fermi* Gamma-Ray Telescope in 2008, the source was continuously monitored by this instrument owing to its wide FoV, with the first *Fermi*-LAT detection reported by Abdo et al. (2010b). 3C 279 showed multiple bright γ-ray flares. An outburst with a complex variability pattern, featuring a sequence of flares, was observed in December 2013, with the maximum γ-ray flux $F(> 100 \text{ MeV}) \simeq 10^{-5}$ ph cm^{-2} s^{-1}, and the shortest flux-doubling time-scale of 2 h (Hayashida et al. 2015). During this high-activity state, the γ-ray spectrum of the source underwent hardening up to $dN/dE \propto E^{-1.7}$, and displayed strong "Compton dominance", with a ratio of the total inverse Compton flux over the total synchrotron flux of the order of 100. Another giant flare of 3C 279 occurred in June 2015 (Cutini 2015), with the peak flux reaching even higher value of $F(> 100 \text{ MeV}) \simeq 3.9 \times 10^{-5}$ ph cm^{-2} s^{-1} (Paliya et al. 2015b).

3.3.2 VHE flares of 3C 279 detected by H.E.S.S. in 2018

3C 279 is one of the objects in the H.E.S.S. Target of Opportunity (ToO) program. Various instruments operating across the electromagnetic spectrum share information on interesting behavior of the sources they monitor, to ensure good MWL coverage during such episodes. *Fermi*-LAT detected 3C 279 in a flaring state in January, February and June 2018, sending an alert and triggering H.E.S.S. follow-up observations of the source. These three observational periods of H.E.S.S. are illustrated in Fig. 3.11.

In January 2018, H.E.S.S. observations were delayed by more than 7 days due to poor weather conditions on the site, and the *Fermi* flare peak was missed. H.E.S.S.

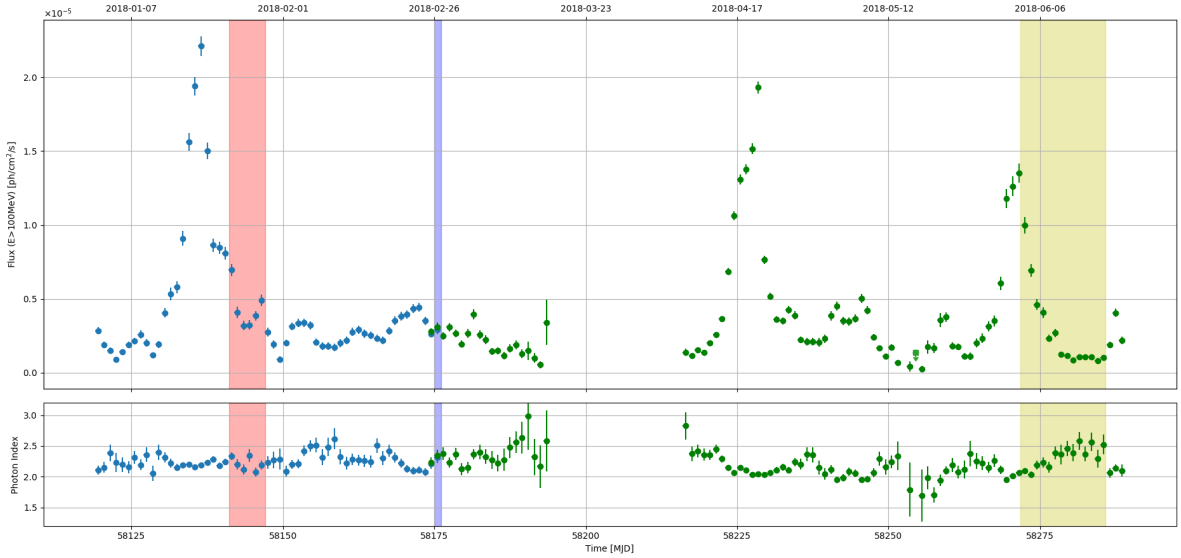


Figure 3.11: Illustration of three periods of H.E.S.S. follow-up observations of 3C 279 in 2018 after an alert from *Fermi*-LAT. Blue and green points in the top part of the figure indicate *Fermi*-LAT γ-ray photon flux of 3C 279 above 100 MeV, and the photon index in the bottom part. Red band marks the January 2018 ToO, blue band shows the February 2018 ToO, and the yellow band displays the June 2018 ToO. (Credit: C. Romoli and H.E.S.S. collaboration)

monitored the source over 6 consecutive nights during the January 2018 campaign, with a total observational time of 5.0 h. Flaring behavior was detected by H.E.S.S. during the night 27-28 January 2018, which appeared to be coincident with a modest secondary flare in the *Fermi* band (see Fig. 3.11). During that night, the source was detected at a significance level of 10.7 sigma accumulated during 1.7 h of observations (Emery et al. 2019).

In February 2018, H.E.S.S. failed to detect any significant signal from the source after 4 hours of observations.

In June 2018, observations by H.E.S.S. were carried out during the peak and decay of the *Fermi* flare (see Fig. 3.11), for 10 consecutive nights, followed by a 2 nights gap, and then monitoring for additional 3 consecutive nights, with the total observational time amounting to 18.7 h. The total significance during the June 2018 campaign is of 11.8 sigma (Emery et al. 2019).

My own contribution to the task force was performing analysis of specific subsets of H.E.S.S. data of January and June 2018 flares of 3C 279. For the January outburst, I focus on the pre-VHE-flare period (first 3 nights in the red band in Fig. 3.11), which

Setting name	Value
Software version	Model++ 0-8-32
Cuts	Mono CT5 Very Loose
Calibration version	default most recent
DST version	Model_HESSII_Hybrid_Prod8
Acceptance calculation	2D acceptance model
Background subtraction method	MultipleOff
Oversampling	0.1°

Table 3.1:: Table summarizing the configuration of the *ParisAnalysis* software used for the analysis of the 3C 279 January 2018 flare data set.

is a phase preceding the secondary *Fermi* flare. For the June high state, I perform a non-standard analysis of the H.E.S.S. data collected during the very peak of the *Fermi* flare (the first night inside the yellow band in Fig. 3.11). The peculiarity of the flare peak night in the June 2018 data set is that H.E.S.S. observations during that night were carried out without a background measurement (unlike the subsequent nights for which the background measurement was present), so my motivation to contribute to the task force by analyzing this specific night was to explore a non-standard analysis approach, as well as to reveal spectral properties of the source during the flare peak.

3.3.3 Analysis of H.E.S.S. 3C 279 data : January 2018 flare

For my analysis, I select H.E.S.S. data in a 3-night period before the source underwent the VHE flare, considering the time interval 23-26 January 2018.

The source has a quite weak flux above 100 GeV, and since the telescopes of the HESS-I array are not sensitive enough to the 3C 279 γ-ray flux at energies of a few tens of GeV, I selected for my analysis CT5 observational data taken in mono-vision regime. The data set comprises 7 observational runs that passed the quality check. The analysis of these runs was done with the help of *ParisAnalysis* chain using Model++ algorithm, with a "CT5 Mono" configuration profile. A summary of the full configuration of the analysis can be found in Tab. 3.1.

As a result, we find a weak indication for a γ-ray signal from the direction of the source at the level of 2.9 sigma for the 1st night, 2.4 sigma for the 2nd night and 0.7 sigma for the 3rd night. The source is therefore very marginally detected during the first two nights, and not detected at all during the third night. Overall, given the low significance at each night, we conclude that the data set under study is dominated by the noise and that the source is essentially *not detected* during the time range of interest.

The overall excess map, significance map and distribution, as well as the θ^2-plot showing the angular distribution of the events around the position of the source, are presented in Fig. 3.12.

The MWL behavior of the source during the entire time span of the January 2018 outburst was investigated in detail within the task force and is presented by Emery et al. (2019). Fig. 3.13 represents the set of MWL light curves during the January 2018 flare of 3C 279. One can notice a quite different flux variation pattern when comparing them between different energy bands.

3.3.4 Analysis of H.E.S.S. 3C 279 data : June 2018 flare

I analyze the H.E.S.S. one-night data during the flux maximum, reached on 1 June 2018. Similarly to the case of the January 2018 outburst, I use CT5 monoscopic data, with two observational runs constituting the data set. Observations during that night happened to be taken in an "ON SOURCE" mode, rather than *with a wobble offset* which is usually necessary for the γ-like background measurement (during the subsequent nights the observations were however carried out with a wobble). The γ-like background (or residual cosmic ray background) represents cosmic ray events misidentified as γ-rays and not rejected in the analysis during the γ-hadron separation step. *Absence of the background measurement* substantially complicates the analysis; *non-standard techniques* have to be employed. In the case of January 2018 data set, all the runs were taken with a wobble, making possible to use the standard algorithm named "MultipleOff", subtracting the background averaged over several off-source regions ("OFF" runs) from the measurements of the flux from the source direction ("ON" runs). In the case of the June 2018 peak flare data set, I apply so-called "ON-OFF" background subtraction technique, in which OFF runs taken at different epochs are used as a background estimate. These OFF runs have to be taken at similar observational conditions, as the analyzed data (for more information on the background subtraction methods, as well as the wobble observational strategy, see e.g. De Naurois (2012)). To find the matching runs, I use a dedicated "*ON OFF Selection*" tool provided in the *RunQuality* (run selection) component of the *ParisAnalysis* software, which searches for OFF runs automatically based on manually specified criteria. The following criteria were applied:

- Same number of telescopes involved in observations (only CT5)
- Zenith angle difference within no more than 5°
- Azimuth difference within no more than 5°

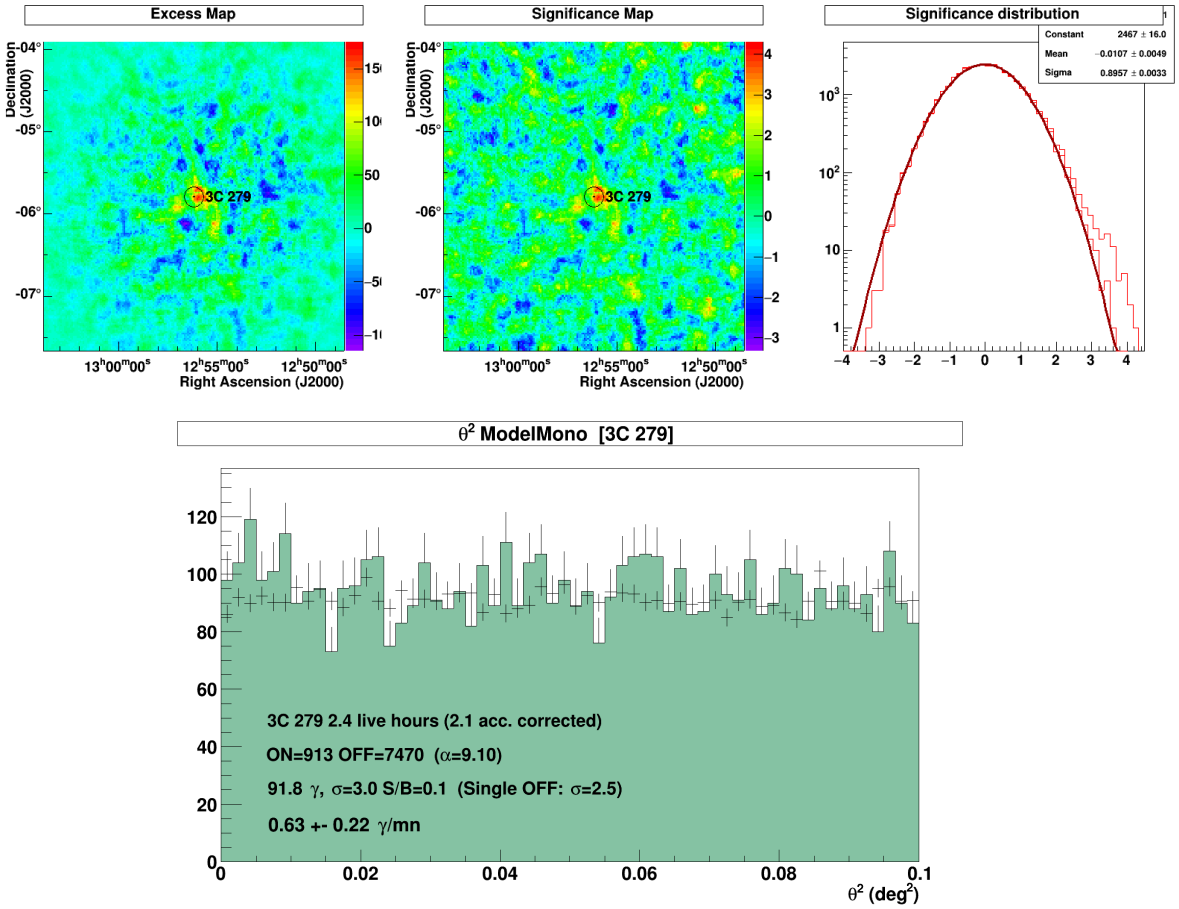


Figure 3.12: Different characteristics representing the event statistics for the analysis of the 3C 279 January 2018 flare data set. *Top panel*: excess counts map (left) in the FoV, significance map (middle) and the 1D significance distribution across the FoV (right). *Bottom panel*: θ^2-plot characterizing the angular (radial) distribution of the source and background events in the 0.1° circle around the source position, together with the information on the live observational time, counts statistics, overall significance, signal to background ratio and count rate. The source is essentially not detected during the studied time interval.

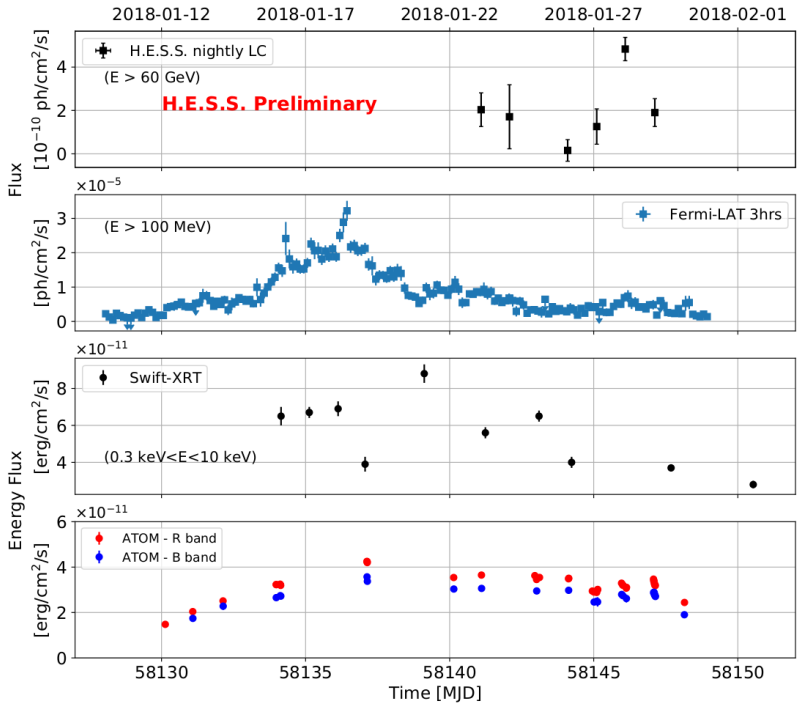


Figure 3.13: A set of MWL light curves of 3C 279 during January 2018 outburst. From top to bottom: (1) H.E.S.S. night-by-night photon light curve above 60 GeV (Mono Very Loose analysis) for the combined time interval including the pre-VHE-flare and the VHE flare periods, (2) *Fermi*-LAT photon light curve above 100 MeV with a 3 h time binning, (3) *Swift*-XRT light curve (energy flux) in the energy range from 0.3 to 10 keV, (4) optical light curve (energy flux) in the R- and the B-band with a nightly binning based on ATOM data. (adapted from Emery et al. (2019))

- Run duration difference within no more than 4 minutes

As a result, two OFF runs matching the two ON runs were found, and were paired into a run list. I analyze this combined data set using the same setup of the *ParisAnalysis* chain, as for the January 2018 flare, with the exception for the background and acceptance calculation methods. The full analysis configuration is summarized in Tab. 3.2. A clear detection of the source is achieved at the level of 11 sigma.

The spectrum was computed in the energy range 0.02 - 10 TeV with 5 bins per decade of energy. I required 2 sigma per point for calculation of the data points. The resulting spectrum is shown in the top panel of Fig. 3.14. The source's spectrum represents a power law with a photon index $\alpha_2 = 3.43 \pm 0.23$. The average source's γ-ray flux above 60 GeV during the night of the June 2018 flare peak is

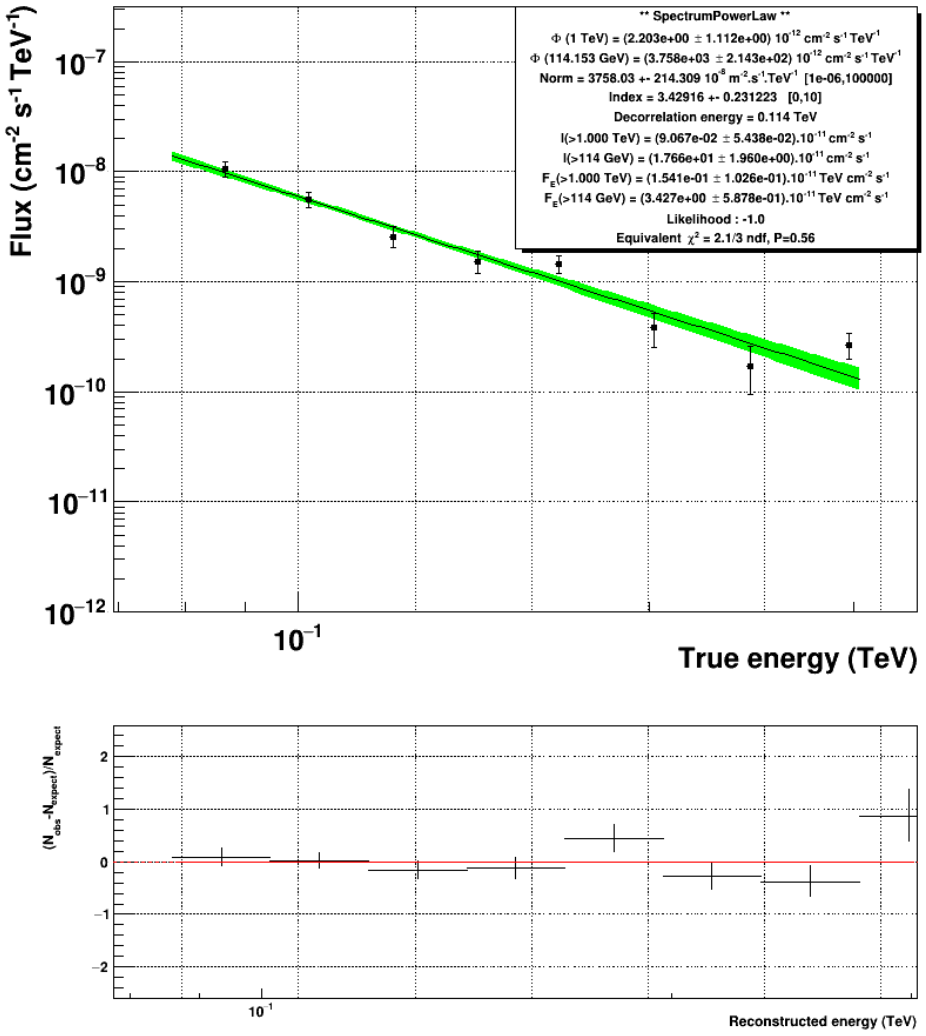


Figure 3.14: HE-to-VHE γ-ray spectrum of the 3C 279 during the peak of the June 2018 flare (1 June 2018).

Setting name	Value
Software version	Model++ 0-8-32
Cuts	Mono CT5 Loose
Calibration version	default most recent
DST version	Model_HESSII_Hybrid_Prod8
Acceptance calculation	Radial Acceptance (no zenith correction)
Background subtraction method	ON-OFF
Oversampling	0.1°

Table 3.2:: Table summarizing the configuration of the *ParisAnalysis* software used for the analysis of the 3C 279 June 2018 flare data set.

$$F_{\mathrm{peak,jun2018}}(>60\,\mathrm{GeV}) = (8.3 \pm 1.2) \times 10^{-10}\ \ \mathrm{ph\,cm^{-2}\,s^{-1}}$$

The MWL variability pattern of 3C 279 during the entire time range of the June 2018 flaring event was studied within the task force and reported in Emery et al. (2019). The corresponding compilation of MWL light curves is presented in Fig. 3.15. One can clearly see a rather strong correlation between the flux behavior in different spectral bands.

3.4 Discussion and perspective

As a member of the dedicated task force within the H.E.S.S. collaboration, I performed analysis of H.E.S.S. data of January and June 2018 flares of the FSRQ 3C 279, focusing on the pre-flaring state of the January 2018 VHE flare and on the very peak of the June 2018 outburst, as well as contributing to the general analysis and discussion of the results.

For the January 2018 pre-VHE-flare state, unfortunately no VHE γ-ray emission was detected with H.E.S.S., however it was possible to deduce upper limits on the source's flux during that time period.

For the June 2018 flare, the source was reliably detected by H.E.S.S. at VHE, both during the peak of the flare and its subsequent decay. I performed non-standard analysis of the peak flare data and calculated the source's spectrum from 20 GeV to 10 TeV, as well as the peak flux above 60 GeV.

The spectral index variation during the June 2018 flare was investigated within the task force. An average H.E.S.S. photon index of 5.0 ± 0.3 was obtained by the task force with the Mono Loose analysis for the period of the flare decay (time range

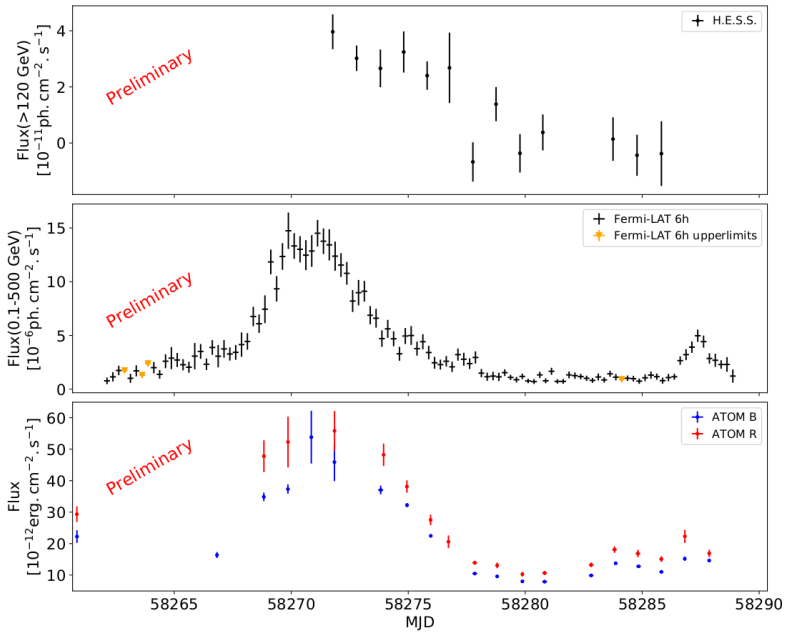


Figure 3.15: A set of MWL light curves of 3C 279 during June 2018 outburst. *Top*: H.E.S.S. night-by-night photon light curve above 120 GeV (Mono Loose analysis, assuming a power law spectrum with an index of 3.7) for the time period 2-16 June 2018 (flare decay, the peak not included). *Middle*: *Fermi*-LAT photon light curve in the 0.1-500 GeV range with a 6 h time binning. *Bottom*: optical light curve (energy flux) in the R- and the B-band with a nightly binning based on ATOM data. (adapted from Emery et al. (2019))

2-16 June 2018). For the flare peak, I measured the H.E.S.S. spectral index $\alpha_2 = 3.43 \pm 0.23$. Thus, spectrum during the flare peak appears to be harder than that during its decay. A noticeable hardening of the photon index also happened in the *Fermi* energy range above 100 MeV (see the bottom part of the Fig. 3.11): the index varied between $\sim$ 2.5 in the pre- and post-flare state, and $\sim$ 2 around the flare peak. The spectral hardening with an increasing flux in the γ-ray band is a trend that has been frequently observed during flaring events of different blazars (see sub-section 5.1.3 and Section 6.1).

One of the most natural explanations of this effect is acceleration of particles inside the source (see sub-section 5.1.3): migration of particles to higher energies leads to decrease of the number of particles in lower energy bins and a pile-up in the higher energy bins, and so the spectral slope becomes harder. However in the case of an FSRQ, another *interesting interpretation* might be considered. The observed spectral hardening might be related to the decrease of the optical depth of γ-γ pair production by VHE γ-rays colliding with low-energy external photons. This results

in a weaker absorption of the VHE flux of the source, which in the moderate optical depth regime, appears as a mitigation of the spectral cutoff strength at the pair production threshold, or, equivalently, hardening of the spectrum. The energy of the soft photons responsible for the effect can be estimated using the kinematic condition for pair-production at the threshold, given by the Eq. 4.22. From this expression, we get $E_{\rm soft} = \frac{m_{\rm e}^2 c^4}{E_\gamma}$, and for $E_\gamma \sim 100$ GeV, one obtains $E_{\rm soft} \sim 3$ eV. Therefore, soft radiation field should peak around the optical spectral domain, and so the target photons are most likely the BLR photons. The temporarily reduced optical depth can be explained by a decrease in the BLR photon field density in the vicinity of the VHE emission zone, arising from e.g. an obscuring event by a "cloud" in between, or, in general, a suddenly increased opacity of the medium between the BLR and the VHE γ-ray emitting region. However the decrease of the BLR radiation field density at the same time leads to the decrease of the intensity of the inverse Compton component (MeV-GeV energy regime), while it clearly strongly enhances (a flare). To produce the required enhancement, one can consider that the obscuring event is accompanied by strong particle injection into the emitting zone, appreciably outweighing the IC flux reduction due to the BLR photon field density decrease.

This part of the work however requires further investigation due to several reasons. Firstly, the MWL coverage during the June 2018 flare was unfortunately quite poor, precluding detailed physical modeling. Secondly, as a result of a cross-check performed by the task force, a non-negligible discrepancy was found between the *ParisAnalysis* and *HAP* results for this flare, namely in the spectral slope value and night-by-night fluxes. The origin of this inconsistency is currently being investigated by the task force. Once the cause of the discrepancy is understood and the results obtained by the two analysis chains converge, it will be possible to interpret the physical processes at work during this flare.

Chapter 4

Modeling of AGN emission: stationary models

As already discussed, blazars show non-thermal emission, spanning from radio frequencies up to TeV γ-rays. The emission, thought to be originating from the jet, is highly variable across the entire electromagnetic spectrum. The energy flux from the source can double over time-scales as short as a few minutes, and as long as a few months or even years. A significant flux increase over relatively short time-scales ($\lesssim 1$ week) is referred to as a flare. The shortest variability is observed in the VHE γ-ray band, with some flares showing a doubling time of only $3-5$ minutes (e.g. PKS 2155-304 (Aharonian et al. 2007) and Mrk 501 (Albert et al. 2007b)). The nature of the flaring behavior and physical processes responsible for initiating it remain unclear. Very strong and rapid variability in the TeV γ-ray range indicates an activity of processes involving very high-energy particles. In order to get a comprehensive picture of the flaring phenomenon, we need to address two broad problems: (*i*) the origin of the blazar broad-band emission, and (*ii*) violent processes in AGN jets driving the dramatic flux amplification.

If sampled at different flux states, the MWL blazar emission encodes the information not only about its nature, but also about the change of the physical conditions in the jet leading to the observed flux variations. Study of the observed emission therefore represents a powerful tool to probe physical conditions and processes inside the blazar emitting zone. Physical modeling of the spectral and timing properties of the blazar emission is a primary method to reveal the emission mechanisms inside the jet, as well as to test various scenarios designed to explain flaring activity.

In this chapter, we will focus on the origin of the steady blazar emission, discussing various mechanisms of production of the radiation in an instantaneous case.

In Chapter 5, we will consider diverse physical processes in the jet that can cause the emission variability and formulate the general time-dependent framework, as well as present a numerical code I developed to simulate the time evolution of the broad-band emission during flaring events, based on the time-dependent approach. Finally, in Chapter 6, I model with my code an exceptional outburst of BL Lac object Mrk 421 detected in February 2010 with a goal to get an insight into the origin of this flaring event and high-energy processes at work.

4.1 Origin of blazar broad-band emission

4.1.1 Leptonic (SSC and EC) and hadronic models

As already mentioned in Chapter 2, a typical broad-band SED of a blazar shows two distinct bumps. The lower-energy bump represents a superposition of the synchrotron radiation of relativistic electrons moving in the jet in the magnetic field, and the host galaxy emission peaking in the optical domain.

The origin of the second bump of the blazar SED is less obvious. In **leptonic** scenarios, the higher-energy component is interpreted as the inverse Compton (IC) scattering of soft photons of a radiation field by the same electron population (see Fig. 4.1). In the **Synchrotron Self-Compton (SSC)** scenario, the high-energy electrons upscatter the photons of the synchrotron emission they themselves produce (Maraschi et al. (1992) ; Bloom & Marscher (1996)). This is now a conventional scenario to explain the emission of BL Lac objects. For the FSRQs, the seed photon fields are dominated by external radiation fields, typically that of the accretion disk, BLR or of the dusty torus (Dermer et al. (1992) ; Sikora et al. (1994) ; Błażejowski et al. (2000)). This scenario is referred to as **External Compton (EC)**.

The models described above are referred to as leptonic, because only leptons are responsible for production of the observed broad-band emission, namely electrons and positrons. However it is not a-priori clear whether the jet is filled with electron-positron or electron-proton plasma. Leptonic models suppose that, even if some amount of protons is present in the jet, their contribution to the observed γ-ray emission is negligible due to their insufficiently high energies, which is explained by a limited power of particle accelerator operating in the jet. This statement if however not certain, as AGN are considered as one of the candidates for being sources of Ultra High Energy Cosmic Rays (UHECR) with energies up to $\sim 10^{20}$ eV. Therefore one could consider that the blazar emitting zone constituent particles include hadrons with extreme energies. So-called **hadronic** emission models assume that the γ-ray

SED component is described as due to synchrotron radiation of extremely high-energy protons ($\sim 10^{19}$ eV) (Aharonian 2000), or interaction of protons with the photon fields, leading to pion production and initiating particle cascades and generation of synchrotron emission by secondary electrons and muons (Mannheim 1993), or combination of the both effects (Mücke & Protheroe (2001) ; Mücke et al. (2003)). The lower-energy bump within the hadronic framework is still interpreted as synchrotron emission of leptons.

Mixed **lepto-hadronic** models assume that the particle population producing the observed emission represents a mixture of electron-positron and electron-proton plasma, and that the leptonic SSC emission and the emission generated by the cascades initiated by p-γ interaction, provide a comparable contribution to the high-energy SED component (Cerruti et al. 2015).

The leptonic description often appears to be preferable, as it possess several important virtues. To produce the observed multi-TeV emission, in leptonic model one needs sub-TeV or TeV energies of electrons in the source (assuming IC scattering in Klein-Nishina regime, $E_\gamma = \delta_\mathrm{b} E'_\mathrm{IC} \simeq \delta_\mathrm{b} E_\mathrm{e}$), which are fairly easily achievable with the shock acceleration mechanism (contrary to the required $\sim 10^{19}$ eV protons). Next, it is easier for leptonic models to explain fast blazar variability. The observed time-scales of flux variations in the GeV/TeV bands (e.g. hours-days) appear to be very similar to the cooling time-scales (in the observer's frame) of high-energy electrons producing this γ-ray emission in the magnetic field of 0.01 - 1 G, while for hadronic models, a quick cooling is rather difficult to achieve with reasonable physical conditions in the source. Especially challenging in the hadronic view is to describe fast ($\lesssim 1$ h) variability (e.g. Gaidos et al. (1996) ; Albert et al. (2007b) ; Aharonian et al. (2007)) (which is however also not so easy to explain in the leptonic scenario). In addition, leptonic models predict faster variability in the higher-energy spectral bands. Indeed, as the cooling time-scale is $t_\mathrm{cool} \propto \gamma^{-1}$ (see Eq. 5.8), and the inverse Compton effect boosts the soft photon energy by a factor of γ^2, $E_\mathrm{IC} = \delta_\mathrm{b} \gamma^2 E'_\mathrm{soft}$, the time-scale of variability at the energy E_IC will be $t_\mathrm{var} \propto E_\mathrm{IC}^{-1/2}$. This implies more rapid variability with increasing photon energy in the γ-ray band, in particular, an order of magnitude shorter time-scales at TeV energies compared to GeV range. Such behavior is very well compatible with available low- and high-energy γ-ray variability patterns of different blazars. Another important consequence of the leptonic SSC models is that the X-ray flux should be very well correlated with the γ-ray one, as the very same electrons are responsible to generate both X-rays and γ-rays, and any change in the electron distribution will be reflected in both bands simultaneously. In hadronic framework, such correlation is generally not expected, as the high-energy component arises due to a different particle population. This strong

correlation predicted by leptonic scenarios was found in various simultaneous X-ray and γ-ray observations of multiple blazars (e.g. Catanese et al. (1997); Sambruna et al. (2000) ; Albert et al. (2007a) ; Fossati et al. (2008) ; Aleksić et al. (2015) ; Ahnen et al. (2016)). It should be noted, that although such behavior was seen in most of the data sets, some exceptions from it exist. For example, during a few flares of Mrk 421 detected at TeV energies, the X-ray flux did not show any variations (so-called "orphan" flares, e.g. Błażejowski et al. (2005) ; Fraija et al. (2015)), meaning that the X-ray/γ-ray correlation was absent.

The leptonic scenario can also naturally explain the so-called *blazar sequence* (see sub-section 2.4.3). Different positions of the synchrotron peak for FSRQs and different types of BL Lacs (LBLs, IBLs and HBLs), can be attributed to shift of the high-energy break/cutoff in the spectrum of electron population in the emitting zone. The position of this break/cutoff is determined by the balance between the acceleration rate and the energy loss rate. One can assume that the time-scales of acceleration process in different blazar types are roughly comparable (e.g. because shocks have similar physical parameters). Electron cooling usually represents a dominant loss channel for high-energy electrons, and its rate increases with the higher density of ambient photon fields (see sub-section 5.1.2 and the Eq. 5.7). In more luminous objects, with higher intensity of the synchrotron bump, a stronger cooling leads to earlier break/cutoff in the electron spectrum, causing the synchrotron and IC peak to move to lower frequencies. In addition, higher radiation densities in more powerful blazars imply also higher opacity for VHE γ-rays due to γ-γ pair production, inducing further shift of the IC peak down to lower energies.

Böttcher & Dermer (2002) connect the different objects of the blazar sequence in an evolutionary scenario, in which the initially powerful accretion of matter on the black hole is gradually becoming weaker. Because of that, the amount of circumnuclear dust or BLR clouds scattering the disk photons is decreasing, leading to lower and lower contribution of external Compton effect. As the result, a blazar experiences transition from FSRQ to LBL and then eventually to HBL. This view is however not unique, many attempts have been made to unify the objects of the sequence with a minimum number of parameters defining the observed difference; also the sequence is a subject of many debates, and we will not review these questions here.

Overall, leptonic models are able to successfully model the MWL emission of different classes of blazars (e.g. Katarzyński et al. (2001) ; Katarzyński et al. (2003) ; Fossati et al. (2008) ; Celotti & Ghisellini (2008) ; Ghisellini et al. (2010) ; Böttcher et al. (2013) ; Aleksić et al. (2015)). In particular, pure one-zone SSC models were found to describe broad-band instantaneous SEDs of the majority of the HBLs detected so far. Despite this success, Boettcher (2012) identified a few HBLs, the MWL

emission of which appeared to be much better reproduced with either SSC with an addition of external photon fields, or with hadronic models. Hadronic models also remain a viable alternative to the SSC approach when modeling FSRQs. For example, the VHE γ-ray emission detected from 3C 279 by MAGIC (MAGIC Collaboration et al. 2008) is quite challenging to explain within the leptonic scenario by BLR soft photons scattering off the high-energy electrons. The VHE γ-rays produced via this process, should experience a severe attenuation due to γ-γ absorption when colliding with the low-energy BLR photons, which does not agree with the observed flux level. The multi-band 3C 279 spectral data is however very well represented by hadronic models, in which photo-hadronic interaction occurs either only on the local synchrotron photons, or on both synchrotron and BLR radiation fields (Böttcher et al. 2009).

Signatures of non-negligible contribution of hadronic processes to the emission are also shown by ultra-high-frequency-peaked BL Lac objects (UHBL; also known as extreme BL Lac). The high-energy bump of these sources peaks in the ~ 1 TeV range, one decade in energy higher than for HBL objects. They also show very hard TeV spectra. In order to describe SEDs of UHBL with leptonic SSC models, one has to assume extreme values of physical parameters, in particular, the Doppler factor of the γ-ray emitting zone has to be (typically) $\delta_{\rm ez} > 50$, and the minimal Lorentz factor of the electron population has to be implausibly high. To overcome those difficulties, Cerruti et al. (2015) propose a lepto-hadronic origin of the UHBL emission. The authors are able to satisfactorily reproduce the observed SEDs of several extreme BL Lacs with a set of reasonable physical parameters.

4.1.2 Blob-in-jet model

As already discussed in Chapter 2, the blazar γ-ray emission is thought to be produced in a compact region inside the jet, a *blob*, having density and magnetic field higher than in the large-scale jet, and filled with high-energy particles (so-called "blob-in-jet" model, presented in e.g. Katarzyński et al. (2001)). This is a quite common physical picture of the blazar γ-ray production site, but not the only one considered in the literature, and alternative options were proposed by various authors. For instance, Ghisellini et al. (2005) advocate a scenario in which the blazar γ-ray emission originates in the spine-sheath structure of the jet, and is produced by the inverse Compton scattering of the Doppler-boosted radiation of the sheath on the electrons in the spine, and vice versa. The beaming of the seed photon field seen in the frames of both spine and sheath appears since these components move with different Doppler factors, and leads to an increased intensity of the inverse Compton emission. In this

thesis, we explore the "blob-in-jet" view and adopt it in our time-dependent model presented in Chapter 5, and examine its applicability for describing the emission of Mrk 421.

Following this scenario, we assume that the VHE γ-ray emission of blazar originates from a compact relatively dense region in the jet (a "blob"), having spherical geometry and a radius $R_{\rm b}$. The blob contains electron-positron plasma having a uniform density, with a tangled magnetic field of a uniform strength B. The plasma blob is traveling along the jet axis at a speed close to the speed of light, with a Doppler factor $\delta_{\rm b} = [\Gamma\,(1 - \beta\cos\theta)]^{-1}$. Electrons in the region are randomly oriented (in the blob frame) due to the magnetic field configuration. For convenience, from this moment we will refer to electrons and positrons as simply electrons.

The electron population residing in the blob, is characterized by a electron spectrum, which is a number of electrons per unit of volume and per unit of Lorentz factors interval.

$$N_{\rm e}(\gamma) = \frac{dN_{\rm e}}{dV\,d\gamma} \tag{4.1}$$

As we suppose an homogeneous plasma inside the blob, the electron spectrum is then simply the number density of electrons per unit of Lorentz factor intervals. The electron distribution may evolve in time (see Chapter 5).

The electron spectrum is non-thermal, and the presence of high-energy particles in the blob is caused by their (pre-)acceleration. We consider here that an acceleration process supplying high-energy electrons, is able to boost particles only up to a certain Lorentz factor, $\gamma_{\rm max}$. At low energies, the minimal Lorentz factor of electrons in the distribution is $\gamma_{\rm min}$, which satisfies $\gamma_{\rm min} \geq 1$. Various efforts on modeling of blazar emission indicate that minimal Lorentz factor in the electron spectrum may be much higher than 1, at the order of $\gamma_{\rm min} \sim 10^3$ (e.g. Abdo et al. (2011)). This situation is discussed in more detail in sub-section 6.2.2, as well as its presumable origin. The minimum and maximum Lorentz factors of electrons $\gamma_{\rm min}$ and $\gamma_{\rm max}$ limit the span of the electron distribution in Lorentz factor space.

Let us consider a few examples of simple parametrizations of particle spectra with a minimal number of parameters. The most simple form is a power law

$$N_{\rm e}(\gamma) = K\,\gamma^{-\alpha} \tag{4.2}$$

where K is the density factor.

Another example is a broken power law

$$N_{\rm e}(\gamma) = \begin{cases} K_1\,\gamma^{-\alpha_1}, & \text{for } \gamma_{\rm min} \leq \gamma \leq \gamma_{\rm br} \\ K_2\,\gamma^{-\alpha_2}, & \text{for } \gamma_{\rm br} < \gamma \leq \gamma_{\rm max} \end{cases} \tag{4.3}$$

where $\gamma_{\rm br}$ is the Lorentz factor of the break. The factors K_1 and K_2 are related via $K_2 = K_1\,\gamma_{\rm br}^{\alpha_2-\alpha_1}$.

Another form is a power law with a cutoff (e.g. an exponential cutoff)

$$N_{\rm e}(\gamma) = K\,\gamma^{-\alpha}\exp(-\gamma\,/\,\gamma_{\rm cut}) \tag{4.4}$$

where $\gamma_{\rm cut}$ is the Lorentz factor of the high-energy cutoff.

The electron spectrum can also have a global significant curvature; a prototypical shape is a log-parabola of the form

$$N_{\rm e}(\gamma) = K\,\left(\frac{\gamma}{\gamma_0}\right)^{-a_{\rm LP}\,-\,b_{\rm LP}\,\log(\gamma/\gamma_0)} \tag{4.5}$$

where γ_0 is the pivot Lorentz factor, $a_{\rm LP}$ is the spectral index at $\gamma = \gamma_0$, and $b_{\rm LP}$ is the curvature parameter telling us the change of the spectral slope as the Lorentz factor increases by a factor of 10.

These and other simple parametrizations of the electron spectrum were used in modeling of different blazar SEDs (e.g. Katarzyński et al. (2001) ; Abdo et al. (2011)).

The particle population inside the blob radiates. As the high-energy electrons are moving in a magnetic field, they emit synchrotron radiation. In the "blob-in-jet" view that we consider, the total synchrotron emission of the jet therefore comprises two synchrotron components: the one produced by the electron population in the large-scale jet, dominating in the radio band (due to a lower magnetic field), and the one generated by the electron population in the blob, prevailing at higher frequencies (due to a higher magnetic field).

We adopt the leptonic scenario for the origin of the γ-ray emission, in which it is produced by the IC scattering process. In the "blob-in-jet" configuration, the high-energy particles of the blob upscatter either synchrotron photons they themselves generate (SSC), or external photon fields (EC). BL Lac objects are typically well-described with the pure SSC model, however in certain types of blazars, in particular LBLs, being intermediate objects between BL Lacs and FSRQs, the synchrotron

radiation of the large-scale jet may be a very important target for upscattering by the ultra-relativistic electrons of the γ-ray emitting zone (Hervet et al. 2015). Also, in two- or multiple-zone models for variable blazar emission (see sub-section 5.3.3), the effects of cross-scattering between the particle populations and emissions could be quite important: electrons of one emitting region can interact with radiation field(s) of another, neighboring emitting region(s). As we apply our modeling to an HBL (Mrk 421), we neglect the upscattering of the extended jet synchrotron emission, and adopt the SSC scenario for the origin of the γ-ray emission, however taking into account possible cross-scattering effects.

A lower bound can be placed on the radius of the blob by recalling the causality condition. A significant variation of the flux on the observed time-scale of $t_{\rm var,obs}$ can only occur if physical conditions change over the entire volume of the emitting region. Such change can only happen within 1 light crossing time of the blob or longer. This leads to

$$R_{\rm b} \; < \; \frac{c\, t_{\rm var,obs}\, \delta_{\rm b}}{1+z} \tag{4.6}$$

where z is the redshift of the source. The variations of physical conditions in the blob, causing the flux variability, can include e.g. a change of the magnetic field or radius of the blob, variations of particle spectrum, etc.

4.1.3 Synchrotron emission

Let us now find out the spectrum of synchrotron emission that a particle population radiates. In the blob frame, due to tangled magnetic field, electrons will produce emission with an isotropic pattern (in the observer's frame the emission will be however strongly beamed towards the observer). For the moment, we do all the computations in the blob frame (all the quantities are also in the blob frame and denoted with primes where important to indicate where the quantity is measured), and transform them to the observer frame later on.

The synchrotron emissivity [J s^{-1} Hz^{-1} sterad^{-1}] of a single electron, averaged over an isotropic pitch angles distribution, is given by (e.g. Crusius & Schlickeiser (1986))

$$P_{\rm s}(\nu', \gamma) \; = \; \frac{3\sqrt{3}\,\sigma_{\rm T}\, c\, U_B\, \zeta^2}{\pi\, \nu_B} \; \times \; \left[K_{4/3}(\zeta)\, K_{1/3}(\zeta) \; - \; \frac{3}{5}\,\zeta \left(K^2_{4/3}(\zeta) - K^2_{1/3}(\zeta) \right) \right] \tag{4.7}$$

where $\sigma_{\rm T}$ is the Thomson cross-section, $\zeta = \nu'/(3\,\gamma^2\,\nu_B)$, $\nu_B = eB/(2\pi\,m_{\rm e})$, $U_B = B^2/(2\mu_0)$, $K_{\rm a}(\zeta)$ is the modified Bessel function of order a.

The synchrotron emissivity [J s^{-1} cm^{-3} Hz^{-1} sterad^{-1}] of the electron population having the electron spectrum $N_{\rm e}(\gamma)$ is

$$j_{\rm s}(\nu') \;=\; \frac{1}{4\pi}\int_{\gamma_{\rm min}}^{\gamma_{\rm max}} N_{\rm e}(\gamma)\,P_{\rm s}(\nu',\gamma)\,d\gamma \tag{4.8}$$

The intensity of the synchrotron emission $I'_{\rm s}(\nu')$ emanating from the blob is governed by the radiation transport equation of a form $\frac{dI'_{\rm s}(\nu')}{dl'} = j_{\rm s}(\nu') - \varkappa_{\rm s}(\nu')\,I'_{\rm s}(\nu')$, where l' is the length along the photons path, $\varkappa_{\rm s}(\nu')$ is the absorption coefficient [cm^{-1}] due to synchrotron self-absorption process. This is however a simple case of the transfer equation for cylindrical geometry. We use here a solution of a more complex case of the transfer equation in spherical geometry (e.g. Bloom & Marscher (1996) ; Kataoka et al. (1999))

$$I'_{\rm s}(\nu') \;=\; \frac{j_{\rm s}(\nu')}{\varkappa_{\rm s}(\nu')}\left(1 \;-\; \frac{2}{\tau^2}\left[1 \;-\; e^{-\tau}\,(\tau+1)\right]\right) \tag{4.9}$$

where $\tau = 2R_{\rm b}\,\varkappa_{\rm s}(\nu')$. This quantity is of the order of the optical depth of the synchrotron self-absorption. The Eq. 4.9 describes the intensity of the synchrotron emission emitted by the electrons in the blob.

The absorption coefficient of the synchrotron self-absorption is given by (e.g. Ghisellini & Svensson (1991))

$$\varkappa_{\rm s}(\nu') \;=\; -\frac{1}{8\pi\,m_{\rm e}\,\nu'^2}\int_{\gamma_{\rm min}}^{\gamma_{\rm max}} \frac{N_{\rm e}(\gamma)}{\gamma\,(\gamma^2\;-\;1)^{1/2}}\,\frac{d}{d\gamma}[\gamma\,(\gamma^2\;-\;1)^{1/2}\,P_{\rm s}(\nu',\gamma)]\,d\gamma \tag{4.10}$$

In the regime where absorption is negligible, by taking a limit $\tau \to 0$ in the Eq. 4.9 (by e.g. performing Taylor expansion of the exponent term), the expression for the synchrotron intensity reduces to:

$$I'_{\rm s}(\nu') \;=\; \frac{4}{3}\,j_{\rm s}(\nu')\,R_{\rm b} \tag{4.11}$$

and the luminosity of the synchrotron emission per unit of the frequency interval in the blob frame is

$$L'_{\rm s}(\nu') = 4\pi^2\,R_{\rm b}^2\,I'_{\rm s}(\nu') \tag{4.12}$$

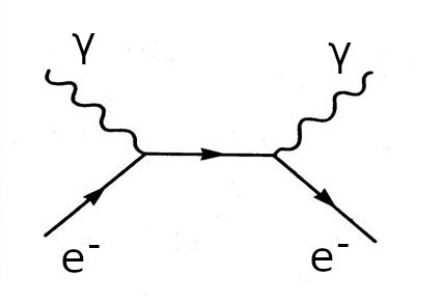


Figure 4.1: Feynman diagram of the (inverse) Compton scattering process

4.1.4 Inverse Compton emission

Let us now calculate the spectrum of the SSC emission that the electron population produces by interacting with the synchrotron spectrum. The relativistic electrons upscatter the synchrotron photons, boosting them in energy. The relevant Feynman diagram is shown in Fig. 4.1. As the diagram has two vertices, the cross-section of the process is proportional to the second order of the coupling constant (fine structure constant), and therefore is of the order of the Thomson cross-section $\sigma_{\rm IC} \sim \sigma_{\rm T}$. We here use a notation $\nu'_{\rm s}$ for the frequency of a seed synchrotron photon, and $\nu'_{\rm IC}$ for the frequency of the photon after undergoing the IC upscattering. The IC emissivity is given by

$$j_{\rm IC} \;=\; \frac{1}{4\pi}\, h\, \eta'_{\rm IC}\, Q_{\rm IC}(\eta'_{\rm IC}) \tag{4.13}$$

where $\eta'_{\rm s/IC} = (h\,\nu'_{\rm s/IC})/(m_{\rm e}\,c^2)$ is the dimensionless energy of a synchrotron or IC photon, and $Q_{\rm IC}(\eta'_{\rm IC})$ is the volumetric rate of production of IC photons per unit of energy interval [$\rm cm^{-3}\ s^{-1}\ eV^{-1}$] described by

$$Q_{\rm IC}(\eta'_{\rm IC}) \;=\; \int d\eta'_{\rm s}\, n(\eta'_{\rm s}) \int d\gamma\, N_{\rm e}(\gamma)\, C(\eta'_{\rm s}\,,\,\gamma\,,\,\eta'_{\rm IC}) \tag{4.14}$$

Here $n(\eta'_{\rm s})$ is the number density of synchrotron photons per unit of energy interval. It is given by

$$n(\eta'_{\rm s}) = \frac{1}{h\,\eta'_{\rm s}}\,\frac{4\pi}{c}\,\tilde{I}'_{\rm s}(\nu'_{\rm s}) \tag{4.15}$$

The synchrotron intensity in fact is not uniform and decreases along the radius of the emitting region (Gould 1979). Because of that, in the calculation of the number density of synchrotron photons, we use $\tilde{I}_{\rm s}(\nu'_{\rm s})$ indicating an "average"/effective

intensity of synchrotron emission over the blob, that undergoes the interaction with relativistic electrons. The effect of non-uniform synchrotron radiation field across the blob can be taken into account in a simple way by scaling the intensity of the central point of the emitting zone by a factor of 3/4 (Kataoka et al. 1999). Then the number density becomes

$$n(\eta'_{\rm s}) = \frac{1}{h\,\eta'_{\rm s}}\,\frac{4\pi}{c}\,\frac{3}{4}\,I'_{\rm s,cen}(\nu'_{\rm s}) \tag{4.16}$$

with $I'_{\rm s,cen}(\nu'_{\rm s}) = \dfrac{j_{\rm s}(\nu'_{\rm s})}{\varkappa_{\rm s}(\nu'_{\rm s})}\,(1\,-\,\exp(-\varkappa_{\rm s}(\nu'_{\rm s})\,R_{\rm b}))$ (Katarzyński et al. 2001).

In order to include a contribution of external photon field, having an intensity $I'_{\rm ext}(\nu'_{\rm ext})$ at the central point of the blob, we simply do a substitution $I'_{\rm s,cen}(\nu'_{\rm s}) \rightarrow I'_{\rm s,cen}(\nu'_{\rm s}) + I'_{\rm ext}(\nu'_{\rm ext})$ in the Eq. 4.16.

The quantity $C(\eta'_{\rm s}\,,\,\gamma\,,\,\eta'_{\rm IC})$ in the Eq. 4.14 is the Compton kernel derived in the framework of quantum electrodynamics (QED) by Jones (1968):

$$C(\eta'_{\rm s}\,,\,\gamma\,,\,\eta'_{\rm IC}) \;=\; \frac{3\,\sigma_{\rm T}\,c}{4\,\gamma^2\,\eta'_{\rm s}}\left[2\chi\ln(\chi)\;+\;(1+2\chi)\,(1-\chi)\;+\;\frac{8\,(\eta'_{\rm s}\,\gamma\,\chi)^2}{1\;+\;4\,\eta'_{\rm s}\,\gamma\,\chi}\,(1\;-\;\chi)\right] \tag{4.17}$$

where $\chi = \dfrac{\eta'_{\rm IC}}{4\,\eta'_{\rm s}\,\gamma\,(\gamma\;-\;\eta'_{\rm IC})}$

This kernel encodes the cross-section of the IC scattering, which takes into account Klein-Nishina (KN) effects. The (nearly) exact cross-section of the IC scattering, calculated with methods of QED is given by (e.g. Coppi & Blandford (1990))

$$\sigma_{\rm IC} = \frac{3\sigma_{\rm T}}{8\chi_1}\left[\left(1-\frac{2}{\chi_1}-\frac{2}{\chi_1^2}\right)\ln(1+2\chi_1)+\frac{1}{2}+\frac{4}{\chi_1}-\frac{1}{2(1+2\chi_1)^2}\right] \tag{4.18}$$

where $\chi_1 = \gamma\,\eta'_{\rm s}$. In the limit of $\chi_1 \ll 1$ (non-relativistic regime) the cross-section tends to the Thomson cross-section $\sigma_{\rm IC} \approx \sigma_{\rm T}$, and in the ultra-relativistic regime $\chi_1 \gg 1$ it decreases as $\sigma_{\rm IC} \approx \frac{3}{8}\,\sigma_{\rm T}\,\frac{\ln(4\,\chi_1)}{\chi_1}$, leading to inefficient upscattering at extremely high energies and a drop in the spectrum.

For a given $\eta'_{\rm s}$ and γ the energy of the outgoing IC photon is not arbitrary, and is bound in an allowed kinematic range

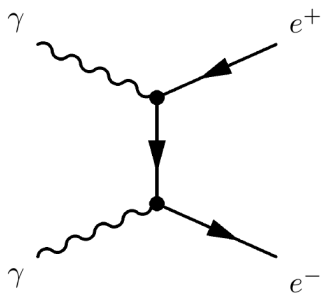


Figure 4.2: Feynman diagram for the γ-γ pair production, a reaction $\gamma + \gamma \rightarrow e^- + e^+$

$$\eta'_{\rm s} \leq \eta'_{\rm IC} \leq \gamma \frac{4\,\eta'_{\rm s}\,\gamma}{1 + 4\,\eta'_{\rm s}\,\gamma} \tag{4.19}$$

The lower bound indicates that we consider only transfer of energy from electrons to photons (inverse Compton effect) and not vice-versa (Compton scattering). The upper bound implies that the photon cannot gain more energy than the electron has (4-momentum conservation law). Thus, the integration in the Eq. 4.14 has to be performed imposing the condition given by the Eq. 4.19.

Neglecting the effect of the internal γ-γ pair production (typically negligible for BL Lac), the intensity of the IC emission $I'_{\rm IC}(\nu_{\rm IC})$ is simply given by an asymptotic solution of the transfer equation in spherical coordinates (analogical to Eq. 4.11)

$$I'_{\rm IC}(\nu'_{\rm IC}) = \frac{4}{3}\, j_{\rm IC}(\nu'_{\rm IC})\, R_{\rm b} \tag{4.20}$$

4.1.5 Gamma-gamma pair production

Two γ-rays colliding with each other can produce an electron-positron pair, i.e. a following reaction occurs: $\gamma + \gamma \rightarrow e^- + e^+$. This reaction is called γ-γ pair production, the corresponding Feynman diagram is depicted in Fig. 4.2. This process is the inverse of the electron-positron annihilation process.

As one can see, the diagram features two vertices, which implies that the cross-section is of the order of the Thomson cross-section $\sigma_{\gamma\gamma} \sim \sigma_{\rm T}$. The (nearly) exact cross-section calculated with the methods of QED is given by (e.g. Aharonian et al. (2008))

$$\sigma_{\gamma\gamma} = \frac{3\sigma_{\rm T}}{2\chi_0^2}\left[\left(\chi_0 + \frac{1}{2}\ln(\chi_0) - \frac{1}{6} + \frac{1}{2\chi_0}\right)\ln(\sqrt{\chi_0} + \sqrt{\chi_0 - 1}) - \left(\chi_0 + \frac{4}{9} - \frac{1}{9\chi_0}\right)\sqrt{1 - \frac{1}{\chi_0}}\right] \tag{4.21}$$

where $\chi_0 = \eta_\gamma\,\eta_{\rm soft}$, η_γ and $\eta_{\rm soft}$ are energies of the high-energy γ-ray and the soft photon in units of the electron rest energy. The threshold of pair production (4-momentum conservation law) is that the energy in the center of mass frame is higher than twice the electron rest energy. This condition is given by

$$\eta_\gamma\,\eta_{\rm soft}\,(1\,-\,\cos\theta)\,\geq\,2 \tag{4.22}$$

for a collision of photons at angle θ.

The maximum value of the pair-production cross-section is $\sigma_{\gamma\gamma,{\rm max}} \approx 0.2\,\sigma_{\rm T}$, achieved at $\chi_0 \approx 3.5$.

This process happens for example in Gamma-Ray Bursts (GRB), where ∼MeV γ-rays collide with each other and produce pairs, posing a problem for γ-ray escape, overcome by relativistic bulk motion of the emitting region. γ-γ pair production can be also important in blazars, especially in FSRQ. A VHE γ-ray colliding with a low-energy photon of external photon field (of BLR, torus or accretion disk), produces an electron-positron pair. This results in the attenuation of the VHE γ-ray flux, characterized by the optical depth $\tau_{\gamma\gamma}$. For BL Lac objects, in most cases, the internal γ-γ absorption is negligible due to insignificance of external radiation fields and due to very low cross-section $\sigma_{\gamma\gamma}$ of the interaction of VHE γ-rays with the local synchrotron radiation (e.g. Katarzyński et al. (2001)).

4.1.6 Transformation to observer's frame

Up to now, all the emission fields (synchrotron and IC) were considered in the blob frame. The source however moves relativistically towards us, and the emission appears to be strongly beamed to the observer. Let us perform transformation of the SSC intensity given in the blob frame, to the observed νF_ν flux (SED) in the observer's frame. First of all, the approaching source boosts the frequency by a factor of $\delta_{\rm b}$, but it suffers from cosmological redshift. Altogether, we apply the following transformation to the photon frequency

$$\nu = \frac{\delta_{\rm b}}{1+z}\,\nu' \tag{4.23}$$

Next, we transform the intensity following the rule $I(\nu) = \delta_{\rm b}^3\, I'(\nu')$. Combining these two effects, we obtain the SED in the observer's frame

$$\nu\, F(\nu) \;=\; \pi \frac{R_{\rm b}^2}{d_{\rm L}^2}\, \delta_{\rm b}^3\, (1+z)\, \nu \, \left[I'_{\rm s}\left(\frac{1+z}{\delta_{\rm b}}\, \nu \right) \,+\, I'_{\rm IC}\left(\frac{1+z}{\delta_{\rm b}}\, \nu \right) \right] \tag{4.24}$$

where $d_{\rm L}$ is the luminosity distance to the source (defined by the redshift z).

4.2 EBL absorption

The VHE γ-rays produced in the blob, after leaving the source, propagate in the intergalactic medium. On the way to the observer, they interact with the soft photons of the Extragalactic Background Light (EBL), which is a ubiquitous ambient radiation field comprising two components: (1) integrated redshifted emission from all stars and galaxies throughout the history of the Universe (optical range), (2) reprocessed stellar emission by dust (IR band). A collision of the VHE γ-ray with an EBL photon leads to production of an electron-positron pair. The (relatively) large cross-section of this process implies that such a process can have high importance, depending on the density of the ambient photons and distance that the high-energy γ-ray travels. In our case, the distances to the observer are cosmological, so that the effect is indeed important. The production of pairs on EBL photons results in disappearance of the initial VHE γ-ray, and leads to attenuation of the VHE γ-ray flux, called EBL absorption effect. This effect starts to be quite important for γ-ray energies (typically) above 100 GeV. Therefore, TeV gamma-rays emitted by a distant blazar cannot propagate over large distances, because of electron-positron pair creation on the optical/infrared photons of the EBL, implying a γ-ray horizon.

The EBL absorption is taken into account by simple exponential suppression of the intrinsic source flux

$$\nu F(\nu)_{\rm obs} = \nu F(\nu)_{\rm intr} \,\cdot\, \exp(-\tau_{_{\rm EBL}}(E,z)) \tag{4.25}$$

where $\nu F(\nu)_{\rm obs}$ is the observed flux, $\nu F(\nu)_{\rm intr}$ is the intrinsic flux from the source, and $\tau_{_{\rm EBL}}(E,z)$ is the optical depth of the EBL absorption which depends on the photon energy and the redshift (distance) to the source. The optical depth is determined by the EBL spectrum and the cross-section of the γ-γ pair production. The latter is well-known from the QED calculations, while the exact shape of the EBL spectrum remains quite uncertain.

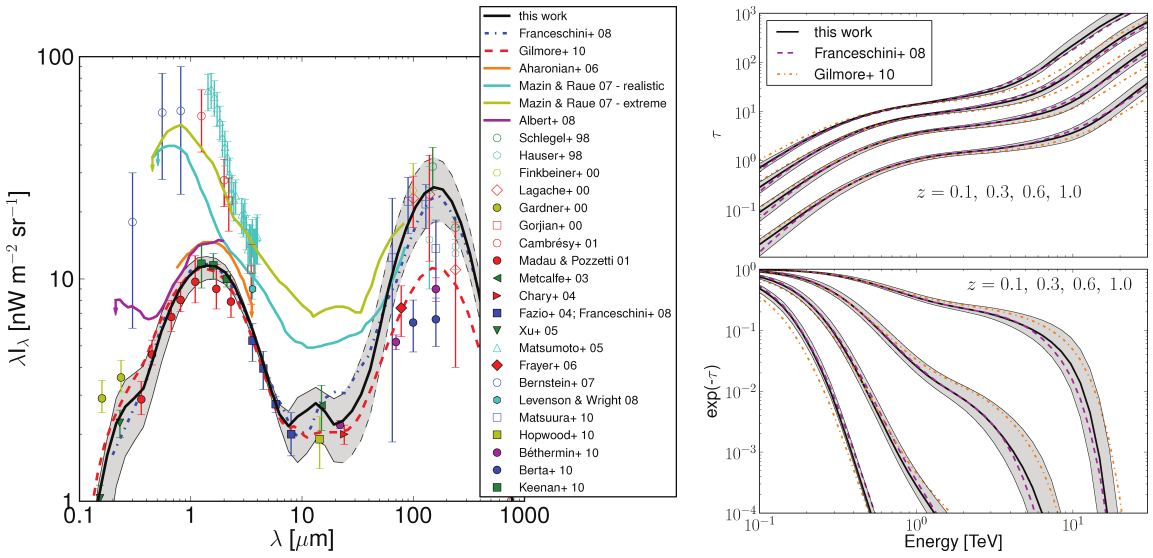


Figure 4.3: Comparison of different EBL models. Black solid line indicates the model by Domínguez et al. (2011). *Left*: Comparison of the EBL spectra deduced using different approaches and data. *Right*: attenuation strength due to the EBL absorption for different EBL models, as a function of the γ-ray energy and of redshift z (top panel – optical depth, bottom panel – flux attenuation factor). (adapted from Domínguez et al. (2011))

There are several methods to estimate it, including the direct measurements from Earth (however they have poor accuracy due to uncertainty in subtraction of the zodiacal light foreground), model description using cosmological models and stellar evolution models (have an uncertainty due to poor knowledge of the initial rate of star formation) and finally, reconstructed from the physical modeling of the blazar spectra. There are currently several "standard" EBL models (e.g. Franceschini et al. (2008) ; Domínguez et al. (2011)), deduced from different approaches, and which provide tabulated values of the optical depth in a 2D matrix of photon energies and redshifts. A comparison of EBL spectra obtained using different methods is illustrated in the left panel of Fig. 4.3. The flux attenuation effect depending on the energy of the γ-ray and the source's redshift for different EBL models is shown in the right panel of Fig. 4.3.

The detection of 3C 279 as a VHE γ-ray emitter, discussed before, besides having implications on blazar radiative models, also has an important impact on EBL models, discussed extensively by MAGIC Collaboration et al. (2008). According to the EBL model by Franceschini et al. (2008), the optical depth $\tau_{\gamma\gamma} = 1$ of the EBL absorption for the source's redshift of $z = 0.536$, is achieved at ~ 200 GeV. However, 3C 279 was detected up to 500 GeV (MAGIC Collaboration et al. 2008), with the optical depth for

γ-rays of this energy of $\tau_{\gamma\gamma} \approx 5$. Also, the results obtained in Chapter 3 (preliminary) show that the H.E.S.S. spectrum of 3C 279 during the strong flare in June 2018 extended up to a few TeV (see top panel of Fig. 3.14). These results indicate that the transparency of the Universe to γ-rays might be higher than previously thought, and we conclude that our understanding of the EBL is very incomplete. Further studies are required to advance in this direction, in particular, future IACT system Cherenkov Telescope Array (CTA) is expected to provide quite tight constraints on the EBL (see sub-section 7.1.2).

Chapter 5

Modeling of AGN emission: time-dependent approach

The framework discussed in the previous chapter, considers only an instantaneous emission from an ensemble of electrons, with an electron spectrum parametrized in a relatively simple, *ad-hoc* way. The chosen shapes are usually motivated by presumed physical processes that are thought to shape the particle distribution, but still, the values of the parameters describing the spectrum profile (e.g. a change of the spectral index after a break) in this case does not appear as a consequence of the considered underlying physics. This treatment is usually satisfactory for the modeling of stationary or quasi-stationary states of blazars (e.g. quiescent emission or long high states / slow flux evolution over time-scales $t_{\rm var} \gg R_{\rm b}/c$) if one has a goal to understand the emission origin. However this approach cannot be applied to (relatively) fast flares, occurring on time-scales $t_{\rm var} \sim R_{\rm b}/c$. A *time-dependent* framework has to be used to describe the variable emission during such flares. A naive and simple way to deal with it would be to either vary parameters of the particle distribution (e.g. a cutoff energy and/or spectral slope), or to vary the physical parameters of the emitting blob (e.g. radius and/or magnetic field), or both at the same time. We focus on moderate perturbations in the source, during which the blob parameters remain constant in time, and consider that the flux variations are caused only by electron spectrum modification. Then one might fit the observed broad-band SED at every moment of time (assuming such information is available) with an instantaneous model and retrieve the electron spectrum at every instance of time, including its parameters. However such information can not tell us a lot about underlying physical processes driving the change of the parameters, at best one could only aspire to interpret the evolution of the parameters only qualitatively by invoking basic physical arguments. Such interpretation may be ambiguous and may lead to inaccurate or even completely

wrong conclusions. A more self-consistent approach to model the varying emission during blazar outbursts, is to describe in a coherent manner how the spectrum of the particle population influenced by different physical processes evolves in time. In this framework, the electron spectrum evolution happens naturally as the consequence of basic physical processes acting in the blob. Such approach allows to explore not only the emission mechanism, but also, very importantly, the causes of the observed variability, in particular, establish which physical process(es) initiate the flux rise and its fall.

In this chapter[1], we first present the general time-dependent approach and discuss in detail various physical processes thought to be responsible for the blazar variability (Section 5.1). Next, in Section 5.2 we present the numerical code I developed based on the time-dependent framework, and finally, in Section 5.3 we discuss various physical scenarios proposed to explain the blazar flaring phenomenon.

5.1 General approach and the kinetic equation

We assume the blob-in-jet model for the VHE γ-ray emitting zone (see sub-section 4.1.2). The blob is considered to be homogeneous, has a radius $R_{\rm b}$, magnetic field of uniform strength B, and a Doppler factor $\delta_{\rm b}$. The particle population in the blazar emitting zone is evolving because of several physical processes. Electrons are injected into the plasma blob with a spectrum $Q_{\rm inj}$ which may be time-dependent, and may gain energy through acceleration by shock (Fermi-I) or stochastic (Fermi-II) mechanisms. The electrons confined in the blob radiate synchrotron and inverse Compton emission (leptonic SSC scenario, see sub-section 4.1.1), and cool via radiative losses, which include synchrotron and inverse Compton cooling losses. The electrons also are able to escape the emission region at a characteristic time-scale $t_{\rm esc}$. We ascribe the flaring activity to the variations of the spectrum of the particle population in the emitting blob. We neglect adiabatic losses and internal γ-γ absorption.

The time evolution of the electron spectrum $N_{\rm e}(\gamma, t)$ in the emitting zone is governed by a kinetic equation representing a continuity equation in phase space, also called Fokker-Planck equation. Its general form, taking into account the above

[1]Partial results of the presented work have been (1) submitted for a publication "Connecting steady emission and the 2010 Very High Energy flaring state in the blazar Mrk 421" by A. Dmytriiev, H. Sol and A. Zech to MNRAS (September 2020), and (2) already published in the proceedings of the 36th International Cosmic Ray Conference (ICRC2019) in a contribution "Time Dependent Modeling of Electron Acceleration and Cooling During Blazar Flares" by A. Dmytriiev, H. Sol and A. Zech (July 2019) (Dmytriiev et al. 2019a)

mentioned physical processes in the blob, is (e.g. Kardashev (1962) ; Tramacere et al. (2011)):

$$\frac{\partial N_{\rm e}(\gamma,t)}{\partial t} = \frac{\partial}{\partial \gamma}\left[\left(b_{\rm c}\gamma^2 - a\gamma - \frac{2}{\gamma}\,D_{_{p,\rm FII}}(\gamma,t)\right)\cdot N_{\rm e}(\gamma,t)\right] + \\ + \frac{\partial}{\partial \gamma}\left(D_{_{p,\rm FII}}(\gamma,t)\,\frac{\partial N_{\rm e}(\gamma,t)}{\partial \gamma}\right) - \frac{N_{\rm e}(\gamma,t)}{t_{\rm esc}(\gamma)} + Q_{\rm inj}(\gamma,t) \quad (5.1)$$

with $D_{_{p,\rm FII}}(\gamma,t) = D_0(t)\,\gamma^q$. For the case of Fermi-II acceleration in "hard-sphere" approximation ($q = 2$) (see sub-section 5.1.3) that we adopt in our modeling (this choice is justified in sub-section 6.5.2), the kinetic equation becomes

$$\frac{\partial N_{\rm e}(\gamma,t)}{\partial t} = \frac{\partial}{\partial \gamma}\left[\left(b_{\rm c}\gamma^2 - a\gamma - 2\gamma D_0(t)\right)\cdot N_{\rm e}(\gamma,t)\right] + \\ + \frac{\partial}{\partial \gamma}\left(D_0(t)\,\gamma^2\,\frac{\partial N_{\rm e}(\gamma,t)}{\partial \gamma}\right) - \frac{N_{\rm e}(\gamma,t)}{t_{\rm esc}} + Q_{\rm inj}(\gamma,t) \quad (5.2)$$

The physical processes behind the different terms, as well as physical quantities appearing in the kinetic equation, are described in the following sub-sections.

5.1.1 Particle injection

Considering the "blob-in-jet" model, we assumed that the blob is filled with relativistic electrons. A natural question arises: "How do these electrons appear in the emitting zone at the first place?".

It is widely accepted that the particles injected into blazar emitting zone are already pre-accelerated, however a direct continuous acceleration of injected cold electrons inside the blob is also considered (e.g. Cao & Wang (2013)). We will focus here on the former view. A source providing the flux of pre-accelerated particles can originate from the central engine (injection at the base of the jet), from within the jet, or in the direct vicinity of the emitting blob. Multiple possible options for the source supplying fresh particles into blazar emitting zone are discussed: (1) a shock leading the blob (Kirk et al. 1998), (2) internal shocks formed due to collision of multiple shells within the jet (Spada et al. 2001), (3) isolated shocks traveling down the jet (Marscher & Gear 1985), (4) shear layer of the spine-sheath jet structure accelerating particles by the Fermi-II mechanism (see 5.1.3) (Ostrowski & Bednarz

2002), (5) magnetic reconnection events (see 5.1.3) (Sikora et al. 2005), (6) central engine (Dermer & Schlickeiser 1993), or (7) hadronically initiated pair avalanches (Kazanas & Mastichiadis 1999). Typically, the injection scenarios involving shocks in the jet or upstream the blob, are regarded as the most conventional.

The particle injection is described in the kinetic equation with the term $Q_{\rm inj}(\gamma, t)$, which is the number of particles injected in a unit volume per unit time and per unit of Lorentz factor interval. The injection spectrum can be time-dependent, and particles may be injected continuously over a given time interval, or in an impulsive manner.

5.1.2 Particle radiative cooling

Particles in the emitting zone radiate, and lose energy, i.e. "cool". The term $b_{\rm c}\gamma^2$ corresponds to the total cooling rate, comprising the synchrotron and the IC cooling rate:

$$-b_{\rm c}\gamma^2 = \dot{\gamma}_{\rm syn} + \dot{\gamma}_{\rm IC} \tag{5.3}$$

The synchrotron cooling is due to emission of photons of synchrotron radiation by the electrons, and its rate is given by (e.g. Chiaberge & Ghisellini (1999)):

$$-\dot{\gamma}_{\rm syn} = \frac{4\sigma_{\rm T}}{3m_{\rm e}c}\cdot\gamma^2\cdot U_B \tag{5.4}$$

where $U_B = \frac{B^2}{2\mu_0}$ is the magnetic energy density.

The inverse Compton cooling is the energy loss due to inverse Compton upscattering of low energy synchrotron photons by the high energy electrons. Its rate is given by (Moderski et al. 2005):

$$-\dot{\gamma}_{\rm IC} = \frac{4\sigma_{\rm T}}{3m_{\rm e}c}\cdot\gamma^2\cdot\int_{\epsilon'_{\rm min}}^{\epsilon'_{\rm max}} f_{\rm KN}(4\gamma\epsilon')\cdot u'_{\rm syn}(\epsilon')\,d\epsilon' \tag{5.5}$$

where $f_{\rm KN}(x)$ includes the full Klein-Nishina inverse Compton cross-section, and is approximated as:

$$f_{\rm KN}(x) \simeq \begin{cases} (1+x)^{-1.5}, & \text{for } x < 10^4 \\ \frac{9}{2x^2}\cdot\left(\ln(x) - \frac{11}{6}\right), & \text{for } x \geq 10^4 \end{cases} \tag{5.6}$$

The term $b_{\rm c}$ in the total cooling rate is thus:

$$b_{\rm c} = \frac{4\sigma_{\rm T}}{3m_{\rm e}c} \cdot \left[U_B + \int_{\epsilon'_{\rm min}}^{\epsilon'_{\rm max}} f_{\rm KN}(4\gamma\epsilon') \cdot u'_{\rm syn}(\epsilon')\, d\epsilon' \right] \tag{5.7}$$

The characteristic cooling time of electrons is:

$$t_{\rm cool}(\gamma) = \frac{\gamma}{b_{\rm c}\gamma^2} = \frac{1}{b_{\rm c}\gamma} \tag{5.8}$$

The inverse Compton cooling starts to be important (in comparison to the synchrotron cooling) if the IC bump in the SED has an intensity similar (or higher) to the one of the synchrotron bump. In BL Lac objects, the inverse Compton cooling is usually negligible in the low state, however it might become important during bright VHE γ-ray flares. For FSRQs, the IC bump is (much) higher than the synchrotron one, so that the inverse Compton cooling represents the dominant cooling process. An important remark has to be made for such case. The kinetic equation in a form as presented by Eq. 5.1 is derived assuming that particles change their energy by a only a small fraction in each interaction, $\Delta E/E \ll 1$. However, electrons undergoing the inverse Compton cooling in Klein-Nishina regime, lose an important fraction of their energy in one electron-photon collision, which means that the Eq. 5.1 is formally not applicable to describe the varying emission of blazars in which the IC cooling dominates over the synchrotron one. A more complex form of the kinetic equation valid for an arbitrary fraction $\Delta E/E$ has to be used in this situation. However, for Mrk 421, this effect should be negligible due to the fact that the IC cooling is subdominant in BL Lac objects, so that the standard kinetic equation can be applied to model the emission of this source.

5.1.3 Particle acceleration

As it was discussed in Chapter 2, various observations indicate the presence of high-energy particles in AGN jets. However, up to now, we did not consider mechanisms that boost particles to very high energies.

Particle acceleration processes in AGN jets leave notable observational signatures. For example, one could notice that in certain energy ranges, the blazar spectra (see e.g. Fig. 2.9) are consistent with a power law (e.g. in radio, optical-to-UV, low-energy γ-rays, depending on the object). This requires a power law shape of the distribution of the underlying particle population, in a wide range of particle energies, which is most naturally explained by acceleration of particles by a certain mechanism(s). Another important observation is that during high states, some blazars

show a clear "harder when brighter" trend, meaning that the photon (and particle) spectrum becomes harder with an increasing flux. One of the possible causes of this phenomenon, is (re-)acceleration of the particle population by the intervening acceleration processes. In this sub-section, we will discuss different acceleration mechanisms that might operate in blazars.

Second-order Fermi acceleration mechanism (Fermi-II)

The Fermi-II mechanism was first proposed by Enrico Fermi-In 1949 (Fermi 1949). He developed this theory in the context of the origin of cosmic rays, considering acceleration of particles due to scattering off the clouds in the interstellar medium of our Galaxy. In Fermi's original view, charged particles are interacting with randomly moving massive "magnetic mirrors", representing the irregularities in the magnetic field of the Galaxy. The collisions of particles with the clouds are assumed to be elastic, so that particles are reflected after hitting the mirror. Fermi demonstrated that particles can stochastically gain energy in these interactions. Here we repeat Fermi's description of this process.

Let us consider an elastic collision of a relativistic particle having mass m and speed v with a massive non-relativistic mirror/cloud having mass M and moving with a velocity U (see Fig. 5.1). The change in the particle energy equals to the change in the cloud energy, due to energy conservation:

$$\Delta E = (\gamma' - \gamma)mc^2 = \frac{1}{2}M(U^2 - U'^2) \tag{5.9}$$

where γ is the initial Lorentz factor of the particle. Unprimed and primed quantities are related to the state before and after the collision, respectively.

The 3-momentum is also conserved in the interaction:

$$\gamma m\mathbf{v} + M\mathbf{U} = \gamma' m\mathbf{v}' + M\mathbf{U}' \tag{5.10}$$

Taking the square of the Eq. 5.10 and re-arranging the terms, we find:

$$M^2(U^2 - U'^2) = m^2(\gamma'^2 v'^2 - \gamma^2 v^2) + 2mM(\gamma' v' U' \cos\theta' - \gamma v U \cos\theta) \tag{5.11}$$

where θ and θ' are the angles between the particle and cloud velocities before and after the collision, respectively.

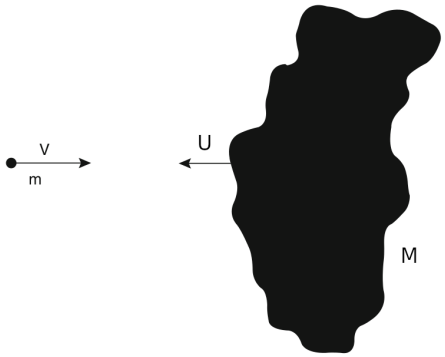


Figure 5.1: Geometry of the collision between a particle and a massive cloud / magnetic mirror. (adapted from Courvoisier (2013))

We consider that the change of the particle energy is small with respect to its initial energy, $\frac{\gamma' - \gamma}{\gamma} \ll 1$. Then, we can approximate the first term in the RHS as (using also $\gamma = 1/\sqrt{1 - v^2/c^2}$): $m^2(\gamma'^2 v'^2 - \gamma^2 v^2) \approx m^2 c^2 \cdot 2\gamma \cdot (\gamma' - \gamma)$. Also, since the cloud is very massive, its velocity change will be almost negligible in the interaction, $U \simeq U'$. That allows us to simplify the last term in the RHS: $2mM(\gamma' v' U' \cos\theta' - \gamma v U \cos\theta) \approx 2mM\gamma v U(\cos\theta' - \cos\theta)$. Substituting the expression for $U^2 - U'^2$ from the Eq. 5.9 to the LHS of the Eq. 5.11, we obtain

$$2mMc^2(\gamma' - \gamma) = m^2 c^2 \cdot 2\gamma \cdot (\gamma' - \gamma) + 2mM\gamma v U(\cos\theta' - \cos\theta) \tag{5.12}$$

Let us consider head-on and rear collisions between the mirror and the particle. In the former case, $\cos\theta = -1$ and $\cos\theta' = 1$, and in the latter case, the inverse, $\cos\theta = 1$ and $\cos\theta' = -1$. We can therefore rewrite the previous equation for both cases at once as follows:

$$2mMc^2(\gamma' - \gamma) = 2m^2 c^2 \gamma(\gamma' - \gamma) \pm 4mM\gamma v U \tag{5.13}$$

where the "+" sign before the last term in the RHS corresponds to the head-on collisions, and the "-" sign refers to the rear ones.

From here, we can express the relative energy change, $\frac{\Delta E}{E} = \frac{\Delta\gamma}{\gamma}$, recalling that the cloud is much more massive than the particle, $m \ll M$:

$$\frac{\Delta E}{E} \simeq \pm 2\,\frac{vU}{c^2} \tag{5.14}$$

After a head-on collision, the particle will gain energy, while after a rear one, the particle will lose energy. A particle traveling through a medium filled with a large number of magnetized clouds, will experience the both types of collisions. The rate of collisions is proportional to the relative velocity of the magnetic mirror and the particle, $R = \sigma n v_{\rm rel}$, where σ is the collision cross-section, and n is the number density of the clouds. For head-on collisions, $v_{\rm rel} = v + U$, and for rear ones, $v_{\rm rel} = v - U$. Therefore, the head-on collisions are more frequent, and the particle, on average, gains energy. The time-average net energy gain rate is:

$$\frac{dE}{dt} = \sigma n (v+U)\cdot E \cdot 2\,\frac{vU}{c^2} - \sigma n (v-U)\cdot E \cdot 2\,\frac{vU}{c^2} \simeq 4\sigma n v E\,\frac{U^2}{c^2} \propto \left(\frac{U}{c}\right)^2 \quad (5.15)$$

As one could see, it is proportional to the second order of the magnetic cloud velocity U. Because of this, such acceleration process is referred to as "second-order Fermi acceleration". The acceleration process is also called "stochastic", since particles experience diffusion in the medium with gaining and losing energy in scatterings, and, gradually gain energy in a stochastic manner. The associated time-scale of the acceleration process, is:

$$t_{\rm FII} = \frac{E}{\dot{E}} = \frac{c^2}{4\sigma n v U^2} = \frac{\lambda_{\rm f}\, c^2}{4\, v\, U^2} \quad (5.16)$$

where $\lambda_{\rm f} = \frac{1}{\sigma n}$ is the mean free path of the particle in the scattering process. The longer the particle scatters off the clouds, the more energy it gains. In real astrophysical conditions, particles will not be accelerated infinitely long, because the region with clouds has limited spatial extension (for the interstellar medium it is the size of our Galaxy). Sooner or later, the particle will escape the acceleration region and therefore the duration of the acceleration episode for the particle will be the escape time-scale.

In this simple picture, we can describe the acceleration process as a competition between the increase of the particle energy, and a decrease of number of particles. Solving the differential equation Eq. 5.16, we obtain that particle energy grows as $E(t) \propto \exp(t/t_{\rm FII})$. The time needed for the particle with the initial energy E_0 to attain energy E will be the inverse of this expression:

$$t = t_{\rm FII} \ln(E/E_0) \quad (5.17)$$

The number of particles in the system declines exponentially, due to the loss term expressed in the Eq. 5.2 as $\frac{dN}{dt} \sim \frac{N}{t_{\rm esc}}$. Solving this equation we get

$$N(t) = N_0 \exp(-t/t_{\rm esc}) \tag{5.18}$$

Plugging the Eq. 5.17 into the Eq. 5.18, we obtain that the number of particles with energy higher than E_0, will be $N(E > E_0) \propto E^{-t_{\rm FII}/t_{\rm esc}}$, which corresponds to a differential particle spectrum

$$\frac{dN}{dE} \propto E^{-(1+t_{\rm FII}/t_{\rm esc})} \tag{5.19}$$

In case the ratio $t_{\rm FII}/t_{\rm esc}$ is energy-independent, the particle spectrum is a power law.

The Fermi-II mechanism is expected to take place in many different astrophysical settings. Most of the astrophysical environments are magnetized, and many of them also feature turbulence. Chaotic motions of highly conducting magnetized plasma at different spatial scales will generate a random component of the magnetic field, due to the fact that the magnetic field lines are moving together with the plasma (frozen-in condition). This phenomenon is called magnetohydrodynamic (MHD) turbulence. Inhomogeneties in the magnetic field structure can be treated as the magnetic clouds, described above. Therefore, the conditions for particle acceleration via Fermi-II mechanism are most likely present in a large number of astrophysical sources.

However, despite its seeming ubiquity, this acceleration mechanism has very low efficiency in many astrophysical environments. Considering the example of interstellar medium, known to be turbulent and magnetized, the time-scale of the stochastic acceleration would be $t_{\rm FII} \sim 10^{19}$ s, which exceeds the age of the Universe. Therefore the Galactic cosmic ray spectrum cannot be produced by means of interstellar turbulence (e.g. Ginzburg & Syrovatskii (1964)). Overall, if the turbulent motions are non-relativistic, the Fermi-II process is expected to be very slow, as it is second order in $\frac{U}{c}$, as is obvious from the Eq. 5.15. Another important remark about the stochastic acceleration is that it does not generate power law particle spectra with a specific "universal" spectral index. Many observations of different cosmic sources (as well as the cosmic ray spectrum) find very similar indices of particle spectra in the range 2 – 2.7. From the Eq. 5.19 one could see that this would require the ratio $t_{\rm FII}/t_{\rm esc}$ to be nearly the same in very diverse astrophysical conditions, which is very unlikely. Thus, we conclude that Fermi-II process cannot be a "universal" ubiquitous mechanism for accelerating particles, and is rather specific to a limited range of cosmic environments.

In the modern view of Fermi-II acceleration, the particles are scattered on various types of plasma waves and are stochastically accelerated. The waves in plasma can

be excited due to MHD turbulence. This phenomenon involves a transfer of energy over a range of spatial scales: the turbulent energy is injected at the largest "stirring scale", and then the energy cascades down to the shortest "damping scale" where the viscosity forces are able to dissipate kinetic and magnetic energy of the fluid.

As previously discussed, the turbulent motions in the magnetized plasma produce a stochastic component of the magnetic field δB in addition to the main (ordered) component B_0. The chaotic fluctuations of the magnetic field at different spatial scales perturb the plasma and excite different kinds of plasma waves, in particular, Alfvén waves. In the quasi-linear framework (e.g. Schlickeiser (1989); Jaekel & Schlickeiser (1992)), the magnetic field fluctuations induced by the turbulence are represented by varying magnetic fields of Alfvén waves, and the MHD turbulence is described by a combination of Alfvén waves with different wavenumbers, forming a continuous wave spectrum. These wavenumbers correspond to spatial scales of turbulent motions. Such a spectrum appears due to cascading, i.e. transfer of energy to smaller spatial scales. The cascading in this description is caused by non-linearity of the MHD equations, and the cascade is developing as follows: first, two waves interact at the largest scale, resulting in a mode with a smaller wavelength (due to non-linear effects), and this process is repeated down to the minimal spatial scale (Achterberg 1979).

The one-dimensional spectrum of the energy density of the MHD turbulence has the following form (e.g. Becker et al. (2006))

$$W(k) \propto k^{-q} \tag{5.20}$$

where $k = 2\pi/\lambda$ is the wavenumber. The power spectrum is normalized as follows

$$\int_{k_{\min}}^{k_{\max}} W(k)\, dk = \frac{\delta B^2}{2\mu_0} \tag{5.21}$$

which is the total energy density stored in the magnetic fluctuations. The minimal and the maximal wavenumbers $k_{\min}$ and $k_{\max}$ correspond to the longest $\lambda_{\max}$ and the shortest wavelength $\lambda_{\min}$ in the Alfvén spectrum accordingly: $k_{\min} = 2\pi/\lambda_{\max}$, and $k_{\max} = 2\pi/\lambda_{\min}$. Taking into account the above-mentioned normalizing condition (Eq. 5.21), the turbulence spectrum is thus

$$W(k) = \frac{\delta B^2}{2\mu_0} \frac{q-1}{k_{\min}} \left(\frac{k}{k_{\min}}\right)^{-q} \tag{5.22}$$

The spectral index $q = 5/3$ for the Kolmogorov turbulence, $q = 3/2$ for the

Kraichnan turbulence, and $q = 2$ for the "hard-sphere" turbulence (see e.g. Zhou & Matthaeus (1990)).

Particles of the plasma interact with the Alfvén waves and may exchange energy and momentum (for a complete description, see e.g. Dermer et al. (1996)). Let us consider this process in more detail. In particular, we are interested in establishing the form of the diffusion coefficient $D_{p,\mathrm{FII}}$, appearing in the kinetic equation (Eq. 5.1). The diffusion coefficient describes momentum/energy gain by a particle during the process of stochastic acceleration. An Alfvén wave is a magnetic field fluctuation propagating in the direction of the (ordered large-scale) magnetic field, and so the vector of the magnetic field fluctuation is always perpendicular to the main component: $\delta\mathbf{B} \perp \mathbf{B_0}$. We focus on a case where $\delta B \ll B_0$. Let us consider motion of a charged particle in the reference frame of the Alfvén wave, where the electric field associated with the wave can be neglected:

$$\frac{d\mathbf{p}}{dt} = e \cdot \mathbf{v} \times (\mathbf{B_0} + \delta\mathbf{B}) \tag{5.23}$$

here $\mathbf{p}$ and $\mathbf{v}$ are vectors of the momentum and the velocity of the particle respectively.

Splitting the momentum into parallel and perpendicular components, $\mathbf{p}_{||}$ and $\mathbf{p}_{\perp}$, we find that the perpendicular component will be governed by the terms $e \cdot \mathbf{v}_{\perp} \times \mathbf{B_0}$ and $e \cdot \mathbf{v}_{||} \times \delta\mathbf{B}$. The first term is the Lorentz force causing the particle to gyrate around the lines of large-scale ordered magnetic field of the plasma B_0 (unperturbed motion), and the second term can be neglected (with respect to the first one), therefore $\mathbf{p}_{\perp}$ does not change its modulus in time. For the parallel component $\mathbf{p}_{||}$, the equation of motion is

$$\frac{d\mathbf{p}_{||}}{dt} = e \cdot \mathbf{v}_{\perp} \times \delta\mathbf{B} \tag{5.24}$$

The modulus of the parallel component of the momentum can be expressed as $p_{||} = |\mathbf{p}| \cos\theta = p \cos\theta$, where the angle θ is a pitch-angle and p is the absolute value of the particle momentum vector, or simply, particle momentum. We will use a notation $\cos\theta = \mu$. Using the relation $p_{||} = p\mu$, we get the differential equation for time evolution of pitch angle:

$$\frac{d\mu}{dt} = \frac{ev}{p}\sqrt{1-\mu^2}\,\delta B \cos(\Omega t - kx + \psi) \tag{5.25}$$

here $\Omega = \frac{eB_0}{\gamma m}$ is the Larmor frequency, $k = \frac{2\pi}{\lambda}$ is the wavenumber of the Alfvén

wave, ψ is the phase. In the reference frame co-moving with the wave, $x = v_{||}t = v\mu t$, and so we obtain

$$\frac{d\mu}{dt} = \frac{ev}{p}\sqrt{1-\mu^2}\,\delta B\cos((\Omega - kv\mu)t + \psi) \tag{5.26}$$

Obviously, the time-averaged variation of the pitch angle will be zero, however the time-averaged square of this quantity is non-zero. Averaging over the random phase of the waves and integrating over time we get:

$$\langle \Delta\mu \cdot \Delta\mu \rangle = \frac{e^2v^2(1-\mu^2)\delta B^2}{2p^2}\int dt' \int dt'' \cos((\Omega - kv\mu)(t'-t'')) = \\ = \frac{e^2v\,(1-\mu^2)\,\delta B^2}{\mu p^2}\,\Delta t\,\cdot\,\delta\left(k - \frac{\Omega}{v\mu}\right) \tag{5.27}$$

where $\delta(x)$ is the Dirac delta function.

From this, one could see that the most efficient interaction occurs when the Larmor radius of a particle, $r_{\text{L}} = \frac{v}{\Omega} = \frac{p}{eB_0} = \frac{\gamma mv}{eB_0}$, "resonates" with a specific Alfvén wavelength, representing a spatial scale of magnetic turbulence: $r_{\text{L}} = \frac{1}{k\mu}$. An additional condition for the resonance to occur is that the Alfvén wave has to be polarized in the sense of the gyrating motion of the particle. If the resonance condition is met, the particle changes its pitch angle in a random direction by $\sim \delta B/B_0$, otherwise, the particle will barely "feel" the wave.

In the reference frame of the wave the resonant scattering resembles to an elastic collision with a "cloud" or magnetic mirror, discussed in the original Fermi version of the theory. The cause of the energy gain or loss in this view is a small-scale electric field, induced by the moving magnetic field. This electric field, appearing in the laboratory frame, either accelerates or decelerates the particle, depending on the orientation of the electric field with respect to the velocity vector of the particle. Similarly to bouncing off the massive clouds, the particle will gain energy in head-on reflections, and lose it in rear ones. Since the head-on collisions are, again, more frequent, than the rear ones, the particle on average gains energy after a large number of scatterings on the waves. The work on acceleration of particles is done by small-scale electric fields induced by moving magnetic field fluctuations. Overall, during the Fermi-II acceleration phase, the particles extract energy from the turbulent motions of the plasma.

Particles interact not with only one wave, but rather with the whole Alfvén wave spectrum. One can repeat the derivation of Eq. 5.27 replacing the δB^2 with

its spectral energy density (per unit of wavenumbers): $\delta B_k(k)^2$. Such an equation describes the diffusion of particles due to the resonant interactions, and is related to the diffusion coefficient $D_{p,\mathrm{FII}}$, which can be derived from the Eq. 5.27 assuming isotropy. We present here the final expression for the diffusion coefficient. It is controlled by the wave-turbulence power spectrum, and is given by (Schlickeiser 1989):

$$D_{p,\mathrm{FII}}(p) \approx \beta_{\mathrm{A}}^2 \frac{\delta B^2}{B_0^2} \left(\frac{r_{\mathrm{L}}}{\lambda_{\max}}\right)^{q-1} \frac{cp^2}{r_{\mathrm{L}}} \tag{5.28}$$

where β_{A} is the Alfvén speed in the units of the speed of light, and $r_{\mathrm{L}} = \frac{p}{eB_0}$ is the Larmor radius of a particle. The quantity $\frac{\delta B^2}{B_0^2}$ is commonly referred to as "turbulence level". One notices that the diffusion coefficient $D_{p,\mathrm{FII}}(p) \propto p^q$. The characteristic time-scale of Fermi-II acceleration process due to particle-wave interactions is (O'Sullivan et al. 2009):

$$t_{\mathrm{FII}} = \frac{p^2}{D_{p,\mathrm{FII}}(p)} = \frac{1}{\beta_{\mathrm{A}}^2} \frac{B_0^2}{\delta B^2} \frac{\lambda_{\max}}{c} \left(\frac{r_{\mathrm{L}}}{\lambda_{\max}}\right)^{2-q} \tag{5.29}$$

It scales with the particle momentum/energy as $t_{\mathrm{FII}} \propto p^{2-q}$. It is interesting to note some similarities with the Eq. 5.16: the velocity of the scattering centers U is now the Alfvén speed, $U = v_{\mathrm{A}} = \beta_{\mathrm{A}} c$, as it is the speed of the magnetic field fluctuations, and, due to the frozen-in condition, also the velocity of the turbulent motions. The mean free path length λ_{f} is now

$$\lambda_{\mathrm{f}} = 4 \frac{B_0^2}{\delta B^2} \cdot \lambda_{\max} (r_{\mathrm{L}}/\lambda_{\max})^{2-q} \tag{5.30}$$

Assuming strong turbulence with $\delta B^2 \sim B_0^2$, one can see that the distance scale between two scattering events is the longest in the "hard-sphere" case, $\lambda_{\mathrm{f}} \sim \lambda_{\max}$ (and is energy-independent), and is the shortest for Kraichnan turbulence, $\lambda_{\mathrm{f}} \sim \sqrt{r_{\mathrm{L}} \cdot \lambda_{\max}}$. Also, the mean free path increases with particle energy in the case of Kraichnan and Kolmogorov turbulence. One also notices that the overall Fermi-II time-scale (Eq. 5.29) is also scaled by the inverse turbulence level compared to the Eq. 5.16, so that the Fermi-II time-scale is shorter for a higher turbulence level. Indeed, the higher is the energy density of the turbulence, the more energy will be transferred to particles, and the more efficient will be the acceleration process.

As one can notice from the Eq. 5.29, the Fermi-II process becomes faster with increasing Alfvén speed, meaning that particle acceleration in (mildly) relativistic MHD turbulence could be quite efficient. The quasi-linear approach of Schlickeiser

(1989) provides rather accurate (with an order of magnitude precision compared to numerical test-particle simulations) description of stochastic particle acceleration in the case of non-relativistic Alfvén speeds $\beta_{\rm A} \ll 1$ and low turbulence levels $\delta B \ll B_0$, and for mildly relativistic Alfvén waves and both small and comparable to unity turbulence levels (O'Sullivan et al. 2009).

In our modeling, presented in Chapter 6, we adopt "hard-sphere" turbulence ($q = 2$), widely assumed in the literature (the motivation of this choice is discussed in sub-section 6.5.2). From the Eq. 5.29 one can see that in the case of "hard-sphere" turbulence the stochastic acceleration time-scale $t_{\rm FII}$ is energy-independent, while for Kolmogorov type turbulence $t_{\rm FII} \propto p^{1/3}$ and for Kraichnan spectrum $t_{\rm FII} \propto p^{1/2}$. This means that with the "hard-sphere" scattering, the shortest Fermi-II acceleration time-scales are achieved at high Lorentz factors, and hence the most efficient particle acceleration. Qualitatively, in the "hard-sphere" case, more energy of magnetic turbulence is concentrated at longer wavelengths that resonate with higher energy particles, which leads to their more efficient acceleration and yields harder particle distributions (see some examples in e.g. Becker et al. (2006)). Thus, if one supposes stochastic particle acceleration as a mechanism powering blazar flares, the "hard-sphere" turbulence is much better at producing strong VHE γ-ray and hard X-ray flares, than the Kolmogorov and Kraichnan turbulence types. The "hard-sphere" approximation is rather frequently used for modeling of Fermi-II acceleration in blazar jets (e.g. Asano & Hayashida (2018)).

For the case of "hard-sphere" scattering, the momentum diffusion coefficient is (Eq. 5.28 with $q = 2$):

$$D_{p,\rm FII}(p) \approx \beta_{\rm A}^2 \cdot \left(\frac{\delta B}{B_0}\right)^2 \cdot \left(\frac{\lambda_{\rm max}}{c}\right)^{-1} \cdot p^2 \equiv D_0\, p^2 \tag{5.31}$$

and the Fermi-II acceleration time-scale in this case is:

$$t_{\rm FII} = \frac{1}{\beta_{\rm A}^2} \cdot \left(\frac{B_0}{\delta B}\right)^2 \cdot \frac{\lambda_{\rm max}}{c} = \frac{1}{D_0} \tag{5.32}$$

The Alfvén speed depends on the physical parameters of the medium: energy density of the electrons and the magnetic field. For the case of relativistic MHD, the Alfvén speed in the units of the speed of light is given by (Gedalin 1993):

$$\beta_{\rm A} = \frac{1}{\sqrt{1 + \frac{\varepsilon + P}{B^2/\mu_0}}} \tag{5.33}$$

where ε is the energy density of plasma particles, and P is the pressure.

The energy density can be evaluated directly from the electron spectrum, assuming that the relativistic particles in the energy range from $\gamma_{\rm min}$ to $\gamma_{\rm max}$ dominate the energy density ε, and neglecting cold particles. For ultra-relativistic particles the pressure $P = \frac{1}{3}\varepsilon$, so for the Alfvén speed we have:

$$\beta_{\rm A} = \frac{1}{\sqrt{1 + \frac{4\mu_0 \varepsilon}{3B^2}}} \tag{5.34}$$

With our assumptions, the value of the energy density is given by:

$$\varepsilon = \int_{\gamma_{\rm min}}^{\gamma_{\rm max}} N_{\rm e}(\gamma) \cdot \gamma m_{\rm e} c^2 d\gamma \tag{5.35}$$

The process of stochastic acceleration of electrons (for $q = 2$) is described in the kinetic equation (Eq. 5.2) by two terms:

(1) $$-\frac{\partial}{\partial \gamma}\left(2D_0 \gamma N_{\rm e}(\gamma, t)\right)$$

is due to the drift of electrons to higher Lorentz factors, with $2D_0\gamma$ being proportional to particle energy gain per unit of time,

(2) $$\frac{\partial}{\partial \gamma}\left(D_0 \gamma^2 \frac{\partial N_{\rm e}(\gamma, t)}{\partial \gamma}\right)$$

describes the diffusion of electron distribution in the Lorentz factor space.

The Eq. 5.32 allows us to express the quantity D_0 appearing in these two terms through the Fermi-II acceleration time-scale $t_{\rm FII}$. The energy-diffusion coefficient is $D_0\,\gamma^2$, and $D_0 = 1/t_{\rm FII}$ is the inverse of the time-scale of the stochastic acceleration process.

Another important parameter, is the duration of the Fermi-II acceleration phase, $t_{\rm dur,FII}$. This parameter appears because the turbulent energy can be injected in a considered region not on a continuous basis, but only during a certain time interval, or even in an intermittent manner.

The term producing the diffusion in the momentum space and broadening of the spectral shape is due to the stochastic nature of the acceleration process (particles experience both energy gains or losses in scatterings). This particularity introduces a curvature in the particle spectrum produced by the Fermi-II acceleration mechanism, and the broad-band particle distribution follows a log-parabola, rather than a power

law (Tramacere et al. 2011).

Also, the Fermi-II mechanism can provide particle spectra harder than $dN_e/d\gamma \propto \gamma^{-2}$ (typical for shock acceleration). This result was demonstrated by Virtanen & Vainio (2005), who performed test-particle simulations of the Fermi-II process in the turbulent downstream of a relativistic shock. In addition, simulations done by Virtanen & Vainio (2005) showed that the stochastic acceleration can be an efficient mechanism in re-acceleration of pre-accelerated particles. This is due to an increase of mean free path of particles with increasing energy in the case of Kolmogorov and Kraichnan turbulences (see Eq. 5.30), and for the "hard-sphere" case, even higher efficiency of re-acceleration is achieved, since, as it was discussed, "hard-sphere" turbulence yields the shortest time-scales of acceleration of high-energy particles among different turbulence types.

First-order Fermi acceleration mechanism (Fermi-I)

This acceleration process will happen if a shock crosses the emitting blob. The Fermi-I (or diffusive shock) acceleration mechanism operates at the fronts of hydrodynamical shock waves, which form in the presence of velocity (and pressure) discontinuities. A shock wave is a disturbance propagating in a medium, characterized by a sharp jump of velocity, pressure and temperature. To illustrate the origin of the shock acceleration term in the kinetic equation (Eq. 5.1), we present here the description of acceleration of relativistic charged particles by a strong shock by Bell (1978).

First, let us focus on the dynamics of a shock itself. Let us consider a shock traveling in plasma with a velocity U, which is highly supersonic or super-Alfvénic, $U \gg v_{\rm A}$. The shock is formed at the contact surface between two media with different parameters: the unperturbed medium ahead of the shock (upstream) with density ρ_1, pressure P_1 and temperature T_1 and the medium behind the shock (downstream) having density ρ_2, pressure P_2 and temperature T_2 (see panel (a) of Fig. 5.2). It is convenient to place the description in the reference frame in which the shock is at rest, then the upstream plasma moves towards the shock front with a velocity $v_1 = U$, and the downstream plasma is receding from the shock with a velocity v_2 (see panel (b) of Fig. 5.2).

We consider a medium with an equation of state $P = (\Upsilon - 1)\, u_{int}$, where Υ is an adiabatic index, and u_{int} is the internal energy of the gas per unit of volume. In the reference frame of the shock, the junction conditions for conservation of mass, momentum and energy across the shock front are:

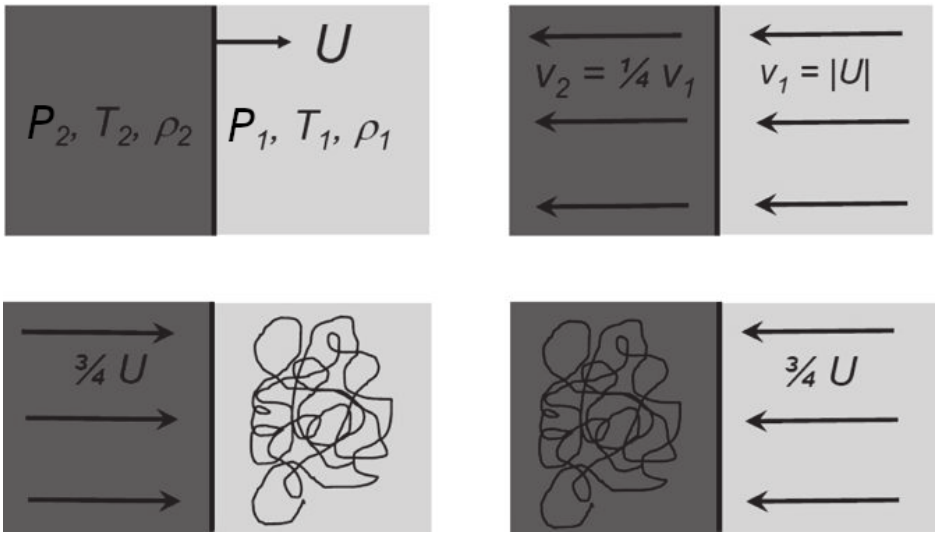


Figure 5.2: The dynamics of the medium in the proximity of a shock wave. (a): The dynamics of a shock as seen in the observer's frame. The shock is moving with a velocity U through a stationary medium. The light gray color indicates the upstream plasma having density ρ_1, pressure P_1 and temperature T_1, and the dark gray – the downstream plasma with density ρ_2, pressure P_2 and temperature T_2. (b): The dynamics of the upstream and downstream media in the reference frame in which the shock is at rest. The downstream gas is receding from the shock wave at a velocity $v_2 = \frac{1}{4}v_1 = \frac{1}{4}U$. ($c$): the same as (b) but in the reference frame of the upstream plasma, in which the particle distribution is isotropic. The downstream medium is approaching the upstream at a velocity $\frac{3}{4}U$. (d): the same as (c) but in the reference frame in which the downstream medium is stationary. The upstream plasma flows towards the downstream at a velocity $\frac{3}{4}U$. (adapted from Longair (2011))

$$\rho_1 v_1 = \rho_2 v_2 \tag{5.36}$$

$$P_1 + \rho_1 v_1^2 = P_2 + \rho_2 v_2^2 \tag{5.37}$$

$$\frac{\Upsilon}{\Upsilon - 1}\frac{P_1}{\rho_1} + \frac{v_1^2}{2} = \frac{\Upsilon}{\Upsilon - 1}\frac{P_2}{\rho_2} + \frac{v_2^2}{2} \tag{5.38}$$

For a case of a "strong shock", the pressure of the upstream medium can be neglected, $P_1 \approx 0$. Let us rewrite the junction conditions in terms of the "compression ratio" $R = \rho_2/\rho_1$:

$$\frac{v_1}{v_2} = R \tag{5.39}$$

$$\frac{P_2}{\rho_2} = v_1^2 \frac{R-1}{R^2} \tag{5.40}$$

$$\frac{2\Upsilon}{\Upsilon - 1} \frac{R-1}{R^2} + \frac{1}{R^2} = 1 \tag{5.41}$$

From this we express the compression ratio:

$$R = \frac{\Upsilon + 1}{\Upsilon - 1} \tag{5.42}$$

For a fully ionized plasma (ideal gas), $\Upsilon = 5/3$, and so $R = 4$, which means that a strong shock wave compresses the medium increasing its density by a factor of 4. It also heats the downstream medium to the temperature

$$T_2 = \frac{m_0}{\rho_2 k} P_2 = \frac{m_0 v_1^2}{k} \frac{R-1}{R^2} = \frac{3\, m_0 v_1^2}{16\, k} \simeq 10^7 \left(\frac{v_1}{10^3\,\mathrm{km/s}} \right)^2 \mathrm{K} \tag{5.43}$$

Let us consider an astrophysical example. Following a supernova explosion, the supernova shell is expanding with a velocity (typically) $v_1 \sim 10^3$ km/s, and, according to the Eq. 5.43, the downstream medium will be heated up to $T_2 \sim 10^7$ K, which implies that supernovae must be prominent X-ray sources. This thermal X-ray emission is indeed detected by X-ray telescopes (e.g. one could see Chandra X-ray images of Cas A).

The medium behind the shock is also highly turbulent. As the upstream plasma arrives at the shock front, its kinematic streaming motion is randomized and converted into turbulent and thermal motions, leading to formation of a turbulent wake in the downstream plasma, where the turbulent energy is dissipated.

Let us now examine the velocity discontinuity in the vicinity of the shock. In the reference frame of the shock, the upstream medium moves faster than the downstream one ($v_1 > v_2$), and the two media are approaching each other at the speed

$$v_{\mathrm{rel}} = v_1 - v_2 = v_1 - \frac{v_1}{R} = \frac{3}{4} v_1 = \frac{3}{4} U \tag{5.44}$$

This is illustrated in panels (c) and (d) of Fig. 5.2.

Now let us consider the dynamics of plasma particles that are crossing the shock. As astrophysical plasmas are magnetized, we focus our attention on a collisionless shock propagating through magnetized plasma. By "collisionless" one means that the thickness of the shock front is much less than the mean free path of plasma particles which cross it. Because of that, these particles hardly notice the shock, and travel freely between the upstream and downstream medium. Let us follow the motion of relativistic particles, initially situated ahead of the shock. Particles from the upstream cross the shock front and enter the turbulent downstream medium. The magnetic field irregularities associated with the turbulent motions, scatter the particles back to the upstream. As the beam of particles re-injected in the upstream plasma is flowing with super-Alfvénic velocity, this leads to excitation of Alfvén waves via streaming instability (Wentzel 1974). The waves scatter the particles and limit their streaming velocity to the Alfvén one. As a result, the particles will be quickly swept by the approaching super-Alfvénic shock, which leads to their return back to the downstream medium. These two scattering processes, firstly, quickly isotropize the velocity distribution of the particles in the frame of reference of the moving plasma on both sides, and secondly, prevent particles from streaming away from the shock. Therefore the plasma particles will find themselves to be confined around the shock front. Particles bounce back and forth between the upstream and the downstream, and are able to recross the shock front a large number of times.

Let us now work out the energy gain of a particle after a crossing. The shock is assumed to be non-relativistic, $U \ll c$. Considering a particle with energy E, as measured in the reference frame of the upstream plasma, we find that, according to Lorentz transformations, it will have energy

$$E' = E\left(1 + \frac{v_{\rm rel}}{c}\cos\theta\right) \tag{5.45}$$

as seen in the downstream medium. Here θ is the angle between the velocity vector of the particle, and the vector normal to the shock front. Therefore, after passing from the upstream to the downstream, the particle will gain

$$\Delta E = E' - E = E\,\frac{v_{\rm rel}}{c}\cos\theta \tag{5.46}$$

We now need to average the energy gain over angles. As the velocity distribution of particles is isotropized, the probability that a particle will move at an angle between θ and $\theta + d\theta$ is proportional to $\sin\theta\, d\theta$. The particle flux flowing towards the shock is proportional to the projection of their velocity vectors on the shock normal, $c\cos\theta$. The overall, normalized to unity, probability of a particle traversing the shock front

is therefore, $dp_{\rm cross}(\theta) = 2\sin\theta\cos\theta\, d\theta$. Averaging the relative energy gain, using this probability, we obtain

$$\left\langle \frac{\Delta E}{E} \right\rangle_{u\to d} = \frac{2}{3}\frac{v_{\rm rel}}{c} = \frac{1}{2}\frac{U}{c} \tag{5.47}$$

This is an average energy gain of a particle after passing from the upstream to the downstream. Behind the shock front, turbulent "magnetic mirrors" scatter particles, randomizing their velocity vectors, and bouncing them to the upstream. After re-crossing the shock front and returning back to the upstream medium, the particle will again increase its energy, gaining the same amount of energy, $\Delta E = \frac{1}{2}\frac{U}{c}E$. Thus, in one round-trip across the shock front, the particle on average gains

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{U}{c} \tag{5.48}$$

A crucial feature of this process is that every time the particle traverses the shock front it gains energy, there are no reflections in which the particle loses energy, unlike the Fermi-II mechanism. As the particles are confined around the shock and are bouncing back and forth across its front, after multiple crossings, the particle will be accelerated to high energies. Since the relative energy gain of the particle in one crossing is first order in the shock velocity (Eq. 5.48), the mechanism is referred to as "first-order Fermi".

Qualitatively, the energy gain during a crossing occurs because the "inhomogeneties" that scatter the particles back, have different velocities in the upstream and downstream plasma. Overall, during the Fermi-I acceleration process, the particles extract kinetic energy from the shock propagating through the plasma.

The same scattering process which entraps the particles near the shock is also responsible for their escape. The downstream plasma is receding from the shock front with a velocity $v_2 = \frac{1}{4}U$, leading to advection of the accelerated particles away from the shock, since their velocity distribution is isotropic in the medium frame. Let us calculate the probability of escape of a particle from the shock. The overall flux of advected particles is, $F_{\rm adv} = \frac{1}{4}nU$, where n is the number density of particles. The flux of particles injected from the upstream to the downstream is, $F_{\rm cross} = \frac{1}{4}nc$ (after averaging over the incidence angles). This is an average flux of particles crossing the shock front in either direction. Therefore, the fraction of particles which are escaping the shock is, $f_{\rm esc} = F_{\rm adv}\,/\,F_{\rm cross} = U/c$.

The number of particles which leave the acceleration region after k crossings, is

$N_k = N_0\,(1 - f_{\rm esc})^k = N_0\,(1 - \frac{U}{c})^k$. The energy of a particle that escapes after k crossings, is $E_k = E_0\,(1 + \frac{U}{c})^k$. Therefore, the number of particles with energy higher than E_k, is

$$N(E \geq E_k) \propto E^{-\frac{U/c}{U/c}} \propto E^{-1} \tag{5.49}$$

This corresponds to a differential energy spectrum

$$\frac{dN}{dE} \propto E^{-2} \tag{5.50}$$

Therefore shock acceleration mechanism yields a power law with an index $\alpha = 2$. This result is obtained for an ideal gas (adiabatic index $\Upsilon = 5/3$), having compression ratio of $R = 4$. For an arbitrary compression ratio (and adiabatic index), the energy spectrum is

$$\frac{dN}{dE} \propto E^{-\tilde{\alpha}}\,, \quad \tilde{\alpha} = \frac{2+R}{R-1} = \frac{3\Upsilon - 1}{2} \tag{5.51}$$

Thus, the energy spectrum of the particles ejected from the shock, is a power law with an index from 2 to 2.5, depending on the gas compression ratio (typically 3 or 4). For a relativistic unmagnetized shock (magnetized shocks are much less efficient at acceleration of particles) a theoretical study by Kirk et al. (2000) predicts the slope of $\alpha = 2.23$, and the numerical simulations by Haugbølle (2011) show a consistent value of $\alpha \sim 2.2$.

Comparing shock and stochastic acceleration, one notices the difference in terms of (*i*) the *slope of the particle spectra*: the Fermi-II process is capable of producing much harder particle spectra than the Fermi-I yielding a slope $\alpha = 2$, and (*ii*) the *spectral curvature*: the Fermi-II mechanism introduces important spectral curvature due to its stochastic nature, which is not the case for the Fermi-I. The discernible features of each acceleration process can be used to disentangle the acceleration scenarios in blazar jets, as well as in a variety of other astrophysical environments.

Shock acceleration is a ubiquitous phenomenon thought to occur in many astrophysical systems, in particular in AGN jets (e.g. Marscher & Gear (1985)). It is therefore natural to consider the possibility that the blazar outbursts could be triggered by a shock passing through the emitting zone.

Continuously operating on a long-term basis, the Fermi-I process could also serve as an efficient pre-acceleration mechanism supplying high-energy particles to the blob

(e.g. Kirk et al. (1998)). Pre-accelerated particles can be further re-accelerated by another shock and/or the second-order Fermi mechanism.

We treat the shock acceleration process with the kinetic approach. The Fermi-I process is considered as a systematic energy gain, and is described in the kinetic equation (Eq. 5.1) by the term $-a\gamma$, which is the Fermi-I acceleration rate, a quantity proportional to the particle energy gain per unit of time. Here $a = 1/t_{\rm FI}$ is the inverse of the characteristic time-scale of the shock acceleration.

Let us make an order of magnitude estimate for the Fermi-I acceleration time-scale depending on the parameters of the shock. The acceleration time-scale is (approximately) the time it takes for the particle to double its energy. The number of shock crossings after which the particle will increase its energy by a factor of 2, is $k_{\rm db} \approx \ln 2 \,/\, \ln(1 + U/c)$. For a relativistic particle, the time of one crossing is $\Delta t_{\rm sfc} \sim \lambda_{\rm f}/c$, where $\lambda_{\rm f}$ is the particle mean free path, defined by the turbulent scattering process (Eq. 5.30) and the plasma streaming instability. The Fermi-I time-scale is then $t_{\rm FI} \simeq k_{\rm db}\, \Delta t_{\rm sfc} \simeq \frac{\ln 2}{\ln(1+U/c)}\, \frac{\lambda_{\rm f}}{c}$. For a non-relativistic shock, $t_{\rm FI} \sim \lambda_{\rm f}/U$, and for a highly relativistic shock, $t_{\rm FI} \sim \lambda_{\rm f}/c$, so that $t_{\rm FI} \sim \lambda_{\rm f}/U$ for an arbitrary value of U. If the mean free path $\lambda_{\rm f}$ of the particle is energy-independent (e.g. in the case of the "hard-sphere" turbulence in the downstream), the Fermi-I acceleration time-scale is also energy-independent.

An additional very important parameter arising when one considers passage of a shock wave through a spatially-limited volume/region, is the duration of the acceleration episode by a shock, $t_{\rm dur,FI}$, which is equal to the time it takes for the shock to cross the region, $t_{\rm cross} = R/U$, where R is the spatial extension of the zone perturbed by a shock.

Magnetic reconnection

Another mechanism of particle acceleration invoked to explain blazar flares, is magnetic reconnection. This is a process occurring in highly conductive magnetized plasma, during which the configuration of the magnetic field quickly rebuilds, leading to particle acceleration and heating of the plasma. Although we do not include the magnetic reconnection process into our modeling, it is nevertheless interesting to consider various properties of this still not fully understood phenomenon, especially because as we will see later on, particle acceleration due to magnetic reconnection can be at the first order described within a kinetic approach with a term having the same form as the one of the shock acceleration.

The first theoretical model for magnetic reconnection was developed by Peter

Sweet (Sweet 1958) and Eugene Parker (Parker 1957) in the framework of resistive MHD. The general view of this physical process is illustrated in Fig. 5.3. The plasma is considered to be ideal everywhere with except for a small diffusion region of width 2δ, where the frozen-in condition is breached. The violation of this condition in this zone happens because of, firstly, strong currents generated in this region due to magnetic field spatial gradient, secondly due to a finite conductivity σ (and hence a non-zero resistivity) of the astrophysical plasma, and thirdly, because this zone features a stagnation point where the speed of plasma $v \sim 0$, so the diffusion term in the MHD equation governing the evolution of the magnetic field becomes dominant relatively to the frozen-in term.

The plasma is considered to have a complex magnetic field, in which there are separate bundles of field lines that have their base points in different poles. Flows of plasma bring together two magnetic field lines with an opposite field direction over a spatial extent of $2L$. Sharp spatial gradients of the magnetic field in the diffusion region generate strong electric currents in it, which can be seen from one of the Maxwell equations, Ampère's circuital law:

$$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{j} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \tag{5.52}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic field vectors, $\mathbf{j}$ is the electric current.

The narrower is the width of the reconnection region, the stronger will be the current. Because of the presence of currents, this region is sometimes also referred to as a current layer/sheet. Within the current sheet, the magnetic field energy is converted into the kinetic energy of plasma motion and heat via the Ohmic dissipation (the rate of which is $\propto j^2/\sigma$) increasing the pressure, and via the magnetic tension due to a large spatial gradient of the magnetic field. As the frozen-in condition is violated inside the current sheet, the magnetic flux diffuses through it, leading to a break of magnetic field lines and their subsequent rejoining in a rearranged topology. As a result, reconnected field lines and plasma are ejected from the diffusion region.

In Sweet-Parker's view, the reconnection process is *stationary*, and is therefore based on the following conditions. *Firstly*, the energetic balance: an important fraction of the magnetic field energy is converted into kinetic energy of the outflow, $\frac{B^2}{2\mu_0} \sim \frac{1}{2}\rho v_{\rm out}^2$, meaning that the plasma is expelled from the reconnection region at the Alfvén velocity $v_{\rm out} \sim v_{\rm A} = \frac{B}{\sqrt{\mu_0\,\rho}}$. Here ρ is the density of the plasma, and $v_{\rm out}$ is the velocity of the ejected plasma. *Secondly*, the mass of plasma is conserved, $v_{\rm in}\, L = v_{\rm A}\, \delta$, with $v_{\rm in}$ being the inflow velocity. *Thirdly*, the electric field is constant in the stationary state, so that Eq. 5.52 becomes: $\nabla \times \mathbf{B} = \mu_0\, \mathbf{j}$, and so $B/\delta \sim \mu_0\, j$.

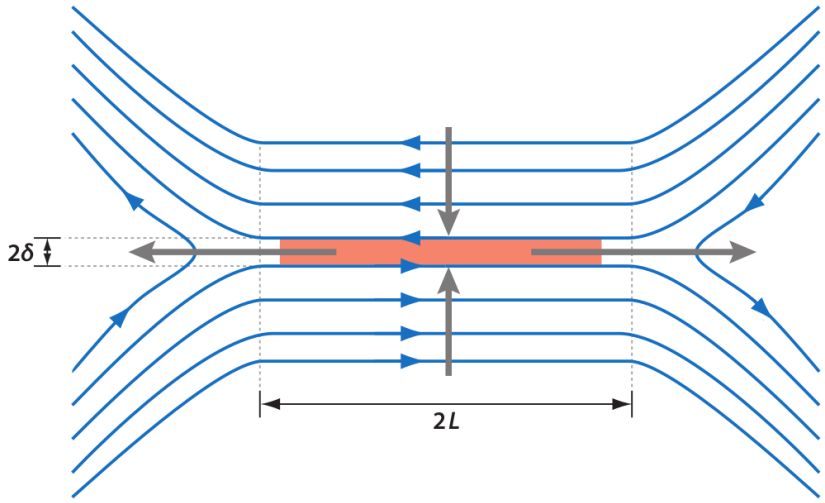


Figure 5.3: Sketch illustrating magnetic reconnection in Sweet-Parker model. Two oppositely directed magnetic field lines carried by the plasma motions approach each other closely and reconnect in a current sheet with a width 2δ. (adapted from Zweibel & Yamada (2009))

From the Ohm's law in the resistive MHD, this electric field is

$$\mathbf{E} + [\mathbf{v} \times \mathbf{B}] = \frac{\mathbf{j}}{\sigma} \tag{5.53}$$

Within the diffusion region the electric field is resistive, so that $v_{\rm in}\, B \sim j/\sigma$.

Now let us estimate the inflow velocity $v_{\rm in}$ and the characteristic reconnection time-scale $t_{\rm rec}$. Combining the expressions formulated in the three conditions above, one obtains

$$v_{\rm in} = \frac{\delta}{L}\, v_{\rm A} = \sqrt{\frac{\eta v_{\rm A}}{L}} \tag{5.54}$$

where $\eta = \dfrac{1}{\mu_0\, \sigma}$ is the plasma diffusivity.

The associated reconnection time-scale is

$$t_{\rm rec} = \frac{L}{v_{\rm in}} = \sqrt{\frac{L^3}{\eta v_{\rm A}}} \tag{5.55}$$

As the Ohmic dissipation rate inside the current sheet is comparable to the flux

of the kinetic energy of the plasma outflow, $\frac{j^2}{\sigma} L \sim \frac{1}{2} \rho v_{\rm A}^2 v_{\rm A}$, fractions of the magnetic energy converted into heat and into kinetic energy of plasma motion are comparable.

Sweet-Parker model predicts reconnection time-scales (∼1 year) significantly shorter than global diffusion of the magnetic field $t_{\rm diff} \simeq \mu_0 \sigma L^2 \sim 10^7$ y, but still much longer than observed during solar flares (∼ 1 h or 1 min) (e.g. Galtier (2016)). The model is therefore too slow, and cannot adequately explain astrophysical reconnection; additional effects have to be included to arrive to a better agreement with the observations.

In 1964, Harry Petschek extended the Sweet-Parker model by assuming that the inflow and outflow regions are separated by slow mode shocks (in an X-point geometry) (Petschek 1964). With such configuration, the aspect ratio of the reconnection layer (ratio of its spatial extension length L to its thickness δ) is substantially decreased (to an order of unity), which, according to Petschek calculations, allows the reconnection to proceed faster. The predictions of this model are more consistent with the observations, however it has a major difficulty: slow mode shocks have never been observed in self-consistent numerical simulations (e.g. Malyshkin et al. (2005)).

Another model proposed to solve the problem of slow reconnection rates is stochastic reconnection model (Lazarian & Vishniac 1999). In this model, the magnetic field of the plasma has a small-scale random (stochastic) component due to turbulence. Magnetic field lines are wandering due to turbulent motions of plasma, and the fluctuations present in the small-scale structure of the magnetic field lines, bring together two magnetic field lines locally, which allows many local reconnection events to occur simultaneously. In addition, as the reconnections proceed at small spatial scales, the transverse scale for reconnection flows is reduced, as compared to Sweet-Parker model. These two effects combined together accelerate significantly the reconnection rate. It appears to be independent of the plasma resistivity, local physics of the reconnection, nature and model of the turbulence, and is controlled only by the level of magnetic field stochasticity.

Various studies (e.g. Furth et al. (1963)) also find that elongated current sheets are prone to tearing-like instabilities, which fragment it into magnetic islands, or plasmoids (e.g. Loureiro et al. (2007)). This effect can speed up the reconnection; also the interactions between these plasmoids can lead to acceleration of particles with resulting spectral indices between 1 and 2 (e.g. PIC simulations by Sironi & Spitkovsky (2014)).

The effect which can boost the reconnection rate is anomalous resistivity of plasma. This phenomenon augments the magnetic diffusivity and also enhances the energy release during a reconnection event. In case waves are present in the plasma,

the collective effects in the medium become much stronger since particles move more coherently, and motion of an individual electron influences a large number of particles rather than only the neighboring ones. This increases the electron scattering rate and the effective resistivity. One of the mechanisms providing an anomalous resistivity is, for instance, the current-driven ion-acoustic instability (Coroniti & Eviatar 1977), which is excited when the drift velocity of plasma exceeds the ion sound speed and leads to rapid growth of the amplitude of ion-acoustic waves, destabilizing the modes.

The magnetic reconnection was proposed as one of the mechanisms powering fast (time-scale of ∼1 min) TeV flares in blazars (e.g. Giannios et al. (2009)). One of the studies testing such physical scenario, conducted by Morris et al. (2019), considers particle acceleration in a plasmoid crossing reconnection layer between regions of opposing magnetic field lines. The plasmoid grows steadily over time due to collisions and merging with other plasmoids. The modeling by Morris et al. (2019) supports the overall feasibility of rapid flares with the magnetic reconnection process, however authors find that it is not possible to produce sufficient amount of IC emission within the conventional SSC scenario (without taking into the account external photons). Another issue of such model is that the size of the plasmoids has to be comparable to the entire jet radius, i.e. unrealistically large.

Within the kinetic approach, the term describing the acceleration process due to magnetic reconnection, can have the same form as for the shock acceleration: $\dot{\gamma}_{\rm rec} = a_{\rm rec}\gamma$, where $a_{\rm rec} \sim 1/t_{\rm rec}$ is the model-dependent reconnection rate, and $t_{\rm rec}$ is the characteristic model-dependent reconnection time-scale. In the scenario by Morris et al. (2019), $a_{\rm rec} = \dfrac{\alpha}{\tau_{\rm merge}}$, where α is the dimensionless free parameter quantifying the efficiency of the acceleration process due to the magnetic reconnection (average energy gain per particle per merging event), and $\tau_{\rm merge}$ is the time-scale of plasmoid merging.

Therefore, the magnetic reconnection in the first approximation can be introduced in the kinetic equation (Eq. 5.1) as an additional Fermi-I-like acceleration term.

5.1.4 Particle escape

The term $\dfrac{N_{\rm e}(\gamma, t)}{t_{\rm esc}}$ describes the escape of particles from the blob, with a characteristic escape time-scale $t_{\rm esc}$, which in general depends on the particle energy. In case of absence of turbulence in the emission zone, particles escape it freely (free-steaming), and the escape time-scale will be $t_{\rm esc} \sim 1\,R/c$. Particles that undergo stochastic

acceleration will escape the region at a longer time scale, due to resonant pitch-angle scattering (diffusion process). The spatial diffusion coefficient and the momentum diffusion coefficient are linked as $D_x\, D_p \approx \beta_{\rm A}^2\, p^2$ (Skilling 1975). The time-scale of escape from a turbulent region is therefore related to the stochastic acceleration time-scale as (Tramacere et al. 2011):

$$t_{\rm esc}^{(turb)} = \frac{R_{\rm t}^2}{c^2 \beta_{\rm A}^2\, t_{\rm FII}} \tag{5.56}$$

, where $R_{\rm t}$ is the size of the turbulent zone.

Then, combining Eq. 5.29 and 5.56, one finds that for an arbitrary slope of the turbulence spectrum q, the escape time-scale is:

$$t_{\rm esc}^{(turb)} = \left(\frac{R_{\rm t}}{c}\right)^2 \left(\frac{\delta B}{B_0}\right)^2 \frac{c}{\lambda_{\rm max}} \left(\frac{r_{\rm L}}{\lambda_{\rm max}}\right)^{q-2} \tag{5.57}$$

As one can see, the escape time scales as $t_{\rm esc}^{(turb)} \propto p^{q-2} \propto \gamma^{q-2}$, which implies faster escape for higher energy electrons for the Kolmogorov and for the Kraichnan type turbulences. For the case of "hard-sphere" turbulence, $t_{\rm esc}$ is energy-independent:

$$t_{\rm esc}^{(turb)} = \left(\frac{R_{\rm t}}{c}\right)^2 \left(\frac{\delta B}{B_0}\right)^2 \frac{c}{\lambda_{\rm max}} \tag{5.58}$$

This expression is used in our modeling to evaluate the time-scale of electron escape from a turbulent region.

5.2 EMBLEM code

The physical processes discussed above, are thought to operate in the blazar emitting zone. In the (quasi-)stationary flux state these processes balance each other, so that the particle distribution does not evolve. A flare emerges if one of the processes becomes dominant, or if a previously not active physical process "switches on". Such imbalance causes evolution of the particle distribution and of the energy flux. Intuitively, particle injection and acceleration cause a flux increase initiating a flare, and particle cooling and escape cause the flux decrease, damping the flare. The time evolution of the particle spectrum is determined by the competition between the above-mentioned physical processes, and is described by the kinetic equation (Eq. 5.1). The associated time-dependent broad-band emission can be computed using a certain

blazar emission model (leptonic SSC/EC or hadronic). This allows to simulate flaring activity of blazars as a function of the parameters of the above-mentioned physical processes.

Following the kinetic approach discussed in the previous section, I developed a numerical code "**EMBLEM**" (*Evolutionary Modeling of BLob EMission*) for time-dependent modeling of blazar emission during flares. In my code, I numerically solve the kinetic equation for the case of "hard-sphere" turbulence (Eq. 5.2) and calculate the associated varying MWL emission using *leptonic SSC model*.

In this section, we first present the EMBLEM numerical code, including its numerical implementation, architecture and input parameters, and then we perform a parameter space study simulating a range of example flares with the code.

5.2.1 Numerical implementation

The code is written in the Python programming language. This language was chosen because of a large variety of tools already available in Python packages, e.g. special functions, in particular Bessel functions needed for computation of synchrotron emission, physical constants, interpolation methods, numerical integration procedures, solvers for system of equations, etc.

Electron spectrum evolution: Chang and Cooper scheme

We use a fully implicit difference scheme by Chang & Cooper (1970) to numerically solve the kinetic equation Eq. 5.2 and retrieve the time evolution of the electron spectrum on a time-grid.

The numerical scheme is designed specifically for solving Fokker-Planck equations of the general form

$$\frac{\partial n(x,t)}{\partial t} = \frac{1}{A(x)} \frac{\partial}{\partial x} \left[B(x,t)\, n(x,t) \, + \, C(x,t) \frac{\partial n(x,t)}{\partial x} \right] \tag{5.59}$$

where t is time, x is usually representing momentum ($0 \leq x < \infty$), $n(x,t)$ is a particle distribution function in momentum space (typically, number of particles per unit of volume and per unit of momentum interval), A is a function of x, and B and C are functions of x and t.

This equation does not include the injection and escape terms, appearing in

the kinetic equation Eq. 5.2. We first consider the numerical scheme for a particle conserving case, and then extend this numerical scheme to take into account injection and escape processes.

The Fokker-Planck equation is a linear partial differential equation describing the time evolution of the probability density function of the momentum of a particle in processes with a stochastic component. Equivalently, (for an ensemble of particles), the equation governs the time evolution of the distribution of particles in momentum space under the influence of systematic energy gain or loss processes, as well as random forces. This equation appears in various problems of statistical mechanics (e.g. Brownian motion) and plasma physics (Boltzmann equation, or its partial case without collision term, Vlasov equation, in the kinetic theory of plasma).

As many other partial differential equations, the Fokker-Planck equation can be solved analytically only in a few special cases. The numerical scheme by Chang & Cooper (1970) aims to find numerical solutions with are (i) non-negative (a negative number of particles or a negative probability is an unphysical situation), and (ii) particle conserving in the absence of external sources and sinks. The number of particles is conserved if we assume no particle reactions, and no particle creation and annihilation; in this case particle conservation law is equivalent to the universal mass-energy conservation corresponding to the time invariance symmetry. As we neglect the internal γ-γ absorption, the number of particles in the emitting zone is conserved.

The first condition is guaranteed if A, B and C are all positive functions, and the second condition is ensured automatically in case proper boundary conditions are set. Let us consider that in more detail. Multiplying both sides of the Eq. 5.59 by $A(x)$ and integrating over x over its entire range from 0 to ∞, we obtain

$$\frac{\partial N}{\partial t} = \left[B(x,t)\, n(x,t) \; + \; C(x,t)\, \frac{\partial n(x,t)}{\partial x} \right] \Bigg|_0^\infty \tag{5.60}$$

with $N = \int_0^\infty n(x,t)\, A(x)\, dx$.

This equation can be interpreted as a kind of continuity equation: change of the quantity N in time is defined by the generalized flux $F = Bn + C\,\frac{\partial n}{\partial x}$ at the two boundaries, $x = 0$ and $x \to \infty$, or more precisely, the difference of the two fluxes, being the net flux representing the exchange of the system with the "environment". If we impose boundary conditions such that the RHS is zero, meaning that we do not allow any flux (inflow or leak) at the boundaries, then we can interpret the quantity N as a total number of particles in the system, which is conserved, $\frac{\partial N}{\partial t} = 0$, fulfilling our second requirement.

The quantities x, t, and $n(x,t)$ are defined on a grid, specific to each quantity. The momentum is defined on a grid with (usually) non-equidistant points, $x = \{x_0, x_1, \ldots, x_j, \ldots, x_{max}\}$; the time grid is homogeneous, $t_k = \Delta t \cdot k$. The function $n(x,t)$ is sampled with the discrete set $u_j^k = u(x_j, t_k)$. The time derivative is chosen to be forward differenced, $\frac{\partial n}{\partial t} = \frac{n^{k+1} - n^k}{\Delta t}$, and the momentum derivatives will be center differenced:

$$\left.\frac{\partial^2 n}{\partial x^2}\right|_j = \frac{n_{j+1} - 2n_j + n_{j-1}}{(\Delta x)^2} \tag{5.61}$$

Replacing all the quantities in the Eq. 5.59 by their samples on the grids, and the derivatives with the finite differences, we will obtain a difference equation, containing n^{k+1}. This equation has to be solved implicitly. The scheme is adjusted in such a way, because the implicit method is known to be faster and more stable than the explicit one.

It is shown in Chang & Cooper (1970) that when expanding the derivative of the generalized flux F in the RHS and evaluating the derivatives of B and C directly (for known analytical functions), while replacing derivatives of n with approximate differencing relations, the scheme can lose its particle conserving property. Therefore, Eq. 5.59 has to be differenced as it is written, i.e. in the RHS, the derivative of the generalized flux has to be replaced with a differencing relation. This ensures that the strict particle conservation is achieved regardless of the grid for x or the number of time steps.

Let us write the differencing equation. We use the centered difference in the momentum for the generalized flux F and a fully implicit representation for $n(x,t)$ that appears in the expression for F

$$\frac{1}{\Delta t}\left(n_j^{k+1} - n_j^k\right) = \frac{1}{A_j\,\Delta x}\left(\tilde{F}_{j+1/2} - \tilde{F}_{j-1/2}\right) \tag{5.62}$$

It is important to note that the step in the momentum Δx in general case depends on the momentum x, as we are dealing with a non-homogeneous grid. Typically, the momentum grid covers a few decades, and because of that, in most cases, x is sampled with an equidistant step in *logarithmic* space, and the sequence of x-grid points thus represents a geometric progression. For our case of centered difference for the momentum derivative, the step of the x-grid is

$$\Delta x = \Delta x_j = x_{j+1/2} - x_{j-1/2} \tag{5.63}$$

The tilde over the generalized flux F in the Eq. 5.62 indicates the implicit aspect of the scheme:

$$\begin{aligned} \tilde{F}_{j+1/2} = B^k_{j+1/2}\, n^{k+1}_{j+1/2} \,+\, C^k_{j+1/2} \left.\frac{\partial n}{\partial x}\right|_{(j+1/2)\,,\,(k+1)} = \\ = \, B^k_{j+1/2}\, n^{k+1}_{j+1/2} \,+\, C^k_{j+1/2}\, \frac{1}{\Delta x}\,(n^{k+1}_{j+1} - n^{k+1}_{j}) \end{aligned} \tag{5.64}$$

Let us now explore the boundary conditions that have to be set for the finite difference scheme. Multiplying the Eq. 5.62 by $A_j\,\Delta x$, and summing up both LHS and RHS over the x-grid, from $j = 0$ to $j = J$, we get

$$\sum_{j=0}^{J} A_j \frac{\Delta x}{\Delta t}\,(n^{k+1}_j - n^k_j) \;=\; F_{J+1/2} - F_{-1/2} \tag{5.65}$$

The LHS is proportional to the change of the total number of particles when going to the next time step, and should be equal to zero to ensure particle number conservation. The RHS is then also equal to zero, which results in a boundary conditions

$$F_{J+1/2} = F_{-1/2} = 0 \tag{5.66}$$

as we require to have no flux through either boundary.

In Eq. 5.64, one could notice the term $n^{k+1}_{j+1/2}$, which is $n(x,t)$ in middle between the x-grid points, x_j and x_{j+1}. To solve the scheme, we need to know $n_{j+1/2}$, which is not the case. One could therefore see the need to "interpolate" this value using the neighboring values. The approach of Chang & Cooper (1970) is to express this quantity as a linear combination of n_j and n_{j+1}:

$$n^{k+1}_{j+1/2} = (1-\delta_j)\, n^{k+1}_{j+1} \,+\, \delta_j\, n^{k+1}_j \tag{5.67}$$

Since δ_j is related to $n^{k+1}_{j+1/2}$, we keep in mind the notation $\delta_j = \delta^{k+1}_j$.

Chang & Cooper (1970) demonstrated that simply taking $\delta_j = 1/2$ (arithmetic average) might yield negative solutions, if the steps in x are not narrow enough. To avoid this problem, simple forward differencing, $\delta_j = 0$, can be used. However it will not give convergence unless, again, the x-grid is fine enough. The weights δ_j

are therefore deduced requiring that the numerical scheme converges, produces non-negative solutions, and works with a satisfactory accuracy. The last constraint is critical, as small errors and inaccuracies induced by the scheme tend to accumulate, degrading the precision of the result. In order to fulfill the last requirement, Chang & Cooper (1970) set a condition that the equilibrium solution ($\frac{\partial n(x,t)}{\partial t} = 0$) of the Fokker-Planck equation (Eq. 5.59), obtained numerically, should match the exact analytical expression. To find the steady solution $n_0(x)$ of the Eq. 5.59, one has to solve the differential equation

$$\frac{\partial}{\partial x}\left[B(x,t)\, n_0(x) \,+\, C(x,t)\,\frac{\partial n_0(x)}{\partial x}\right] \,=\, 0 \tag{5.68}$$

Taking into account the boundary conditions requiring no flux through either boundary, we get

$$B(x,t)\, n_0(x) \,+\, C(x,t)\,\frac{\partial n_0(x)}{\partial x} = 0 \tag{5.69}$$

Its solution is

$$n_0(x) = \alpha \cdot \exp\left(-\int \frac{B(x,t)}{C(x,t)}\, dx\right) \tag{5.70}$$

Obviously, the stationary solution exists in case $\frac{B(x,t)}{C(x,t)}$ is time-independent. Now let us compare $n_0(x_{j+1})/n_0(x_j)$ estimated analytically and numerically. We evaluate this quantity analytically using Eq. 5.70:

$$\frac{n_0(x_{j+1})}{n_0(x_j)} \,=\, \exp\left(-\int_{x_j}^{x_{j+1}} \frac{B(x,t)}{C(x,t)}\, dx\right) \,\simeq\, \exp\left(-\frac{B(x_{j+1/2})}{C(x_{j+1/2})}\,\Delta x\right) \tag{5.71}$$

To obtain the numerical estimate, we first derive the generalized flux taking into account the weighted average for $n_{j+1/2}$. For that, we use the Eq. 5.64, in which we substitute $n_{j+1/2}$ from the Eq. 5.67

$$\tilde{F}_{j+1/2} \,=\, \left((1-\delta_j)\, B^k_{j+1/2} \,+\, \frac{1}{\Delta x}\, C^k_{j+1/2}\right) n^{k+1}_{j+1} \,-\, \left(\frac{1}{\Delta x}\, C^k_{j+1/2} \,-\, \delta_j\, B^k_{j+1/2}\right) n^{k+1}_{j} \tag{5.72}$$

In the equilibrium case, according to the Eq. 5.69, $F_{j+1/2} = 0$. This gives

$$\left((1-\delta_j)\,B^k_{j+1/2} + \frac{1}{\Delta x}\,C^k_{j+1/2}\right) n_0(x_{j+1}) - \left(\frac{1}{\Delta x}\,C^k_{j+1/2} - \delta_j\,B^k_{j+1/2}\right) n_0(x_j) = 0 \tag{5.73}$$

From here one finds

$$\frac{n_0(x_{j+1})}{n_0(x_j)} = \frac{\frac{1}{\Delta x}\,C^k_{j+1/2} - \delta_j\,B^k_{j+1/2}}{\frac{1}{\Delta x}\,C^k_{j+1/2} + (1-\delta_j)\,B^k_{j+1/2}} \tag{5.74}$$

Equating the analytical estimate (Eq. 5.71) to the numerical one (Eq. 5.74), we retrieve the weights δ_j:

$$\delta_j = \delta_j^{k+1} = \frac{1}{\tilde{w}^k_j} - \frac{1}{\exp(\tilde{w}^k_j) - 1} \tag{5.75}$$

where $\tilde{w}^k_j = \dfrac{\Delta x\,B^k_{j+1/2}}{C^k_{j+1/2}}$

For $\tilde{w}^k_j$ increasing from 0 to ∞, these weights decrease monotonically from 0.5 to 0, i.e. shift from centered difference towards the forward difference. It is demonstrated in Chang & Cooper (1970) that implementing these weights (Eq. 5.75) in the Eq. 5.67 guarantees non-negative, stable and accurate solutions.

We now can write the final difference equation (for the general case) to solve the Fokker-Planck equation (Eq. 5.59). Taking the Eq. 5.62, and plugging in the Eq. 5.72, we obtain

$$\begin{aligned}\frac{1}{\Delta t}\,(n^{k+1}_j - n^k_j) = \frac{1}{A_j\,\Delta x_j}\Bigg(&\left[(1-\delta^{k+1}_j)\,B^k_{j+1/2} + \frac{1}{\Delta x_j}\,C^k_{j+1/2}\right] n^{k+1}_{j+1} - \\ - &\left[\frac{1}{\Delta x_j}\,(C^k_{j+1/2} + C^k_{j-1/2}) + (1-\delta^{k+1}_{j-1})\,B^k_{j-1/2} - \delta^{k+1}_j\,B^k_{j+1/2}\right] n^{k+1}_j + \\ &+ \left[\frac{1}{\Delta x_j}\,C^k_{j-1/2} - \delta^{k+1}_{j-1}\,B^k_{j-1/2}\right] n^{k+1}_{j-1}\Bigg)\end{aligned} \tag{5.76}$$

where $\Delta x_j = x_{j+1/2} - x_{j-1/2}$. If the x-grid is sufficiently fine, $\Delta x_j \approx (x_{j+1} - x_{j-1})/2$.

We apply this differencing equation to solve the Eq. 5.2. It includes also injection and escape terms, therefore we need to adapt the scheme. We add the injection and

the escape terms, discretized on the time and momentum grids, to the RHS of the Eq. 5.76. Also, in our case, $A(x) = 1$, and x stands for the Lorentz factor γ. As a result, we obtain the difference equation for solving the kinetic equation Eq. 5.2:

$$\frac{1}{\Delta t}\left(n_j^{k+1} - n_j^k\right) = \frac{1}{\Delta\gamma_j}\left(\left[(1-\delta_j^{k+1})\,B_{j+1/2}^k + \frac{1}{\Delta\gamma_j}\,C_{j+1/2}^k\right] n_{j+1}^{k+1} - \right.$$
$$- \left[\frac{1}{\Delta\gamma_j}\left(C_{j+1/2}^k + C_{j-1/2}^k\right) + (1-\delta_{j-1}^{k+1})\,B_{j-1/2}^k - \delta_j^{k+1}\,B_{j+1/2}^k\right] n_j^{k+1} +$$
$$\left. + \left[\frac{1}{\Delta\gamma_j}\,C_{j-1/2}^k - \delta_{j-1}^{k+1}\,B_{j-1/2}^k\right] n_{j-1}^{k+1}\right) + Q_j^k - \frac{n_j^{k+1}}{t_{\rm esc}} \quad (5.77)$$

The B and C coefficients, in our case, are

$$B(\gamma,t) = b_{\rm c}(\gamma,t)\,\gamma^2 - [a(t) + 2D_0(t)]\,\gamma \quad (5.78)$$
$$C(\gamma,t) = D_0(t)\,\gamma^2 \quad (5.79)$$

We employ a Lorentz factor grid with equidistant logarithmic step, as prescribed by e.g. Chiaberge & Ghisellini (1999):

$$\gamma_j = \gamma_{\rm min}\left(\frac{\gamma_{\rm max}}{\gamma_{\rm min}}\right)^{(j-1)\,/\,(J-1)} \quad (5.80)$$

where J is the number of points on the Lorentz factor grid.

The time grid we use, as previously discussed, has a constant time-step:

$$t_k = k\,\Delta t \quad (5.81)$$
$$\Delta t = \frac{t_{\rm end} - t_{\rm start}}{\mathcal{K}-1} \quad (5.82)$$

where $\mathcal{K}$ is the number of points on the time grid.

We multiply both sides of the Eq. 5.77 by Δt, and rewrite it in the following form

$$V1_j\,n_{j-1}^{k+1} + V2_j\,n_j^{k+1} + V3_j\,n_{j+1}^{k+1} = S_j^k \quad (5.83)$$

The coefficients $V1$, $V2$ and $V3$ are expressed as (performing some transformations)

$$V1_j = (-1)\,\frac{\Delta t}{\Delta\gamma_j}\,\frac{C_{j-1/2}}{\Delta\gamma_{j-1/2}}\,W^-_{j-1/2} \tag{5.84}$$

$$V2_j = 1 + \frac{\Delta t}{t_{\rm esc}} + \frac{\Delta t}{\Delta\gamma_j}\left[\frac{C_{j-1/2}}{\Delta\gamma_{j-1/2}}\,W^+_{j-1/2} + \frac{C_{j+1/2}}{\Delta\gamma_{j+1/2}}\,W^-_{j+1/2}\right] \tag{5.85}$$

$$V3_j = (-1)\,\frac{\Delta t}{\Delta\gamma_j}\,\frac{C_{j+1/2}}{\Delta\gamma_{j+1/2}}\,W^+_{j+1/2} \tag{5.86}$$

where

$$W^{\pm}_{j\pm1/2} = \frac{w_{j\pm1/2}\,\cdot\,\exp(\pm\,w_{j\pm1/2}\,/\,2)}{2\sinh(w_{j\pm1/2}\,/\,2)} \tag{5.87}$$

$$w_{j\pm1/2} = \frac{B_{j\pm1/2}}{C_{j\pm1/2}}\,\Delta\gamma_{j\pm1/2} \tag{5.88}$$

$$\Delta\gamma_{j\pm1/2} = \gamma_{j\pm1/2+1/2} - \gamma_{j\pm1/2-1/2} \tag{5.89}$$

Here we omitted upper indices (on the time grid) for B and C coefficients: e.g. $B_{j+1/2} = B^k_{j+1/2}$.

The term S on the RHS is

$$S^k_j = n^k_j + Q^k_j\,\Delta t \tag{5.90}$$

This formulation in terms of $W^{\pm}_{j\pm1/2}$ was introduced by Park & Petrosian (1996). This way of expressing the V coefficients is convenient to explore the importance of the Fermi-II acceleration process relative to the cooling or shock acceleration when performing computations: in case $|w| \gg 1$, the stochastic acceleration is much weaker than cooling or shock acceleration.

In fact, Eq. 5.83, represents a system of equations. These equations are coupled, i.e. it is not possible to solve each individual equation separately from the other ones, instead, the equations have to be solved all together. The system of equations 5.83 represents a tridiagonal matrix, and to solve it, we apply the numerical algorithm by Press et al. (1989). This algorithm is available in the standard Python mathematical package *NumPy*, through the method `np.linalg.solve(A,b)`, where A and b are the matrix and free-term vector respectively (in a system $A \cdot \mathbf{x} = \mathbf{b}$).

One also should not forget about the boundary conditions, that are imposed to ensure particle number conservation. As shown in Park & Petrosian (1996), the boundary conditions (Eq. 5.66) are enforced in the numerical scheme by setting

$$V1_0 = 0 \tag{5.91}$$
$$V3_{J-1} = 0 \tag{5.92}$$

For the case of the kinetic equation (Eq. 5.2) without the Fermi-II acceleration term, when the stochastic acceleration is not present or deactivated, the equation is now first order only. One does not really need to re-derive the entire scheme for this particular case, it is possible to simply take a limit of $t_{\rm FII} \to \infty$ (or $D_0 \to 0$) in the system 5.84. This particular case of the Chang and Cooper scheme (for only injection, escape and cooling processes) is described in Chiaberge & Ghisellini (1999). The V coefficients in this case are (not neglecting the shock acceleration though)

$$V1_j = 0 \tag{5.93}$$
$$V2_j = 1 + \frac{\Delta t}{t_{\rm esc}} + \frac{\Delta t\, B^*_{j-1/2}}{\Delta\gamma_j} \tag{5.94}$$
$$V3_j = (-1)\,\frac{\Delta t\, B^*_{j+1/2}}{\Delta\gamma_j} \tag{5.95}$$

where $B^*(\gamma, t) = b_{\rm c}(\gamma, t)\,\gamma^2 - a(t)\,\gamma$

Since the differential equation is first-order, there is only one boundary condition that has to be imposed

$$V3_{J-1} = 0 \tag{5.96}$$

This condition means no particle flux through the maximal Lorentz factor boundary, and is equivalent to the condition that at high Lorentz factors the electron spectrum should fall to zero.

This numerical scheme was implemented in the EMBLEM code. As a basis, we took a publicly available Python code by C. Nigro[2], who implemented a particular

[2]The code was previously publicly available on GitHub platform, however the author has removed his code from public domain in 2019.

case of the scheme for the first order equation with only injection, escape and cooling terms (differencing scheme presented in Chiaberge & Ghisellini (1999)). We have extended this code by adding shock and stochastic acceleration, as well as including treatment of the inverse Compton cooling.

The part of the code computing the evolution of the electron spectrum was verified using test equations from Park & Petrosian (1996), as well as by comparing to analytical solutions for simple cases of the kinetic equation.

Emission evolution

The SED of the emission from the blob is calculated for the underlying electron spectrum at each time step. The computation is performed according to the leptonic SSC scenario presented in Chapter 4 (sub-section 4.1.1). The synchrotron emission is calculated using the expressions from sub-section 4.1.3, and the IC emission is computed following sub-section 4.1.4. Finally, we transform the emission from the blob reference frame to the observer's frame following sub-section 4.1.6. We adopt a value for the Hubble constant of $H_0 = 70$ km s^{-1} Mpc^{-1}. The absorption of γ-rays in the intergalactic medium due to their interaction with the Extragalactic Background Light (EBL) is taken into account, with the use of a publicly available module [3], including several different EBL models (optical depth depending on the γ-ray energy and the redshift). We compare these models by simulating an example VHE γ-ray spectrum typical for HBLs, and absorbing it according to EBL models available in the module. The difference in absorption effect induced by different EBL models is depicted in Fig. 5.4. For the current application, we use the EBL model by Domínguez et al. (2011).

The light curves are calculated by integrating over the time-dependent emission in the energy range of interest via a separate script.

The description of the radiative emission was cross-checked with the output of the code by Katarzyński et al. (2001), and shows a good agreement with it.

5.2.2 Input parameters

A parameters file is used to specify the input parameters of a single simulation run. The code takes the following input parameters:

[3]Developed by M. Meyer, https://github.com/me-manu/ebltable

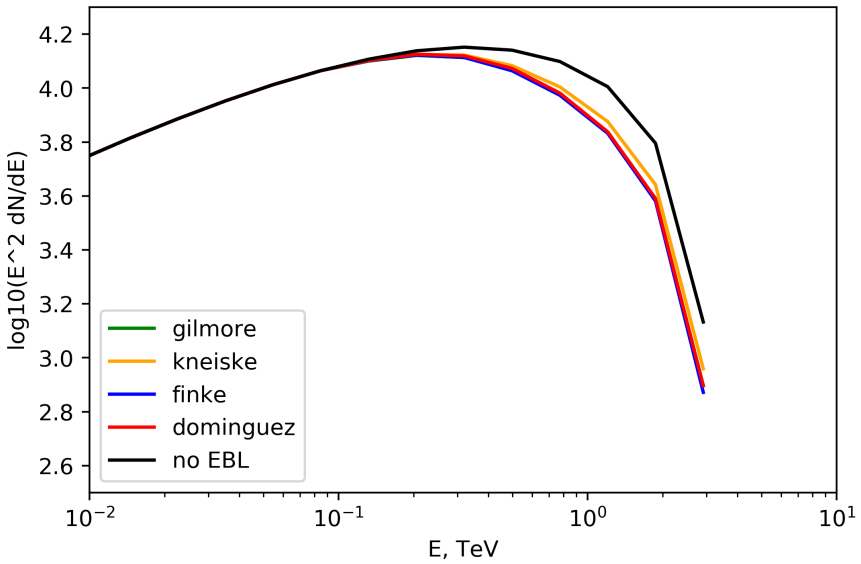


Figure 5.4: Comparison of different models of EBL absorption in the module by M. Meyer. Black curve represents a simulated example intrinsic source's spectrum in the VHE γ-ray band, colored curves display the spectrum affected by the absorption on EBL described with a respective model.

- Physical parameters of the emitting blob
  - B (magnetic field)
  - $\gamma_{\mathrm{min,inj}}$ (minimal Lorentz factor of injected electrons)
  - R_{b} (radius of the blob)
  - δ_{b} (Doppler factor of the blob)
  - z (redshift of the source)

- Evolution parameters
  - t_{inj} (duration of particle injection)
  - t_{esc} (time-scale of particle escape)
  - $Q_{\mathrm{inj}}(\gamma, t)$ (injection function/spectrum)
  - $t_{\mathrm{FII}} = 1/D_0$ (time-scale of stochastic acceleration)
  - $t_{\mathrm{FI}} = 1/a$ (time-scale of shock acceleration)
  - $t_{\mathrm{dur,FI}}$ (duration of Fermi-I acceleration phase)
  - $t_{\mathrm{dur,FII}}$ (duration of Fermi-II acceleration phase)

- SED parameters

  - EBL model name (Dominguez/Finke/Inuoe/Gilmore/Kneiske or "none")

- Grids parameters
  - J (number of points on the Lorentz factor grid)
  - $\mathcal{K}$ (number of points on the time grid)
  - $N_{\mathrm{p,SED}}$ (number of points on the photon energies grid used for calculation of the SED)
  - Δt_{out} (time step for the SED and LC computation)
  - $\gamma_{\mathrm{gg,min/max}}$ (minimum and maximum Lorentz factors in the Lorentz factor grid)
  - $E_{\mathrm{ph,min/max}}$ (minimum and maximum photon energies for computation of the SED (in the observer's frame))

The most common injection function would be a power law in γ. A slightly more complex case is a power law with an exponential cutoff, parametrized with normalization A_{inj}, slope α_{inj} and cutoff Lorentz factor γ_{cut}:

$$Q_{\mathrm{inj}}(\gamma) = A_{\mathrm{inj}}\,\gamma^{-\alpha_{\mathrm{inj}}}\exp(-\gamma/\gamma_{\mathrm{cut}})$$

The injection term could be also an arbitrary function in Lorentz factor and time, as long as it has a physical meaning; this can be used to implement different injection scenarios.

The time behavior of the acceleration processes in the code can be controlled: it is possible to activate acceleration of particles lasting for only a certain period of time $t_{\mathrm{dur,FI/FII}}$. An arbitrary parametrization of time-dependent acceleration process could be also implemented, such flexibility allows to model and test various acceleration scenarios.

These input "parameters" are flags allowing to enable/disable certain processes:

- Enable/disable Fermi-I or Fermi-II acceleration process ("yes"/"no")
- Take/not take into account IC cooling ("yes"/"no")

If one estimates that the IC cooling is negligible, it is possible to deactivate it in order to speed up the code. For HBLs in the stationary state this effect is typically minor, however during flaring states it can be rather important.

If the Fermi-II acceleration is deactivated using the respective flag in the parameter file, or if the Fermi-II process at a given time step is found to be "weak", the code

calculates the electron spectrum evolution according to the first-order expressions in the Chang & Cooper (1970) scheme, given by equations from Eq. 5.93 to Eq. 5.96. The condition for "weak" Fermi-II acceleration process is set in the code as following: the time-scale of Fermi-II acceleration at a given time step should be longer than the escape time-scale by a factor of 100 or more, $t_{\rm FII}(t_k) \geq 100\, t_{\rm esc}(t_k)$. This condition means that the Fermi-II process will not have time to significantly change energies of particles over the time they reside in the blob, and therefore will not appreciably affect the spectrum of particle population. Imposing this condition allows to speed up the code and avoid overflowing exponential terms in Eq. 5.87.

5.2.3 Code architecture

The EMBLEM code's hierarchy in a most general case can be represented as illustrated in Fig. 5.5:

A single simulation run is performed as follows:

- Specify input parameters in a file
- Run the code with the parameter file (as a command line argument):
  python emblem.py input_parameters.py
- Grids are initialized
- Electron spectrum is computed at each time step → output to files
- The "history" of the electron spectrum serves as an input for the SED computation procedure
- SED is computed at a time step $\Delta t_{\rm out}$, and then is converted into the observer's frame and EBL absorbed (if selected)
- The SED evolution is written to the output files
- To compute the light curves, run a separate script, indicating the input parameters file, energy range of interest, and whether the output flux is in photons or in ergs (as command line arguments), e.g.:
  `> python comp_lc.py input_parameters.py 1.0e12 10.0e12 photons`
- The electron spectrum (in the frame of the blob) and the SED evolution (in the observer's frame), as well as the light curves (in the observer's frame), will be located in the output folder

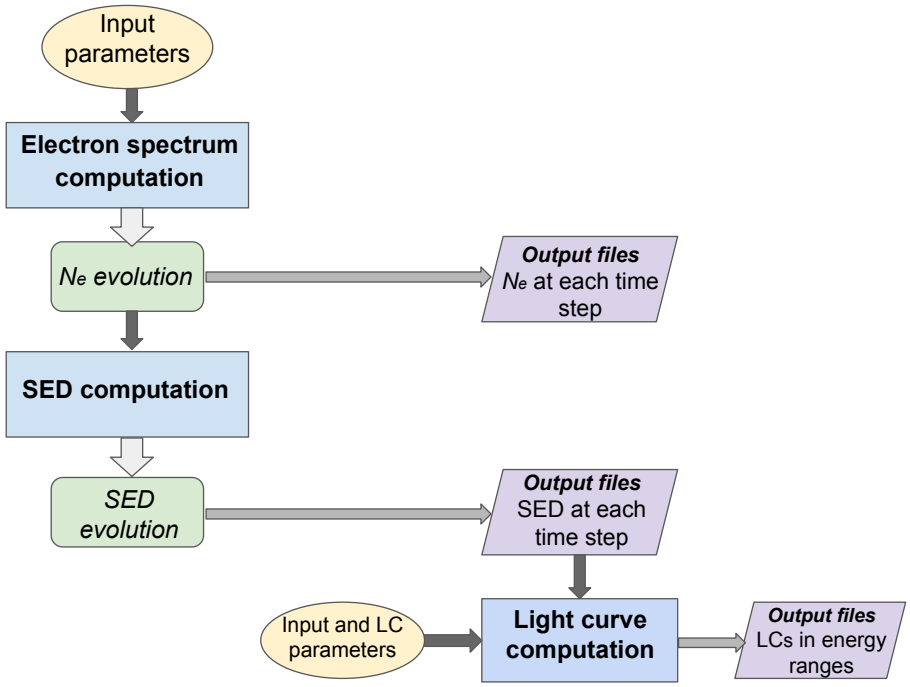


Figure 5.5: Scheme representing the architecture of the EMBLEM code

5.2.4 Examples of simulated flares

In order to demonstrate the domain of the code application, we perform simulations of flaring events and *explore the parameter space*. Such a study illustrates also the way key parameters of the code influence the SED and the light curves.

We simulate flaring events, according to two simple scenarios:

1. Electrons are injected *continuously* into the blob, cool and escape the emitting zone and undergo Fermi-II acceleration process during a certain time interval
2. Electrons are injected *instantaneously* into the blob at $t = 0$, and undergo continuous Fermi-I and Fermi-II acceleration processes acting at the same time (together with cooling and escape).

In the first case, the flare is triggered by an intervening Fermi-II acceleration process, acting during a certain time period. During the acceleration phase, the physical parameters of the blob remain unaltered. The rise and fall of the light curve is due to the competition between the gain processes (injection and acceleration), and loss processes (cooling and escape). In the second case, the origin of the flare is the injection pulse, magnified by persistent Fermi-I and Fermi-II acceleration processes. The

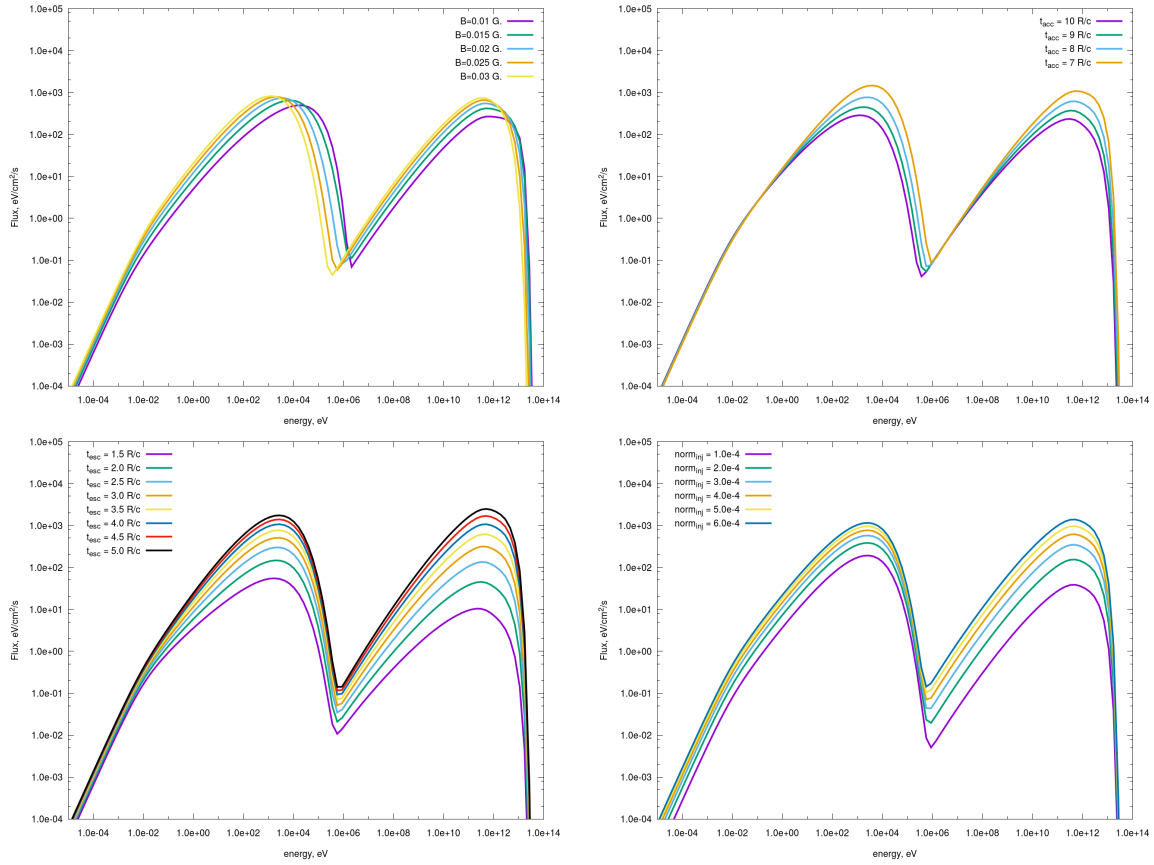


Figure 5.6: Effect of the magnetic field (top-left), Fermi-II time-scale (top right), escape time-scale (bottom-left) and injection spectrum normalization (bottom-right) on the peak SED during the flare. A flare is triggered by Fermi-II acceleration process lasting for a limited time interval, during a continuous electron injection phase.

subsequent flux evolution is determined by the competition between the acceleration, and losses – cooling and escape. In both cases we neglect the IC cooling effect and use a power law with a cutoff as an injection spectrum. We re-run the simulation for both scenarios trying different values of a studied parameter, scanning over a certain domain, keeping all the other parameters unchanged. We present the peak SED during the flare produced according to the scenario 1 in Fig. 5.6, and the VHE light curve profiles (flux in the range 1-10 TeV), generated according to the scenario 2 in Fig. 5.7. These figures demonstrate the effect of the studied parameters on the SED and the light curve shape.

From the upper-left panel of the Fig. 5.6 one can see that a higher magnetic field shifts the synchrotron peak to lower frequencies. This may appear contrary to intuition, as for one electron, peak energy of the spectrum of its synchrotron emission

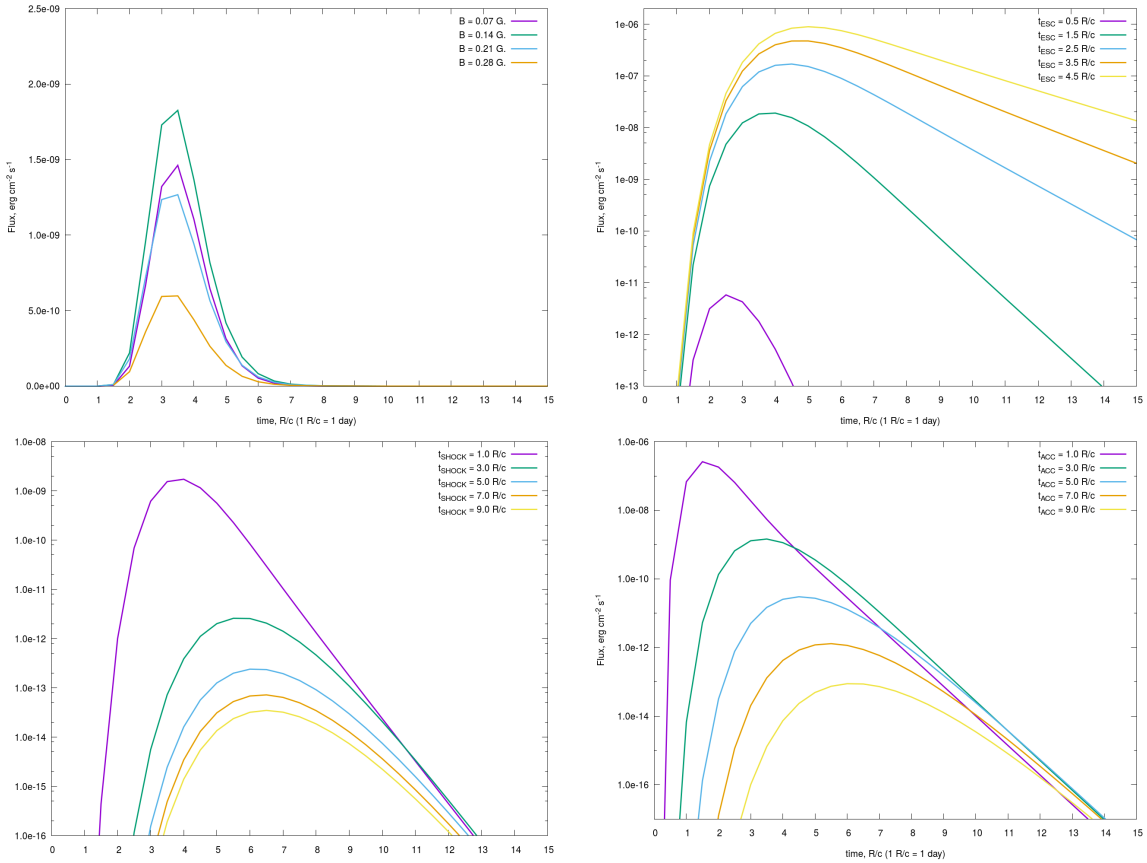


Figure 5.7: Effect of the magnetic field (upper-left), escape time-scale (upper-right), Fermi-I acceleration time-scale (bottom-left) and Fermi-II acceleration time-scale (bottom-right) on the profile of the light curve in VHE band (1-10 TeV). A flare is triggered by an injection pulse followed by a continuous acceleration phase.

is proportional to the magnetic field. However, we are dealing here with a time-dependent phenomenon: the electrons undergo cooling during their evolution in the emitting zone. The Lorentz factor of the cooling break is determined by the balance between the cooling rate (Eq. 5.4) and the acceleration or escape rate:

$$t_{\rm cool} = min(t_{\rm esc}\,,\,t_{\rm acc}) \;\rightarrow\; \frac{1}{b_{\rm c}\,\gamma} = min(t_{\rm esc}\,,\,t_{\rm acc}) \;\rightarrow\; \gamma_{\rm br} \propto \frac{1}{B^2} \tag{5.97}$$

Very roughly, the peak energy of the SED will be

$$\epsilon_{\rm peak} \propto B\,\gamma_{\rm br}^2 \propto \frac{1}{B^3} \tag{5.98}$$

It is however important to note, that the Eq. 5.98 is valid only in a specific

regime, when the peak frequency of the synchrotron bump is determined by the break Lorentz factor (and not the one of the high-energy cutoff), $\gamma_{\rm br} < \gamma_{\rm cut}$. Therefore, for flaring states in this regime, the SED peak indeed shifts to lower energies for a higher magnetic field. Also, this plot reveals that in the optical-to-X-ray domain of the SED (strictly speaking, the power law part of the synchrotron bump), the flux augments with higher magnetic field (as expected), which is not the case in the hard X-ray (strictly speaking, "after-peak" falling part of the synchrotron bump), where the opposite behavior is seen. This happens due to stronger radiative cooling. The same trend is observed in the IC part of the SED, due to the same reasons. However, the shift of the IC peak to lower energies, and the earlier cutoff, are not that pronounced due to Klein-Nishina effects.

From the upper-right panel of the Fig. 5.6 it is visible that, as expected, shorter acceleration time-scales cause harder particle distribution and photon spectra. The bottom-left panel of the Fig. 5.6 shows the effect of the escape time-scale on the SED. Longer particle escape time not only results in accumulating more electrons in the emitting zone and amplification of the spectrum normalization, but also causes harder spectra because the time during which particle undergoes acceleration in the emitting zone is increased (escape time-scale is roughly, on average, the time particle spends in the blob). The bottom-right panel of the Fig. 5.6 displays simple linear rise of the synchrotron flux with increasing injection spectrum normalization and quadratic rise of the IC flux.

The most interesting behavior of the light curve profile is in the upper-left panel of the Fig. 5.7. One can see that, at first, as the magnetic field augments, the total VHE flux increases, however after a certain critical value, the VHE flux starts to decrease with further growing magnetic field. Such transition occurs due to competition between enhancing IC peak flux as the magnetic field increases, and the stronger cooling, which displaces the IC peak position to lower frequencies and causes the cut-off in the VHE part of the spectrum to happen earlier (in energy). The upper-right panel of the Fig. 5.7 demonstrates that the longer is the escape time, the slower is the fall of the light curve. The bottom-left panel of the Fig. 5.7 illustrates that shorter Fermi-I acceleration time-scales produce sharper rising part of the light curve profile, as low-energy particles faster "migrate" to high Lorentz factors. The VHE flux increases with shorter acceleration time-scale because the peak spectra are harder. The same behavior is seen in the bottom-right panel of the Fig. 5.7, where the Fermi-II time-scale effect is depicted. One could also notice in the bottom panel, that the flux decay is becoming flatter with increasing Fermi-I or Fermi-II time-scale. This occurs due to the fact that, with longer acceleration time-scales, there are more particles that are still in the process of being accelerated, i.e. still "flow" to high energies during

the fall of the flux where losses due to radiative cooling and escape dominate; such higher inflow to high Lorentz factors mitigates the losses and leads to a slower fall of the light curve.

5.3 Physical modeling of AGN flares

In the previous sections, we briefly mentioned a couple of possible processes that may trigger the flares. These processes were considered as isolated; so far we did not try to construct a general picture, in which these processes arise as a consequence of some other, more complex physical processes occurring in the jet environment of the emitting blob. Such processes are linked to (still not well understood) properties of the jet, or even the entire AGN system. In this section, we give a review of various physical scenarios proposed to explain the blazar flaring activity.

5.3.1 What causes the observed variability in the VHE γ-ray regime ?

Variability in all wavebands is one of the key properties of AGN. The time-scales of flux variations observed in various sources are ranging from as short as few minutes to as long as few months and even years. The most rapid variability is observed in the VHE γ-ray band. Most likely, the nature of variability and physical processes involved are different for different time-scales. One can differentiate three typical time-scales of flux variations:

- **A few years – 1 month**: The variability on these time-scales at VHE is very poorly studied due to observational possibilities of currently operating Cherenkov telescopes. The origin of such flux variations can be jet precession (e.g. Abraham & Romero (1999)), jet inhomogeneties (e.g. Li et al. (2018)), binary supermassive black holes (e.g. Villata et al. (1998)), ablation of a gas cloud entering into the jet (Zacharias et al. 2017), etc.

- **1 week – 1 day**: Flares occurring on these time-scales are due to the dynamics, or macro- or micro-physics of the emitting region: a geometrical effect (e.g. Abdo et al. (2010a) ; Casadio et al. (2015)), or perturbing physical processes within the emitting zone or its vicinity (e.g. Marscher & Gear (1985) ; Chiaberge & Ghisellini (1999)).

- **1 hour – 1 minute**: The fastest variability discovered at VHE is quite challenging to explain with standard emission models since these time-scales are shorter than the light-crossing time of the SMBH. Different sophisticated scenarios have been developed to model such outbursts, possible options include turbulent cells (e.g. Marscher (2014)), magnetic reconnection inside the jet (e.g. de Gouveia Dal Pino et al. (2010)) or near its base (e.g. Kadowaki et al. (2015) ; Khiali et al. (2015)), pulsar-like acceleration in the SMBH magnetosphere (e.g. Levinson & Rieger (2011)), a red giant star crossing the jet close to the SMBH (Barkov et al. 2012), etc.

There are two broad classes of models of the origin of flaring emission. In the *first class*, flares originate from the region *in the vicinity of the central engine*. That comprises accretion disk instabilities (e.g. Czerny (2006)), acceleration in vacuum gaps in the BH magnetosphere (e.g. Ptitsyna & Neronov (2016) ; Levinson & Rieger (2011)), magnetic reconnection events at the base of the jet (e.g. Khiali et al. (2015)), etc.

In the *second class*, flares are initiated *inside the relativistic jet*, in the vicinity of the emission zone. To explain the emergence of flares within the relativistic jet, one generally distinguishes two possible types of scenarios. *In the first type*, variations of macrophysical properties of the emitting region, like its global geometry and kinematics, are responsible for launching outbursts. For instance, the relativistic Doppler factor can be increased by a change of viewing angle or bulk Lorentz factor, which can lead to stronger emission boosting and launch flares (e.g. Casadio et al. (2015) ; Larionov et al. (2016) ; Raiteri et al. (2017)). This happens if the emitting region is moving along a curved trajectory, due to jet bending (Abdo et al. 2010a) or its helical configuration (Villata & Raiteri 1999). *In the second type*, the observed flux variability is considered to be due to the micro-physics inside the VHE emitting zone and the subsequent evolution of the population of radiating particles caused by various physical processes, e.g. enhanced injection (Mastichiadis & Kirk 1997), and different particle acceleration mechanisms due to the development of shocks (e.g. Marscher & Gear (1985) ; Sikora et al. (2001)), turbulence (e.g. Tammi & Duffy (2009) ; Tramacere et al. (2011)) or magnetic reconnection (Giannios et al. 2009).

Emission models often combine the two types of mechanisms in order to reproduce extreme flares. E.g. Marscher (2014) developed a physical scenario, in which a turbulent flow of relativistic plasma is traversing a conical recollimation shock, which compresses the plasma and accelerates particles to high energies. Hence, both particle acceleration and modification of the emitting region geometry (its squeezing) are invoked. Occurrence of a quite similar physical situation might explain for instance

the strong 1997 outburst of Mrk 501 the properties of which suggested a decrease of the size of the emitting region, an enhancement of the particle density, and an increase of the Lorentz factor of the break in the electron spectrum (Katarzyński et al. 2001). This effectively corresponds to complex scenarios in which the emitting zone is compressed by environmental forces, and the particle population gains energy by an accompanying acceleration process.

In my study, for the reasons explained below, I focus on scenarios in which the flaring emission originates from *inside the jet* (rather than from a zone in the vicinity of the SMBH), either from the blob itself or from the jet environment in its proximity. This choice is justified as follows. As the flare I model (Mrk 421 February 2010 flare, see Section 6.3) is the brightest up to now in the VHE γ-ray range, it is clear that in order to reach such extreme VHE flux one needs (*i*) strong Doppler boosting and (*ii*) low opacity for VHE γ-rays escaping the source. The former effect is typical for AGN jets, and the latter condition is very difficult to achieve for VHE γ-rays escaping from a compact region close to the SMBH. Therefore, the observational properties of the flare under study are the most easily accommodated in a scenario in which the flaring emission production site is located inside the jet. Finally, such scenarios are up to now the most developed and considered in the literature, and are often favored, although a number of questions still remain open.

In addition, since the studied flare occurs on a time-scale of a few days, I explore the approach based on *evolution of particle distribution*, assuming constant physical parameters of the emitting zone (like its size and magnetic field) and propose time-dependent SSC scenarios to try to connect self-consistently the long-term low-state emission with the flaring one. This appears to be a powerful way to drastically reduce the number of free parameters of the flaring event, and also provides a direct explanation to the fact that active AGN states often show SED evolving continuously from the quiescent ones.

Flare scenarios of the considered type may involve only one active zone inside the jet or multiple, often connected active zones.

5.3.2 One-zone models

The most basic model for AGN VHE flaring emission with a minimal number of parameters is the one-zone model. In this model, the short-term flaring emission is originating from the blob, also responsible for production of the steady-state emission. The flare is caused by a non-destructive perturbation of the emission region (and of its steady state). The possible perturbations were discussed in the previous sub-section,

the most relevant ones seem to be, in the case of rather fast flares, an *injection pulse*, or *acceleration processes*, acting directly in the emitting blob. The short-term electron injection beam may originate, for instance, from the base of the jet. The blob can be also disturbed by an isolated shock traveling down the jet and traversing the emitting region, or by crossing a standing shock with a diamond structure. Another option is spontaneous development (and subsequent decay) of turbulence inside the emitting zone due to various causes (instabilities). In the two last cases, the particle population in the blob is perturbed by an intervening Fermi-I or Fermi-II acceleration process, which initiates the flare due to distortion (hardening and/or normalization increase) of the steady-state electron spectrum. As both shocks and turbulence are ubiquitous in AGN jets, the one-zone scenario appears to be the most natural way to produce a flare if its origin is attributed to the microscopic properties of the emitting zone.

5.3.3 Multiple-zone models

Due to the growing complexity of the observed data, and a "multi-layer" character of the processes that are thought to launch blazar outbursts, one could expect to have a more complex geometric configuration in the physical scenarios invoked to explain flares.

A two-zone model is becoming more and more widely discussed option in the literature. For instance, the injection "flash" discussed before, may not originate from a very distant source (like the base of the jet), but from the vicinity of the emitting zone: e.g. from an acceleration region suddenly forming near the blob. Particles in this region may be accelerated by e.g. Fermi-II mechanism if the region is turbulent, or by shocks or magnetic reconnection. One of the first scenario of this kind was proposed by Kirk et al. (1998), with a shock acceleration region located in the vicinity of the emitting zone. Furthermore, a two-zone configuration seems to be a natural choice to explain rapid flares on top of the slowly varying steady-state flux, if the flaring emission is produced in a smaller region than the low-state one.

Another example of a two-zone model is the configuration in which the flaring emission is emanating from another zone, than the quiescent emission. Such "flaring blob" may suddenly appear due to various instabilities in the large-scale jet, turbulence arising around the emitting blob, two blobs catching up, etc. There are even two-zone models, in which the steady-state emission is originating from two emission regions of different size, and a flaring event is caused by e.g. turbulence arising in one of the zones (e.g. Cao & Wang (2013)).

Finally, some multiple-zone models assume that the observed quiescent blazar

emission is coming from the entire jet, and the observed variability is due to various physical processes acting at small-scales. For instance, in the turbulent multi-zone model by Marscher (2014), the total emission is the sum of fluxes from a large number of individual small cells, comprising the large-scale blazar jet. In this model, the noise-like fluctuations of the flux at short time-scales are ascribed to turbulence at small spatial scales.

In the following, we consider a remarkable flare of the archetypal BL Lac Mrk 421. We first apply the simplest scenario in order to limit the number of free parameters, but allow new levels of complexity step by step when observational constraints impose it. We therefore start with the modeling of the quiescent state by a quasi-stationary one-zone SSC scenario and then perturb it to give rise to the active state.

Chapter 6

Modeling of a MWL flare of Mrk 421

Having developed the EMBLEM code, we are now going to apply it to a real data set of a flare, with the goal to get an insight about physical processes and change of physical conditions during flaring activity[1]. Revealing the processes that trigger and damp the flare, as well as understanding the time evolution of the system during the outburst, allows to construct a physical scenario of the flaring behavior. With the EMBLEM code we are able to test various scenarios of flaring activity for a given flare data set, in which the outburst is initiated via an injection pulse and/or acceleration processes acting within or in the vicinity of the emission region. We focus specifically on these causes of flaring activity, as they are among the most common ones assumed for the origin of blazar flares. Clearly, to achieve minimal ambiguity in interpreting the observational data and to ensure that the physical model found to explain the data set is realistic and justified, one needs a flare data set with as good MWL and time coverage as possible, i.e. measurements of the timing and spectral properties of the varying signal, in a form of spectra in different flux states and light curves in different wavebands. This is because more measurements place more constraints on the model.

Markarian 421 (Mrk 421), a BL Lac object, being (generally) the brightest extra-

[1]Partial results of the presented work have been (1) submitted for a publication "Connecting steady emission and the 2010 Very High Energy flaring state in the blazar Mrk 421" by A. Dmytriiev, H. Sol and A. Zech to MNRAS (September 2020), and (2) already published in the proceedings of the 36th International Cosmic Ray Conference (ICRC2019) in a contribution "Time Dependent Modeling of Electron Acceleration and Cooling During Blazar Flares" by A. Dmytriiev, H. Sol and A. Zech (July 2019) (Dmytriiev et al. 2019a).

galactic source in the VHE γ-ray band, offers one of the most high-quality MWL flare data sets. This source is also one of the best studied blazars (more information on the source in Section 6.1). In February 2010, Mrk 421 showed a particularly strong MWL flare, which is *the brightest ever observed* in the VHE γ-ray range up to now. The outburst was monitored by various instruments across the globe, from radio frequencies to VHE γ-rays. The resulting data set (presented in Section 6.3) features one of the most complete measurements of varying source flux in different energy ranges (MWL light curves). The spectral coverage is quite decent as well, with rather good-quality spectral measurements at or around the flare peak in optical, X-ray and gamma-ray bands. Thus, we have chosen to study this specific flare because of its extreme nature and a particularly rich MWL coverage, with a goal to reveal the physical processes that led to the exceptionally high VHE flux. In addition, because of a very good coverage of the timing characteristics of the source flux during this outburst, we put special emphasis on describing shapes of the MWL light curves. Such effort, in fact, is one of relatively first attempts to reproduce the MWL light curve profiles by means of physical modeling.

We stick to an approach, in which we consider that the blazar flaring activity is due to a perturbation of the steady state of the VHE source. This assumption means that the flaring emission is not originating from an arbitrary location in the jet and is not just superimposed on the quiescent one. Instead, this assumption implies that the flaring emission comes directly from the quiescent VHE γ-ray emitting zone, or from a zone *physically connected* to it. Such assumption imposes important additional constraints on the physical model and results in reduction in the number of free parameters.

As already pointed out in 5.3.1, the nature of the disturbance causing flaring behavior is widely discussed. Typically, when considering a single perturbation arising in the system, the most common scenarios include an injection pulse or, more generally, varying injection rate (e.g. Mastichiadis & Kirk (1997)), a shock crossing the emitting zone (e.g. Marscher & Gear (1985) ; Sikora et al. (2001)), various instabilities in the medium of the jet (Meliani & Keppens (2009), Bromberg & Tchekhovskoy (2016)), or star-jet interaction (Barkov et al. 2012), etc. In our work, we will focus our attention on physical scenarios described in 5.3.2 and 5.3.3, assuming that particle acceleration processes (shock and/or stochastic mechanism) acting in an intermittent mode, represent the perturbation responsible to trigger the flare.

We aim at *connecting* self-consistently the long-term low-state emission to the flaring one, within one coherent scenario, using the EMBLEM code and analytical methods. In order to explain the flare within this framework, we first model the quiescent state of the source (Section 6.2), and then introduce a disturbance in the

description of the steady state of the system and try to reproduce the MWL flare data set. With this approach, we first derive a general analytical criterion to test one-zone flare models with a passing shock (Section 6.4), and then proceed to physical modeling of Mrk 421 flare data set (Section 6.5). First, one-zone models are explored (sub-section 6.5.1), and after concluding their inability to describe the data, we focus on two-zone models and, as a result, find a reasonable solution explaining the Mrk 421 flaring behavior (sub-section 6.5.2). Finally, in Section 6.6 we discuss the results and various implications.

6.1 The studied source: Mrk 421

We focus on one of the most comprehensively studied blazars, Mrk 421. It shows one of the most complete data sets for both different flaring states and for the low state, thanks to numerous multi-wavelength campaigns on this source.

Markarian 421 (Mrk 421 or Mkn 421 ; RA = 11h 04m 27.3s , DEC = $+38^\circ\,12'\,31.8''$ in J2000) is a high-synchrotron-peaked BL Lac (HBL), with a synchrotron peak in the X-ray band. It is generally the brightest extragalactic TeV emitting source, and also the closest TeV blazar to the Earth (redshift $z = 0.0308$). This object was actually the first extragalactic source detected in the VHE γ-ray sky (Punch et al. 1992).

The source is strongly variable across the entire electromagnetic spectrum. Ahnen et al. (2016) examined multi-band variability of Mrk 421 on a time-scale of 1 year. The authors find that in the VHE γ-ray regime (observations with MAGIC), the flux level from the object varies by a factor as high as $\sim 20-30$ over 2.3 years period, with a minimal flux above 400 GeV of $\sim 10^{-11}$ ph cm^{-2} s^{-1}, and a maximal flux in the same range of $\sim 3 \times 10^{-10}$ ph cm^{-2} s^{-1}. A similar variability pattern (low state and flaring emission) was identified in the X-ray light curves, well-correlated to the one in VHE γ-rays. Such correlation is typically explained by a synchrotron self-Compton (SSC) emission scenario. The variability in the optical band appeared to be more modest, with a very weak or almost no correlation with emission variations in higher energy bands. Based on the ultra-long-term light curve of Mrk 421 in the optical B-band, spanning almost 100 years, Liu et al. (1997) revealed two kinds of variability behaviors: non-periodic fast and strong flux variations on time-scales ranging from hours to days, and periodic variations with a period of 23.1 ± 1.1 years. The latter behavior was ascribed by the authors to accretion disk instabilities. In the 1-10 GeV energy band (long-term observations by *Fermi*-LAT), the source displays variability in a stochastic manner (Acciari et al. 2014); the same behavior was also reported by *Whipple* in the 0.1-1 TeV range based on its long-term flux measurements. The

flux variability strength for Mrk 421, on the whole, tends to enhance with increasing energy (Ahnen et al. 2016). The source is also one of the most rapidly varying TeV emitters, showing flares on time-scales as short as 30 min (Gaidos et al. 1996).

Multiple observations and subsequent studies of prominent individual outbursts of Mrk 421 were done. For example, Whipple observations of a flare in 1997 (McEnery et al. 1997) with a peak flux above 500 GeV around 10 Crab units; exceptional long-lasting high state during January – May 2001, reported by Krennrich et al. (2002), characterized by variability of spectral index between 1.89 (high state) and 2.72 (low state); an outburst in June 2008 (Pittori et al. (2008) ; Cao & Wang (2013)) during which no significant correlation between flux behavior in the optical and X-ray or VHE γ-ray bands was found. The source underwent an extreme flare in February 2010 (Shukla et al. (2012); Singh et al. (2015)) which is, up to now, the brightest known AGN flare in the VHE γ-ray range, as well as one of the strongest in hard X-rays. Other major outbursts were detected in March 2010 and in April 2013, and intensively studied afterwards (March 2010 flare: Aleksić et al. (2015) , April 2013 flare: Paliya et al. (2015a); Sinha et al. (2015) ; MAGIC collaboration et al. (2020)). Since then, Mrk 421 showed a few of other VHE flares, detected by various IACTs, however no profound studies of these events were carried out, due to absence of coordinated MWL campaigns on the source during the high states.

Overall, during the flaring activity, the source spectrum in both X-rays and γ-rays is hardening with an increasing flux – so-called "harder-when-brighter" trend (Aharonian et al. (2002); Albert et al. (2007a); Fossati et al. (2008); Acciari et al. (2011b)).

A few TeV flares detected from Mrk 421, did not display a counterpart in the X-ray band, or in other lower energy bands (so-called "orphan" flares, e.g. Błażejowski et al. (2005); Fraija et al. (2015)). These outbursts challenge the existing emission models, especially the standard one-zone SSC scenario. For instance, in the study by Sahu et al. (2016), authors invoke a hadronic model and explain an "orphan" TeV flare that occurred in April 2004 by interaction of Fermi-accelerated multi-TeV protons with the low-energy background SSC photons in the compact flaring region of the jet, having enhanced photon density.

Albert et al. (2007a) conducted an investigation of the properties of Mrk 421 emission, based on MAGIC data taken during an, on average, relatively low flux state of the source, with flux variations up to a factor of 4. The authors revealed that the SED of Mrk 421 above 100 GeV deviates from a power law, exhibiting a curved shape, remaining even after correction for the EBL. The study also confirmed the correlation between X-ray and VHE γ-ray signal, as well as the absence of clear cor-

relation between optical and γ-ray intensities. Another important conclusion made by the authors was that the different flaring states are due to variations of the underlying spectrum of the electron population, rather than due to a change of physical parameters of the emitting zone.

6.2 Modeling of the low state of the source

As previously discussed, to describe flares in our approach, we first have to develop a physical model of the quiescent state of the source, i.e. find a scenario for production of emission in unperturbed source state, which includes establishing the physical conditions in the emitting region. The constructed model will serve as a "baseline": by perturbing the low-state configuration, a flaring state will be achieved.

6.2.1 Mrk 421 low-state emission: archival data

For our modeling of the steady state of the source, we use the measurements of the quiescent SED of Mrk 421 presented in Abdo et al. (2011). In this paper, authors report the results of the 4.5-month-long observational MWL campaign on the source during its quiescence. The dedicated observations were performed from January 19 to June 1, 2009, when Mrk 421 showed a quite low level of flux, in particular, one of the lowest detected by MAGIC, and virtually no activity and/or variability across the entire electromagnetic spectrum. During the campaign, different instruments were monitoring the source, including the VLBA at $\sim$GHz frequencies, radio telescopes from the F-GAMMA program, near-infrared and optical telescopes from the GASP-WEBT program, *Swift* and RXTE in X-rays, *Fermi*-LAT in GeV energy range and MAGIC in the VHE γ-ray regime. The authors compiled time-averaged spectral measurements done in different wavebands, into a composite MWL data set, spanning from radio frequencies up to VHE γ-ray band (shown in Fig. 6.1). This unprecedented broad-band measurement serves as an excellent proxy for the long-term quiescent/low state SED of the source.

The obtained low-state SED presents the usual two-bumped structure, with the low-energy component peaking in $0.2 - 0.4$ keV range, and the higher-energy bump having its peak around $100 - 200$ GeV. Abdo et al. (2011) have modeled the combined broad-band data set with (stationary) leptonic (one-zone SSC) and hadronic (synchrotron-proton-blazar) scenarios. Both models are found to describe quite well the data (the leptonic SSC fit is shown in Fig. 2.9), and result in a similar power of the jet emission. However, the physical parameters of the best-fit hadronic model

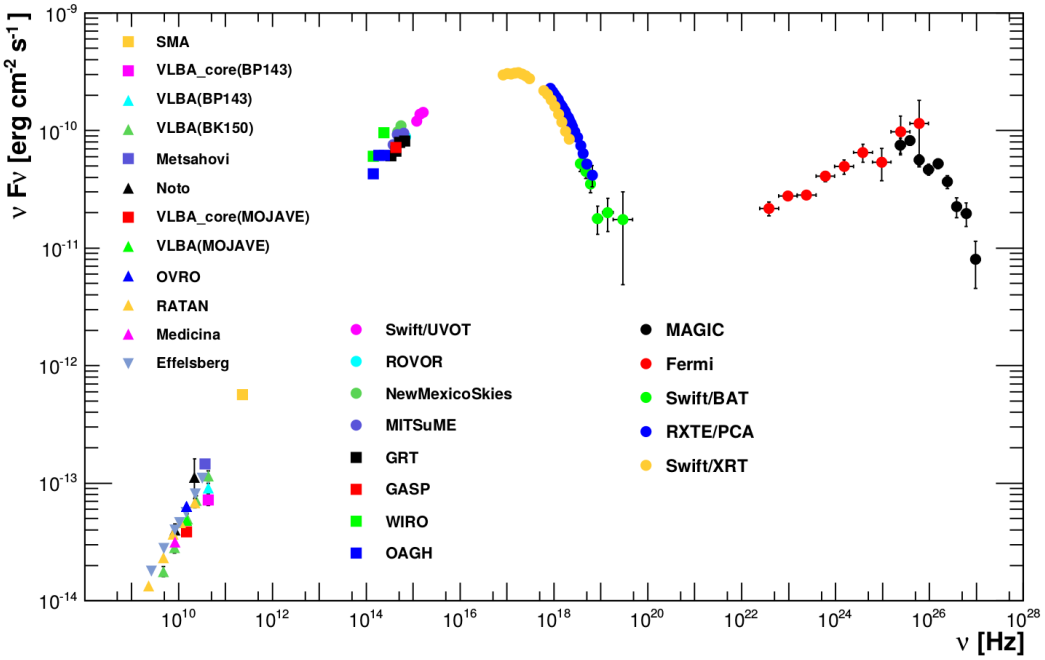


Figure 6.1: Composite MWL data set combining spectral measurements of Mrk 421 low-state emission performed by different instruments from radio band to VHE γ-ray range. The measurements are averaged over the time period of the campaign (January 19 – June 1, 2009). The host galaxy flux has been subtracted, and the optical and X-ray measurements were corrected for the Galactic extinction. The VHE γ-ray data by MAGIC were EBL-deabsorbed using the EBL model by Franceschini et al. (2008). The radio measurements were performed for the most compact core region. (adapted from Abdo et al. (2011))

appeared to be rather extreme, in particular, the magnetic field strength as high as ~ 50 G, a quite compact size of the emission region of only a few Schwarzschild radii $R_{\rm g}$, and maximal energies of protons $\sim 10^{18}$ eV, which all-together requires very particular conditions for particle acceleration and their confinement in the emitting zone. On the other hand, the deduced parameters within the leptonic framework seem to be much more "comfortable" and realistic: the magnetic field of ~ 0.04 G, emitting blob radius $10^4\,R_{\rm g}$, and electron energies up to 50 TeV.

The authors also attempt to interpret the shape of the underlying electron distribution in the case of the leptonic scenario. They found that the particle spectrum is consistent with a broken power law with an index 2.2, which is a typical index resulting from the shock acceleration process. In addition, the electron distribution should feature *two* breaks in order to reproduce accurately the X-ray and hard X-ray parts of the SED. Authors relate the second break to a cooling break, and explain

the first break as resulting from the acceleration process.

6.2.2 Mrk 421 low-state emission: physical modeling and discussion

Physical modeling

Abdo et al. (2011) perform an *instantaneous* (or stationary) modeling of the MWL data set, fitting the broadband emission, and provide a tentative interpretation of the underlying electron spectrum and its features.

In our work, we endeavor to reproduce the low-state SED of Mrk 421, presented in Abdo et al. (2011), going one step further, i.e. describing within a coherent physical scenario both the production of the electron spectrum and its associated emission. We assume that the steady-state spectrum of the electron population in the blob is formed as a result of the physical processes discussed in Section 5.1, and that the steady-state itself is an asymptotic state ($t \to \infty$) of the *time-dependent* evolution of the system. This situation corresponds to the stationary solution ($\frac{\partial N_e}{\partial t} = 0$) of the kinetic equation (Eq. 5.2).

We suppose the following physical scenario of the steady state of Mrk 421. Pre-accelerated electrons are continuously injected into the emitting blob in the form of a steady particle flux. The electron population in the emission region radiates in agreement with the SSC scenario, and cools. Particles also escape the blob, we assume an escape time-scale of $t_{\rm esc} = 1\, R_{\rm b}/c$. The stationary state of the source in our model corresponds to the asymptotically established equilibrium between the gain processes (injection) and losses (cooling and escape). The scenario is schematically illustrated in Fig. 6.2.

We assume that the injected particles are (pre-)accelerated *outside* of the emission zone. As discussed in the sub-section 5.1.1, there are multiple possible scenarios of particle injection. In the study by Yan et al. (2013), authors find that a power law with an exponential cutoff model of the underlying electron spectrum, associated with the Fermi-I acceleration mechanism, is preferable to fit the low-state spectral properties of Mrk 421, over the log-parabola model related to Fermi-II process. Thus, we suppose that the flow of injected particles is generated by the shock acceleration mechanism, and follows a power law with exponential cutoff:

$$Q_{\rm inj,q}(\gamma) = A_{\rm inj}\, \gamma^{-\alpha_{\rm inj}} \exp(-\gamma\,/\,\gamma_{\rm cut}) \tag{6.1}$$

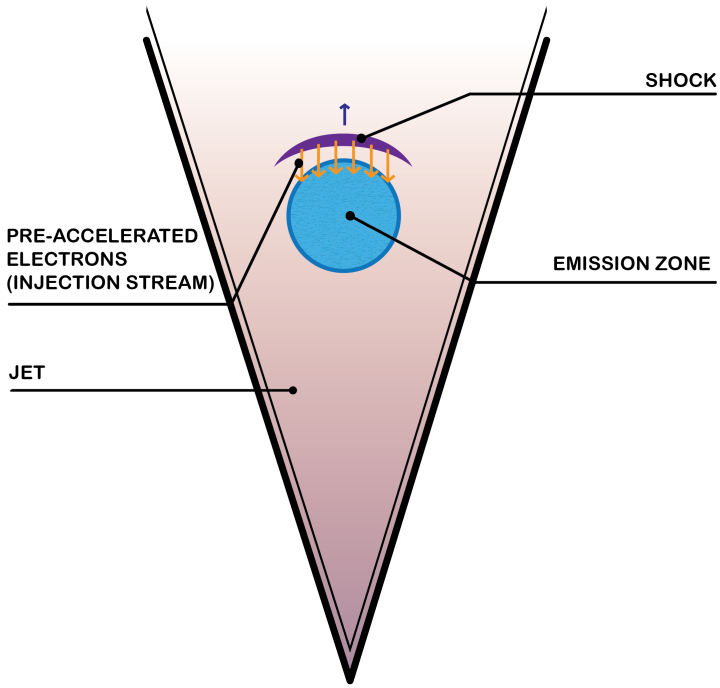


Figure 6.2: Sketch representing a physical scenario for the long-term steady-state emission of Mrk 421. The blue filled circle indicates the VHE γ-ray emission zone (the blob) traveling along the jet. The violet curve shows the stationary shock leading the blob, which accelerates the particles of the upstream plasma, and subsequently injects them into the downstream emitting zone (injection flux indicated in orange arrows). This continuous inflow of pre-accelerated particles is responsible for the long-term steady-state emission of the source. Electrons injected into the blob radiate according to the SSC scenario and cool, and escape from the blob.

where $A_{\rm inj}$ is the injection spectrum normalization, $\alpha_{\rm inj}$ is the power law index, and $\gamma_{\rm cut}$ is the Lorentz factor of the high-energy cutoff. We also assume that the particles are injected only above a certain Lorentz factor $\gamma_{\rm min,inj}$, so that the injection spectrum $Q_{\rm inj,q}(\gamma) = 0$ if $\gamma < \gamma_{\rm min,inj}$.

We simulate the steady-state electron spectrum and SED with our EMBLEM code, following the scenario described above. The radius of the emitting zone is constrained by invoking the usual causality argument, assuming a typical variability time-scale of ∼1 day. This time-scale is used by Abdo et al. (2011) in their instantaneous modeling, and also during the February 2010 flare the flux evolves at a similar time-scale. Using this, we get an estimate $R_{\rm b} \sim 10^{16} \frac{\delta_{\rm b}}{10}$ cm. The typical range for the blob Doppler factor is $10 \lesssim \delta_{\rm b} \lesssim 40$.

As already discussed, the steady-state is achieved asymptotically after the system evolves for the time $t \to \infty$. Clearly, when running the code, one cannot evolve the system for infinitely long, therefore it is necessary to explore for how long one has to evolve the electron distribution, so that after this time duration $t_{\rm evol}$ the distribution does not appreciably evolve anymore, and is very close to the exact asymptotic analytical solution, $\frac{\partial N_{\rm e}}{\partial t} = 0$. By trying different values of $t_{\rm evol}$, we find that the steady-state is established and the electron spectrum and the SED experience almost no more variations after $t_{\rm evol} \simeq (10 - 20)\ t_{\rm esc}$. The relative difference between electron spectra that have evolved for $10\, t_{\rm esc}$ and for $20\, t_{\rm esc}$ is $< 6 \times 10^{-4}$ % across the Lorentz factor grid. Also, we have verified that the electron spectrum that has evolved for $20\, t_{\rm esc}$ is very close to the analytical asymptotic solution given by the Eq. 6.15 (performing the comparison for the case when the IC cooling is not taken into account). The relative difference is < 0.2 % below the high-energy cutoff, and is gradually increasing from 0.2 to ≈ 5 % beyond the high-energy cutoff, which happens due to the decrease of the numerical scheme accuracy in the domain where the electron spectrum sharply drops by a few orders of magnitude.

We vary physical parameters of the model until a reasonable description of the low-state MWL data set is achieved.

It should be noted that in this situation (as well as in further modeling efforts presented in this thesis), we are not actually aiming to make a fit of the data set using statistical methods, but rather endeavor to find a model that can describe the data set reasonably well. In this way, one gets more flexibility to better reproduce the subsets in the energy ranges in which the quality of the data is better and in which one can be reasonably certain that the physical model should work well. In the particular case of Mrk 421, various authors advocate that the extended jet might provide a non-negligible contribution to the optical flux (e.g. Cao & Wang (2013)),

Parameters of quiescent state	Symbol	Our model	Abdo et al. (2011)
Magnetic Field [G]	B	0.04	0.038
Comoving blob radius [cm]	$R_{\rm b}$	2.8×10^{16}	5.2×10^{16}
Doppler Factor	$\delta_{\rm b}$	29	21
Time-scale of electron escape	$t_{\rm esc}$	$1\,R_{\rm b}/c$	not defined
Spectrum of injected electrons	$Q_{\rm inj,q}(\gamma)$	$A_{\rm inj}\,\gamma^{-\alpha_{\rm inj}}\exp(-\gamma/\gamma_{\rm cut})$ for $\gamma \geq \gamma_{\rm min,inj}$ 0 for $\gamma < \gamma_{\rm min,inj}$	not defined
- Injection spectrum normalization [cm^{-3} s^{-1}]	$A_{\rm inj}$	2.63×10^{-3}	
- Injection spectrum slope	$\alpha_{\rm inj}$	2.23	
- Min. Lorentz Factor in inj. spectrum	$\gamma_{\rm min,inj}$	800	
- Lorentz Factor of exp. cutoff in inj. spectrum	$\gamma_{\rm cut}$	5.8×10^{5}	

Table 6.1:: Physical parameters of the source in the low state. 3$^{\rm rd}$ column: parameters in our steady-state model, 4$^{\rm th}$ column: parameters of the instantaneous model by Abdo et al. (2011)

on the basis of a weak correlation of the optical flux with that in X-ray and γ-ray bands often observed in Mrk 421. Therefore, when performing the modeling of the emission of this source (low state and the flare), we focus on reproducing well the X-ray and γ-ray measurements, while allow a certain moderate discrepancy in the description of the optical data. In conclusion, the results of the modeling that we perform in this thesis can be viewed as more qualitative than quantitative.

The best-fit parameters are listed in the Table 6.1, and are in good agreement with the parameters of the instantaneous model of Abdo et al. (2011). The related variability time-scale in our model is $t_{\rm var} \approx 0.4$ d. The comparison between our steady-state model and the spectral measurements of the quiescent emission from Abdo et al. (2011) is illustrated in Fig. 6.3. The model shows a very good consistency with the data.

Discussion

One could notice a discrepancy between the model and the spectral measurement in the energy range 10^{-5} – 10^{-3} eV (radio band). The source appears as extended in the radio band, and the flux was measured for the core region of the jet. It is widely considered that the emission of the core region in this frequency domain is dominated by the synchrotron emission of the large-scale jet. Beyond 10^{-3} eV the contribution of the extended jet radiation may be negligible, and the emission of the blob prevails in the source flux. In our modeling, we do not take into account the radio emission of the extended jet, and that is why our model underpredicts the flux in that energy band.

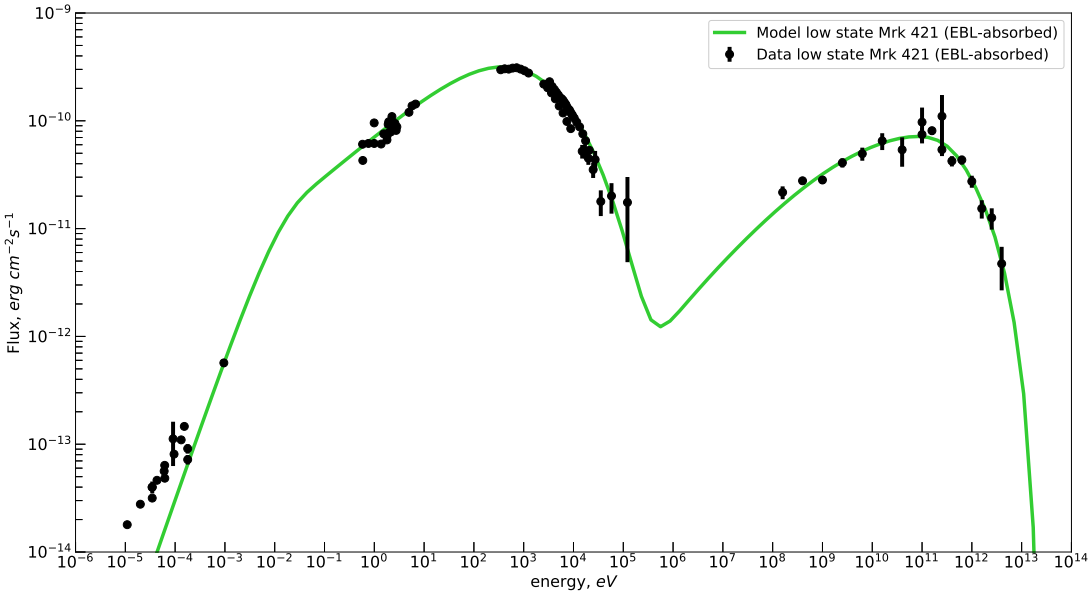


Figure 6.3: Physical modeling of the Mrk 421 SED in the low state. Black data points represent the MWL data set from Abdo et al. (2011), the green curve indicates the model – simulated SED of the source in the asymptotically stationary state. The flux is shown in νF_ν representation. The model curve is EBL-absorbed using the EBL model of Domínguez et al. (2011), and the data in the VHE γ-ray band are not corrected for EBL absorption. The host galaxy emission was subtracted from the measurements in the optical band, and optical-to-X-ray flux was corrected for the Galactic extinction. The flux in the radio band was measured for the most compact core region. An additional component appears at low radio frequencies due to the extended radio source.

Contrary to the modeling done by Abdo et al. (2011), where authors arbitrarily parametrize the generic broken power law profile, in our modeling the electron spectrum, including in particular the cooling break, arises naturally as a consequence of basic physical processes (injection, radiative cooling and escape) taking place in the emitting zone. However, in our model we still need to specify the injection function, as we do not model pre-acceleration of injected particles.

The retrieved power law index of the spectrum of injected electrons, $\alpha_{\rm inj} = 2.23$ is fully compatible with the initial assumption about Fermi-I mechanism responsible for pre-acceleration of injected particles. Indeed, the value of $\alpha_{\rm inj}$ agrees very well with the theoretical predictions for the slope of the spectrum of particles accelerated at the front of the relativistic electron-positron shock (e.g. Sironi et al. (2015)).

There are two commonly accepted options for the shock acceleration site located outside of the emitting zone: (1) the base of the jet, (2) the vicinity of the emission region. We assume the latter, adhering to the scenario by Kirk et al. (1998), in which the injection flow is produced by a shock, located upstream of the emitting zone. As the blob travels along the jet, the upstream plasma overflows and "hits" the blob on a continuous basis, which results in a formation of a stationary shock in the blob frame at the contact surface. The upstream plasma passing through the shock, undergoes Fermi-I acceleration, and the high-energy particles are injected into the downstream emitting region.

The high-energy *cutoff in the injection spectrum* could be due to (1) incapability of the Fermi-I acceleration process to boost particles further in energy under the physical conditions in the acceleration region, (2) a competition between the shock acceleration process and radiative cooling losses, and (3) inefficiency of acceleration of high-energy electrons by a shock. Let us consider each option in detail.

In the case (1), the Hillas criterion for the maximal Lorentz factor an electron can attain in an arbitrary particle accelerator, is

$$\gamma_{\rm max,acc} \sim \frac{e\,c\,B_{\rm acc}\,R_{\rm acc}}{m_{\rm e}\,c^2} \tag{6.2}$$

where $B_{\rm acc}$ is the magnetic field scale in the accelerator, and $R_{\rm acc}$ is the size of the accelerator.

Let us find out physical conditions in the shock acceleration zone, under which the shock would be unable to boost electrons beyond $\gamma_{\rm max,acc} = \gamma_{\rm cut} = 5.8 \cdot 10^5$. Assuming that the magnetic field in the accelerator is $B_{\rm acc} \sim 10^{-3}$ G (typical values measured in VLBI jets), from the Eq. 6.2 one obtains an estimate of the spatial extent

of the accelerator, $R_{\rm acc} \sim 10^{12}$ cm $\sim 10^{-4}\, R_{\rm b}$. As the shock acceleration region has a smaller size and magnetic field compared to the blob, the accelerator produces only a negligible amount of emission.

If in reality the magnetic field in the shock zone is superior to the value assumed before, or the size of the accelerator is larger than the value deduced from Hillas criterion, the shock will ideally be able to accelerate particles to higher Lorentz factors, than $\gamma_{\rm cut}$. In this case, the cutoff at $\gamma_{\rm cut}$ can be explained by radiative losses dominating over the acceleration process above $\gamma_{\rm cut}$ (case (2)). The cutoff position is determined by the balance between the acceleration and radiative cooling losses: $\gamma_{\rm cut}/t_{\rm FI,az} = b_{\rm c}\gamma_{\rm cut}^2$, and so $\gamma_{\rm cut} = b_{\rm c}^{-1} t_{\rm FI,az}^{-1}$, where $t_{\rm FI,az}$ is the Fermi-I time-scale in the shock region and $b_{\rm c}$ is given by the Eq. 5.7. As the cooling is expected to be rather weak (due to a relatively low magnetic field in the acceleration zone), the time-scale of acceleration in the shock region $t_{\rm FI,az}$ will be rather long. Even for an order of magnitude stronger magnetic field $B_{\rm acc} \sim 10^{-2}$ G, one gets $t_{\rm FI,az} \sim 0.5$ years, which is five orders of magnitude longer than the light crossing time of the accelerator size from the previous estimate ($R_{\rm acc} \sim 10^{12}$ cm). Therefore, it is unlikely that the cutoff in the injection spectrum results from the radiative cooling losses.

Finally, in the case (3), the cutoff might originate from a rapid drop in the efficiency of acceleration of electrons by a shock after certain Lorentz factor. In subsection 5.1.3 it was discussed that one of the effects responsible to confine electrons near the shock and causing them to recross the shock front multiple times is scattering on the turbulent motions in the downstream medium. If the electron has a sufficiently high energy, so that its Larmor radius is much larger than the characteristic spatial scale of turbulent motions in the downstream, $r_{\rm L} \gg \lambda_{\delta B}$, the electron instead of scattering back in the upstream, will escape from the downstream region and leave the shock region. Therefore, the shock cannot confine and further accelerate particles with energy higher than the critical, which yields a high-energy cutoff in the spectrum. The value of the critical Lorentz factor is given by $\gamma_{\rm cut} \approx \gamma_{\rm min,az}\, \sigma_{\rm m}^{-1/2}$ (Martin Lemoine, private communication), where $\gamma_{\rm min,az}$ is the minimal Lorentz factor of a particle in the shock region, and $\sigma_{\rm m}$ is the magnetization of the plasma in the vicinity of the shock, which is the ratio of the magnetic energy density to the particle energy density. This allows to estimate the magnetization of the shock. In our case, $\gamma_{\rm min,az} = \gamma_{\rm min,inj}$, which yields $\sigma_{\rm m} \sim 10^{-6}$. The shock is therefore *weakly magnetized*. This agrees well with the expectations: as discussed in sub-section 5.1.3, weakly magnetized shocks are much more efficient at acceleration of particles, than magnetized.

One could also notice that in order to describe well the quiescent state, we have to suppose that no particles with Lorentz factor below $\gamma_{\rm min,inj} = 800$ are injected into the blob. Because of that, the electron distribution in the emitting zone has a

minimal Lorentz factor of electrons of $\gamma_{\rm min} = \gamma_{\rm min,inj} = 800$. Such rather high minimal Lorentz factor of the electrons in the injection flux could arise from their acceleration on the front of weakly magnetized ion-electron relativistic shock. Particle-in-cell (PIC) simulations by Spitkovsky (2007) and Sironi & Spitkovsky (2010)) show that collective effects in the plasma crossing the front of weakly magnetized relativistic shock, lead to transfer of energy from ions to electrons in the downstream medium and eventual rough equipartition of their energy densities. After crossing the shock front, more energetic protons go through the turbulence in the downstream medium without noticing it, while electrons undergo diffusion. The charge separation induces an electric field that pulls the electrons through the turbulent layer and accelerates them. This leads to an increase of the minimal Lorentz factor of electrons (Martin Lemoine, private communication). Since the spectrum of injected electrons has a slope softer than 2 ($Q_{\rm inj} \propto \gamma^{-2.23}$), the bulk of their energy density is concentrated at low Lorentz factors. That is because the energy density "per decade of energy", associated with the injection spectrum, is $\propto \gamma^2\, Q_{\rm inj,q}(\gamma) \propto \gamma^{-0.23} \exp(-\gamma/\gamma_{\rm cut})$. Let us evaluate the total energy density contained in the injected electrons:

$$\varepsilon_{\rm e,inj} \propto \int_{\gamma_{\rm min,inj}}^{\gamma_{\rm cut}} Q_{\rm inj,q}(\gamma) \cdot \gamma\, m_{\rm e} c^2\, d\gamma \propto \frac{\gamma_{\rm cut}^{2-\alpha_{\rm inj}} - \gamma_{\rm min,inj}^{2-\alpha_{\rm inj}}}{2-\alpha_{\rm inj}} \tag{6.3}$$

For our case, $\alpha_{\rm inj} = 2.23$, and as $\gamma_{\rm cut} \gg \gamma_{\rm min,inj}$, this expression simplifies to

$$\varepsilon_{\rm e,inj} \propto \gamma_{\rm min,inj}^{2-\alpha_{\rm inj}} \tag{6.4}$$

and indicates that the energy density in this case is controlled by $\gamma_{\rm min,inj}$. An idea about the value of the energy density associated with the accelerated electrons allows us to assess $\gamma_{\rm min,inj}$.

The values of fraction of the energy density contained in electrons in the downstream plasma observed in PIC simulations (e.g. Spitkovsky (2007) ; Sironi & Spitkovsky (2010)) is $0.1 - 0.3$, which means $\gamma_{\rm min,e} \sim (0.1-0.3) \cdot \frac{m_{\rm i}}{m_{\rm e}} \gamma_{\rm sh}$, where $m_{\rm i}$ is the mass of ion (e.g. proton) and $\gamma_{\rm sh}$ is the Lorentz factor of the shock. For a relativistic shock with $\gamma_{\rm sh} \sim 3$, we obtain $\gamma_{\rm min,inj} \sim 10^3$.

The ion-electron shock should however inject also protons into the emitting zone, therefore, the interpretation of the rather high $\gamma_{\rm min,inj}$ as due to the particle acceleration at the front of the ion-electron shock, is only possible within the hadronic scenarios, while we assume purely leptonic matter content in the blob. This interpretation is however not unique and another possibilities exist. The appearance of $\gamma_{\rm min,inj} = 800$ in a purely leptonic scenario can be explained, for example, by pre-

acceleration of electrons and positrons to this Lorentz factor occurring before they reach the shock, e.g. near the base of the jet. In this thesis we adopt the latter view and further assume purely leptonic blob.

6.3 Observational data set of the February 2010 flare of Mrk 421

Having developed the model of the VHE source in its unperturbed state, we now can proceed to time-dependent modeling of the system, disturbed by various physical processes in and/or near the emitting zone. Our goal is to connect the "long-term" low-state of the source to the exceptional flaring state observed in February 2010, in a coherent and self-consistent framework, in which the flare arises naturally from basic physical ingredients. A special emphasis is put on reproduction of the source MWL variability pattern during the outburst, i.e. fitting the light curves in different wavebands. We present in this section the flare data set under study.

6.3.1 Archival data

We focus on the *brightest* VHE flare of Mrk 421 detected up to now, with the peak flux above 1 TeV of about 27 Crab units. The source underwent this huge flare during the period February 10-23, 2010 (MJD 55237 – 55250). A number of instruments from the radio band to VHE γ-ray regime were monitoring the source in the high state, providing overall an impressive MWL coverage.

Variability was detected from the optical up to the VHE γ-ray band; no significant variability was seen at the radio frequencies. The flare reached its peak on February 16, 2010, (MJD 55243.0 – 55243.5) in the optical, hard X-ray and gamma-ray bands, and on February 17, 2010, (MJD 55244.0) at soft X-ray energies. The positions of the peaks however have an uncertainty around ± 0.5 d, due to 1 d time resolution of the light curves. The strongest variability was observed in X-rays (in particular, hard X-rays), and in the VHE γ-ray range, where the flux level augmented by a factor of 5 – 10. The X-ray flux experienced a sharp increase by more than a factor of 2 over only one day. Also, "harder-when-brighter" behavior was observed in X-ray and gamma-ray energies. At TeV energies, variability at 1 h. time-scale (intra-night variability) was reported (Abeysekara et al. (2020); Shukla et al. (2012)). Variability in the optical V-band appeared to be rather weak, with flux increasing only by $20 - 30\%$ and the light curve in the radio band was found to be consistent

with a constant, thus variability in this band is negligible (Shukla et al. 2012). This however does not necessarily imply that the radio emission of the blob was not variable, as the measured radio flux is dominated by the large-scale jet emission, which is relatively constant in time.

The source has shown a spectacular variability pattern, comprising two flaring events: a secondary, more modest flare is observed after the main one in X-rays, but is not seen at GeV γ-rays. The secondary event starts at around MJD 55247, reaches its peak near MJD 55249.0 – 55249.5, and fades in $\sim$ 3 days.

We have selected for our modeling the archival light curves of the outburst from the optical up to VHE γ-ray band, as well as the available spectral measurements during the high state in X-rays, γ-rays, and in the optical spectral range.

The light curves collection includes the following data. Measurements in the radio band done by OVRO at 15 GHz (Shukla et al. 2012), optical data taken by the SPOL telescope, published in Shukla et al. (2012) with the host galaxy flux subtracted. Measurements in the X-ray band comprise the ones by *Swift*-XRT (Shukla et al. (2012) ; Singh et al. (2015)), by *Swift*-BAT (Shukla et al. 2012), by MAXI (Singh et al. 2015), and by RXTE-PCA and RXTE-ASM (Shukla et al. 2012). The data at MeV-GeV energies collected by *Fermi*-LAT telescope is presented in Singh et al. (2015) (unbinned likelihood analysis with IRFs `P7SOURCE_V6`) and Abeysekara et al. (2020) (binned likelihood analysis with newer IRFs `P8R2_SOURCE_V6`). In the VHE γ-ray range, the flare was monitored by several instruments, including H.E.S.S., VERITAS, TACTIC and HAGAR. H.E.S.S. took data in the time interval MJD 55245 to MJD 55247 (Tluczykont 2011). VERITAS performed observations around MJD 55244.3, and then sampled the Mrk 421 VHE flux during the three subsequent nights Abeysekara et al. (2020). Unfortunately, both observatories started following the source already after the (estimated) flare peak: VERITAS $\sim$1 d, and H.E.S.S. $\sim$1.5 d after the supposed peak position. As a result, only part of the flare decay is available in their measurements. Observations with a better, nearly full time coverage of the outburst in the VHE regime were done by the TACTIC Cherenkov telescope (Singh et al. 2015) and the HAGAR array (Shukla et al. 2012), though with a lower sensitivity. HAGAR performed data taking during the period February 13 – 19, and TACTIC during even longer period, February 10 – 23.

We also use different spectral information collected for the high flux state. The selection includes spectral measurements near the peak of the flare (February 16, 2010) by *Swift*-XRT at keV energies, and *Swift*-BAT in the hard X-ray range (Singh et al. 2015). In the VHE band, VERITAS probed the source spectrum 1 day after the estimated peak (on February 17, 2010, MJD 55244.3) based on 5 h of observations

done that night (Fortson et al. 2012), and H.E.S.S. measured time-averaged spectrum for the period February 17-20, 2010 (Tluczykont 2011), uncovering thus spectral characteristics of the source at extreme γ-ray energies during the fall of the flare. The overall compilation of spectra provides additional constraints on the model of the time evolution of the SED during the outburst, in particular, the exact spectral shape around the flare peak.

Some of the light curves were published in instrumental units. We perform conversion to physical units for these measurements.

The data of RXTE-ASM in the range 1.5-12 keV represents an instrumental count rate depending on time. We convert the count rate to the energy flux in erg cm^{-2} s^{-1} using a conversion formula $F(\text{erg cm}^{-2}\text{ s}^{-1}) = 3.2 \times 10^{-10} \times R(\text{cts/s})$ (Grimm et al. (2002) ; Chitnis et al. (2009)). This conversion factor is deduced for the Crab-like spectrum, and the energy range 2-10 keV. We take into account the difference between Crab and Mrk 421 photon indices into the systematic error. The photon index of Crab in the range 1-100 keV is $\Gamma_{\text{crab}} = 2.1$ (Madsen et al. 2015), while for Mrk 421 in the quiescent state the slope is softer, $\Gamma_{\text{mrk421}} = 2.41$ in the range 3-32 keV (Abdo et al. 2011). The ratio of energy fluxes in the 2-10 keV band, associated with the two photon indices yields a systematic uncertainty of 2%. We neglect the small difference in the energy bands extension, as we have no possibility to include the IRFs in our calculation, and apply the conversion factor of 3.2×10^{-10} to the count rate, including the obtained 2% systematic error to the total uncertainty of the energy flux.

We also check the importance of the effect of photoelectric absorption in the range 1.5-12 keV. Using the HEASARC online tool[2], we retrieve the neutral hydrogen column density in the direction of Mrk 421: $n_{\text{H}} \simeq 1.33 \times 10^{20}$ cm^{-2}. For such a value, the photoelectric absorption is indeed negligible in the energy range of interest.

The data of *Swift*-BAT in the range 15-50 keV are in instrumental units [counts cm^{-2} s^{-1}]. We convert them to physical units [ph cm^{-2} s^{-1}] by using BAT peak measurement of the SED presented in Singh et al. (2015) (in physical units). We fit the hard X-ray spectrum by a simple power law of a form $dN/dE = A\,(E/E_0)^{-\Gamma}$, where $E_0 = \sqrt{15 \cdot 50} \approx 27$ keV (pivot energy), and integrate the best-fit model in the energy range of interest (15-50 keV), which yields a physical flux of the source at the peak of the flare in this band. From this, comparing the obtained value with the flux in counts, we evaluate the relevant conversion factor between counts and photons. We neglect the dependence of the conversion factor on the spectral slope that is

[2] https://heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3nh/w3nh.pl

changing during the flare, assuming constant "peak" value for the entire time range. We estimate the uncertainty on the conversion factor, attributing it to the error on the spectral slope in our fit and hence to the error on the physical flux, resulting from the integration of the spectrum. The conversion factor (counts-to-photons) we get is $C_{\text{c-p}} = 1.3 \pm 0.1$. We apply it to the light curve in counts, and reconstruct the physical light curve in [ph cm^{-2} s^{-1}]. We include the systematic uncertainty to the total uncertainty of the physical flux by propagating the uncertainty on the conversion factor.

The flux at TeV energies measured with the TACTIC telescope was found to be systematically lower than the one by the H.E.S.S. array, when computed for the same energy range and compared for the same night. We recalculated H.E.S.S. fluxes above 2 TeV for the TACTIC energy range (1.5 – 11 TeV), using the average spectrum obtained by H.E.S.S., assuming variable nightly amplitude but constant spectral shape (given by the average SED), during the period of overlap. The TACTIC flux appeared to be substantially smaller than the one by H.E.S.S. for all three nights. We consider that this divergence stems from inaccurate calibration of absolute physical fluxes measured by TACTIC[3], and apply rescaling to the TACTIC light curve by a constant factor of 5.7. After that upscale, the two light curves appear to be in a good agreement with each other (cf. Fig. 6.4).

6.3.2 Analysis of *Fermi*-LAT gamma-ray data

We cross-check the light curves obtained with an unbinned likelihood analysis by Singh et al. (2015) and with a binned likelihood analysis by Abeysekara et al. (2020), by using the aperture photometry method to extract the source light curves. We analyze publicly available data of *Fermi*-LAT telescope. For our analysis we select the Mrk 421 February 2010 flare data spanning 16 days, from MJD 55237 to MJD 55253. We process the data in a standard way following the recommendations of the Fermi Science Support Centre [4].

We select events in the energy range 0.1-100 GeV. The angular extent of the point spread function (PSF) of LAT is a strong function of energy. We take it into account by choosing different sub-classes of the SOURCE class for the analysis in different

[3]TACTIC is a single IACT with a relatively small mirror collection area of 9.5 m^2. Accurate calibration of absolute physical flux for such setups is much more difficult compared to large IACT arrays.

[4]https://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/aperture_photometry.html

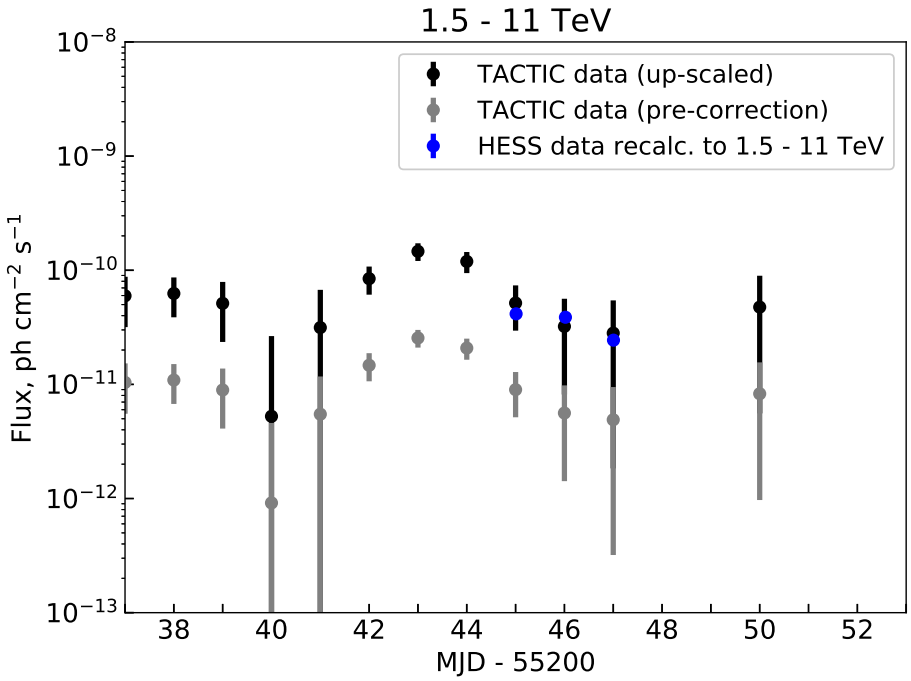


Figure 6.4: VHE light curve by TACTIC in the energy range 1.5-11 TeV before the correction (gray points, Singh et al. (2015)) and after the correction by a factor of 5.7 (black points), compared to the light curve by H.E.S.S., recalculated for the TACTIC energy range (blue points). The log scale is applied to the y-axis.

energy ranges. In the range 0.1-1 GeV, we use only events from the FRONT subclass (γ-rays converting to electron-positron pairs in the front section of the telescope tracker, `evtype=1`). In the range above 1 GeV, we use all the SOURCE class events, including FRONT and BACK converting events (`evtype=3`). Also, we select only those events coming from zenith angles below 105° to avoid the Earth albedo. Since we are not modeling the background (contrary to the likelihood approach), we apply a narrow circular ROI of 1° in order to reject the majority of background events, but at the same time to collect most of the source events and have them dominate the signal. We process the events selected from a 1° circle around the source using Science Tools package v10r0p5 and the `P8R2_SOURCE_V6` instrument response functions (same as used by Abeysekara et al. (2020)) with the help of the *gtselect* - *gtmktime* - *gtbin* - *gtexposure* tool chain. The *gtexposure* tool is used with the `apcorr=yes` to account for the fact that the 1° circle contains only an energy-dependent fraction of the γ-ray events from the source. Also, we use the spectral index of 1.77, which is an average spectral index of the low state of the source in the range 0.1 - 100 GeV.

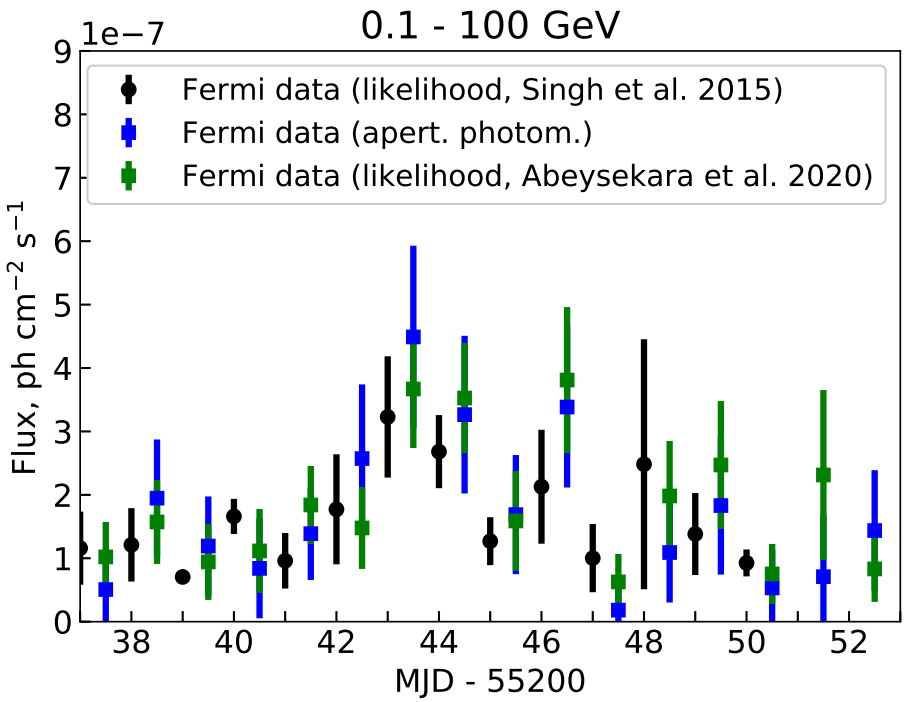


Figure 6.5: Comparison of the *Fermi*-LAT light curves in the energy range 0.1 – 100 GeV, obtained with the unbinned likelihood analysis with IRFs `P7SOURCE_V6` done by Singh et al. (2015) (black points), the binned likelihood analysis with more recent IRFs `P8R2_SOURCE_V6` done by Abeysekara et al. (2020) (green points), and aperture photometry with IRFs `P8R2_SOURCE_V6` done by the author of this thesis (blue points). All the light curves show rather limited quality and appear to be consistent within the error bars.

As the result, we get the measurement of the flux from the direction of the source in the energy range of interest. In addition to the source counts, the diffuse Galactic and isotropic backgrounds are contributing to this measurement. We subtract the backgrounds by extracting the background flux estimates from four 2° source-free circles with centers at the distance 5° from the source position at different sides of the source. Finally, we bin the net source events in time with bin width of 1 day. We have verified that the aperture photometry measurements obtained in this way are fully consistent with the unbinned likelihood analysis results reported by Singh et al. (2015) and binned likelihood results presented by Abeysekara et al. (2020) (cf. Fig. 6.5).

6.4 A general criterion to test one-zone flaring scenario with a passing shock wave

As discussed in sub-section 5.3.2, a one-zone model is the most basic scenario proposed for flaring activity and it is tempting to adopt it in as much cases as possible. However, the improving MWL and time coverage challenge such simple scenarios and we examine here their limits.

Spectral hardening in the high state can be explained in a most simple way by re-acceleration of the electron population in the emitting zone. The most standard scenario in which the particle re-acceleration is involved, is a one-zone SSC model with a transient shock wave. As stated in sub-section 5.3.1 and 5.3.2, shocks are thought to be common phenomena in the violent environment of AGN jets, forming due to diverse physical processes in the relativistic plasma. Two configurations of the interaction between the blob and the shock are possible: (1) a fast shock traveling down a jet and traversing the emitting region (Marscher & Gear 1985), (2) the blob crossing a standing shock in the jet (see Fig. 6.6). In both cases, the transient shock re-accelerates the particle population in the emitting region via Fermi-I mechanism during the time of the crossing, which results in modification of the electron spectrum, in particular its hardening, and most likely also a change in the spectral break position, or the emergence of a new break. Accelerated low-energy particles migrating to higher Lorentz factors, create an excess of electrons at higher-energy part of the electron spectrum, which leads to enhanced emission intensity, i.e. a flare. Due to spectral hardening, the increase of flux level will be larger with higher photon energy. As the shock exits the blob, the initial configuration of the system is restored, the shocked ensemble of particles escapes the blob and cools, and the source returns back to the quiescent emission state.

The one-zone scenario described above has a minimal number of free parameters. In the kinetic approach we are following, only two parameters govern the flaring behavior: shock acceleration time-scale $t_{\rm FI}$ – (roughly) how long does it take for a single particle to appear at the next decade of energy, and transit time $t_{\rm cs}$ – duration of the acceleration phase (same as $t_{\rm dur,FI}$ in sub-section 5.1.3), i.e. for how long the shock keeps transferring its energy to the particles and power the flaring activity. The latter is equal to the time it takes for the shock to cross the emitting blob. The other usual free parameters of the SSC scenarios are fixed by their values deduced from the quiescent state analysis, under our assumption that the flare is just a moderate perturbation of the stationary configuration.

We assume that the time-scale of particle escape from the emission region during

the passage of the shock is not significantly altered, and remains $t_{\rm esc} = 1\,R_{\rm b}/c$. Indeed, the particles may need a longer time to be advected away from the shock and leave the blob, however since we generally consider a relativistic shock, the probability of escape from it in one or very few crossings is quite high (see sub-section 5.1.3), and on average particles easily leave the relativistic shock which justifies to keep the escape time-scale constant in this scenario. To minimize the number of free parameters, we also suppose that the passing shock does not affect the macroscopic properties of the emitting zone, like its radius (we neglect the possible compression) and the magnetic field, which is consistent with the flare being just a relatively small perturbation. In addition, here we neglect the inverse Compton cooling losses, so that the cooling rate is equal to the synchrotron cooling rate, which is constant in time.

We develop here a general criterion to allow or dismiss the one-zone model described above. Since in most cases, the observational data of the outbursts are of the highest quality around the peak of the flare, we base the verification of the validity of the one-zone scenario on its ability to reproduce the observed peak fluxes from the source at different frequencies at a given variability time-scale. The information on the maximal flux and variability time-scale in multiple wavebands is indeed often used to constrain models for the flare origin. However the physics of interaction between a shock and an emitting zone may be way more complex, we do not try in this section to reproduce detailed shape of the light curve, and focus our attention on the multi-band peak flux.

We aim to formulate our criterion in the framework of analytical calculations, to ensure its universality and straightforward applicability, so that one does not require to run any specific emission code to test the one-zone model with a transient shock. In order to predict the flux for an arbitrary frequency value / spectral range within this scenario, we first have to estimate in the general case the flux variations caused by an arbitrary perturbation of the electron spectrum, and then establish how the electron spectrum evolves with time during the passage of a shock. We can then apply it to the real data: using the deduced general form of the time-dependent electron distribution, we adjust the model parameters in a way to fit the maximal flux in one spectral range (e.g. in X-rays), and predict the flux enhancement in other energy band(s) (e.g. in the optical band) with the inferred parameters. The theoretical value is then compared to the one from the data. In case they match each other quite well, this would be an argument in favor of the one-zone scenario, and in case strong discrepancy between the model and the observations is found, this shows that basic one-zone scenarios fail to provide a reasonable description of the data.

Let us first connect the increase of photon and energy flux to varying electron distribution. In our approach we consider only the MWL variability of the synchrotron

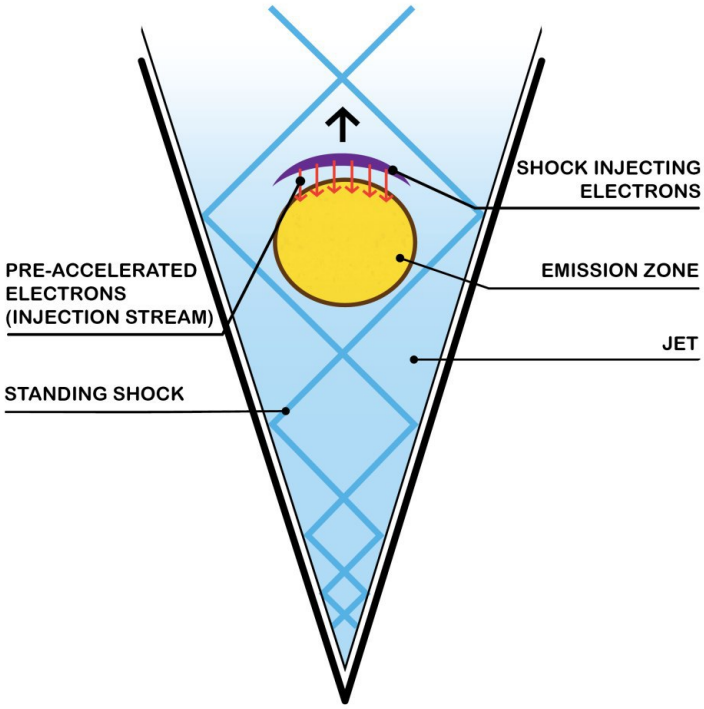


Figure 6.6: Sketch representing a one-zone flaring scenario with a passing shock wave. In this model, the outburst occurs due to an interaction between the shock and the emitting region. Upon entering the blob, the shock re-accelerates particles confined in it, boosting them to higher energies, and therefore perturbs the electron spectrum, which leads to a flaring event. An example of such a setting is a passage of the emitting zone through a knot of a standing shock with a so-called "diamond structure". During this interaction, the particle population is re-accelerated by transient Fermi-I process. The violet curve shows the stationary shock leading the blob, injecting pre-accelerated particles into it.

emission, e.g. for Mrk 421 the synchrotron component of the SED stretches from radio frequencies to hard X-ray energies. To simplify our analytical calculations, we use the δ-approximation for the synchrotron emissivity of a single relativistic electron. In this approximation an electron with a Lorentz factor γ produces radiation at only one frequency, called critical frequency $E_{\rm s}$, related to the electron Lorentz factor in the following way (e.g. Rybicki & Lightman (1979))

$$E_{\rm s} \simeq \frac{5 \times 10^{-9}\, B_{\rm G}\, \delta_{\rm b}}{1+z}\, \gamma^2 \tag{6.5}$$

where $B_{\rm G}$ is the magnetic field in Gauss, z is the redshift.

The total synchrotron emissivity of a single electron is then given by

$$P_{\rm s}(E,\gamma) \propto \gamma^2\, \delta(E - \kappa\gamma^2) \tag{6.6}$$

where $\kappa = 5 \times 10^{-9} B_{\rm G} \delta_{\rm b}/(1+z)$. The factor γ^2 before the δ-function stems from the normalization condition that the total power (integrated over all the frequencies) of synchrotron radiation emitted by a relativistic electron is $P_{\rm tot} \propto \gamma^2$.

The synchrotron emissivity per solid angle of an electron population with an electron spectrum $N_{\rm e}(\gamma)$ is

$$j_{\rm s}(E) = \frac{1}{4\pi} \int_{\gamma_{\rm min}}^{\gamma_{\rm max}} N_{\rm e}(\gamma)\, P_{\rm s}(E,\gamma)\, d\gamma \tag{6.7}$$

where $P_{\rm s}(E,\gamma)$ is given by the Eq. 6.6. We are interested in the SED of the synchrotron emission, the $\nu F_\nu = E^2 \frac{dN_{\rm ph,syn}}{dE}(E)$ flux is $\propto E\, j_{\rm s}(E)$, evaluating the integral for $j_{\rm s}(E)$ we get

$$E^2 \frac{dN_{\rm ph,syn}}{dE}(E) \propto \bar{\gamma}^3 \cdot N_{\rm e}(\bar{\gamma}) \tag{6.8}$$

where $\bar{\gamma} = \bar{\gamma}(E) = \sqrt{E/\kappa}$. The quantity $\bar{\gamma}$ represents the Lorentz factor associated with a given photon energy, i.e. a Lorentz factor of an electron radiating synchrotron photons of energy E. The result in Eq. 6.8 is identical to the one obtained by Dermer & Schlickeiser (2002). This equation shows how perturbations of the electron spectrum at a specific Lorentz factor are reflected on the spectral density of the flux at the corresponding (critical) frequency; this dependence appears to be linear.

Now let us find the similar link for light curves flux. The energy and photon flux of the synchrotron emission in a spectral range spanning from $E_{\rm min}$ to $E_{\rm max}$ is obtained by integration of the spectrum in this range

$$F_{\rm erg} \propto \int_{E_{\rm min}}^{E_{\rm max}} E \frac{dN_{\rm ph,syn}}{dE}(E)\, dE \propto \int_{E_{\rm min}}^{E_{\rm max}} E^{1/2}\, N_{\rm e}(\bar{\gamma}(E))\, dE \tag{6.9}$$

$$F_{\rm ph} \propto \int_{E_{\rm min}}^{E_{\rm max}} \frac{dN_{\rm ph,syn}}{dE}(E)\, dE \propto \int_{E_{\rm min}}^{E_{\rm max}} E^{-1/2}\, N_{\rm e}(\bar{\gamma}(E))\, dE \tag{6.10}$$

These equations connect the electron spectrum changes to the variability of flux in a certain energy band. The key quantity, that can be measured from the light curves of the flaring event is a flux increase factor ξ, ratio of the peak flux to the pre-flare (low-state) flux. Using the expressions 6.9 and 6.10 we find that this factor is

$$\xi = \frac{F_{\rm peak}}{F_0} = \frac{\int_{E_{\rm min}}^{E_{\rm max}} E^s \cdot N_{\rm e,peak}(\bar{\gamma}(E))\, dE}{\int_{E_{\rm min}}^{E_{\rm max}} E^s \cdot N_{\rm e,0}(\bar{\gamma}(E))\, dE} \tag{6.11}$$

where $s = 1/2$ for the increase factor of the energy flux, and $s = -1/2$ for the one of the photon flux. $N_{\rm e,peak}(\gamma)$ is the electron distribution during the peak of the outburst, and $N_{\rm e,0}(\gamma)$ is the pre-flare (steady-state) electron spectrum.

In case one has high-quality spectral measurements for the peak of the flare, it is easier to use those data, to avoid the need to compute an integral over the photon energies. We write out the analogous expression for ξ for the spectral flux increase factor, using the Eq. 6.8:

$$\xi_{\rm spec} = \frac{dN_{\rm ph,syn,peak}(E)/dE}{dN_{\rm ph,syn,0}(E)/dE} \propto \frac{N_{\rm e,peak}(\bar{\gamma}(E))}{N_{\rm e,0}(\bar{\gamma}(E))} \tag{6.12}$$

Thus, flux increase factors allow to probe the variations of the electron spectrum at corresponding Lorentz factors or in corresponding range of electron energies.

The goal is now to find a way to establish whether the variations inferred from the data could be caused by a passing shock wave. This can be achieved by comparing the observed/measured flux enhancement factors to those expected if the shock wave is responsible for initiating the outburst. In order to evaluate the expected flux increase factor we have to know $N_{\rm e,peak}(\gamma)$, which can be found using the time-dependent electron distribution disturbed by a shock. In other words, we have to connect the

steady-state electron spectrum $N_{\rm e,0}(\gamma)$, to the one around the peak $N_{\rm e,peak}(\gamma)$. We then focus on the analytical derivation of the time evolution of the electron spectrum $N_{\rm e,FI}(\gamma, t)$ during interaction of the shock with the emission region. The evolution of $N_{\rm e,FI}(\gamma, t)$ obeys the kinetic equation with the transient Fermi-I acceleration term:

$$\frac{\partial N_{\rm e,FI}(\gamma,t)}{\partial t} = \frac{\partial}{\partial \gamma}\left[W(\gamma)\, N_{\rm e,FI}(\gamma,t)\right] - \frac{N_{\rm e,FI}(\gamma,t)}{t_{\rm esc}} + Q_{\rm inj}(\gamma) \tag{6.13}$$

where $W(\gamma) = b_{\rm c}\gamma^2 - \frac{\gamma}{t_{\rm FI}}$

We set the time when the shock enters the emission region to $t = 0$. We have solved this equation analytically, the full derivation is presented in the Appendix A:

$$N_{\rm e,FI}(\gamma,t) = N_{\rm e,0}(\gamma) + \int_0^t \frac{\Gamma(\gamma,t,t')\cdot e^{(1/t_{\rm esc}\,-\,1/t_{\rm FI})\cdot(t'\,-\,t)}}{b_{\rm c}\, t_{\rm FI}\, \gamma^2} \times \\ \times \left[Q_{\rm inj}(\Gamma(\gamma,t,t')) + \left(b_{\rm c}\Gamma(\gamma,t,t') - \frac{1}{t_{\rm esc}}\right)\cdot N_{\rm e,0}(\Gamma(\gamma,t,t'))\right] dt' \tag{6.14}$$

here t is the time elapsed after the shock enters the blob, and

$$N_{\rm e,0}(\gamma) = \frac{1}{b_{\rm c}\gamma^2}\int_{\gamma}^{\gamma_{\rm max}} Q_{\rm inj}(\gamma')\cdot \exp\left(\frac{1/\gamma' - 1/\gamma}{b_{\rm c} t_{\rm esc}}\right) d\gamma' \tag{6.15}$$

is the steady-state particle spectrum, and

$$\Gamma(\gamma,t,t') = \frac{\gamma\cdot e^{(t'-t)/t_{\rm FI}}}{1 + \gamma b_{\rm c}\, t_{\rm FI}\,(e^{(t'-t)/t_{\rm FI}} - 1)} \tag{6.16}$$

Now, as we know how the particle distribution evolves during the passage of the shock, we can compute the expected enhancement of the flux in different energy bands. First, we need to calculate the steady-state electron spectrum. This can be achieved by the modeling of the available pre-flare data, similarly to the approach presented in Section 6.2, which yields the physical parameters of the source (B, $\delta_{\rm b}$, $R_{\rm b}$, etc.) and the injection spectrum $Q_{\rm inj}(\gamma)$. Having this information, we compute $N_{\rm e,0}(\gamma)$ using the Eq. 6.15. Since the low-state spectrum does not depend on the parameters of the flare scenario ($t_{\rm FI}$ and $t_{\rm cs}$), we can evaluate the integral 6.15 and then either fit it with a certain function (e.g. a broken power law with a cutoff, or a log-parabola), or simply tabulate it for further use.

To determine the spectrum during the flare peak, we need to know the transit time of the shock through the blob, and the shock acceleration time-scale. The crossing time of the shock induces the duration of the activity of the transient Fermi-I acceleration process in the emitting zone, which is most plausibly dictated by the observed time-scale of the flux rise during the outburst, corrected for the relativistic effects, $t_{\rm rise} \sim t_{\rm cs}/\delta_{\rm b}$. However this is not so simple and requires further analysis as follows.

Let us check this guess and estimate the rise time of the flux, until it reaches maximal value. As already discussed, within the one-zone scenario with the transient shock, the flux growth is caused by inflow of lower-energy electrons to a specific higher energy bin. The migration of particles in Lorentz factor space is governed by the "flux" term in the kinetic equation, $\frac{\partial}{\partial\gamma}\left(-\dot{\gamma}\, N_{\rm e}(\gamma)\right)$:

$$\dot{\gamma} = -W(\gamma) = -b_{\rm c}\gamma^2 + \frac{\gamma}{t_{\rm FI}} \tag{6.17}$$

Solving this equation, we get an evolution of a Lorentz factor of a single particle with a starting Lorentz factor γ_0, which gains energy via shock acceleration and loses energy via synchrotron cooling

$$\gamma(t) = \frac{1}{b_{\rm c} t_{\rm FI}(1 - e^{-t/t_{\rm FI}}) + \frac{1}{\gamma_0} e^{-t/t_{\rm FI}}} \tag{6.18}$$

From this expression one can see that first, the Lorentz factor of a particle grows with time, and then, after long enough time $t \to \infty$, the Lorentz factor stalls tending to a critical value of $\gamma_{\rm crit} = \frac{1}{b_{\rm c}\, t_{\rm FI}}$, which is the Lorentz factor at which the cooling time-scale equals the Fermi-I acceleration time-scale, $t_{\rm cool} = t_{\rm FI}$. Such stagnation due to cooling losses can be reached only if $t_{\rm cs} > t_{\rm FI}$, and precludes the acceleration process boosting particles further in energy, inducing a break-like feature in the particle spectrum. Let us invert the Eq. 6.18 and find out how much time an electron needs to reach a certain Lorentz factor γ, starting from γ_0

$$t = t_{\rm FI} \cdot \ln\left(\frac{\gamma}{\gamma_0} \cdot \frac{\gamma_{\rm crit} - \gamma_0}{\gamma_{\rm crit} - \gamma}\right) \tag{6.19}$$

The flux rise will stall and the light curve will start showing a plateau, after the lowest-energy electron with the initial Lorentz factor $\gamma_0 = \gamma_{\rm min}$ has approached the cooling break $\gamma_{\rm crit}$ very closely. This effect will occur after the time period (taking into account that usually $\gamma_{\rm crit} \gg \gamma_{\rm min}$)

$$t \approx t_{\rm FI} \cdot \ln\left(\frac{\gamma}{\gamma_{\rm min}} \cdot \frac{1}{1-\gamma/\gamma_{\rm crit}}\right) \tag{6.20}$$

This expression is only valid for $\gamma < \gamma_{\rm crit}$. For $\gamma_{\rm crit} \simeq 10^2\,\gamma_{\rm min}$, to attain $\gamma = 0.5\,\gamma_{\rm crit}$, the particle needs $\simeq 5\,t_{\rm FI}$, and to reach $\gamma = 0.9\,\gamma_{\rm crit}$, the particle needs $\simeq 7\,t_{\rm FI}$. Evidently, the particle needs infinite amount of time to reach $\gamma_{\rm crit}$, so we consider that the flux stagnation will happen somewhere after $7\,t_{\rm FI}$. Therefore, in case the transit time of the shock is longer than $\sim 7\,t_{\rm FI}$, the light curve will first show a sharp rise at a time-scale of few of $t_{\rm FI}/\delta_{\rm b}$, and then a transition in the shape will occur towards a plateau-like profile until the moment when the shock exits the emitting zone. For light curves of such type, one can estimate roughly the shock acceleration time-scale from the measurement of the time-scale over which the flux experiences a steep rise. The total duration of the "non-falling" part of the light curve in this case equals to the transit time of the shock in the observer's frame, $t_{\rm cs}/\delta_{\rm b}$. If the crossing time of the shock is shorter than $\sim 7\,t_{\rm FI}$, then the light curve rise will not have time to reach the plateau yet, the flux rise will be ceased at the moment when shock leaves the blob, and the total duration of the "non-falling" (here also rising) part of the light curve will be $t_{\rm cs}/\delta_{\rm b}$.

This considerations are valid for the light curves at characteristic photon energies below those corresponding to critical Lorentz factor, $E < E_c = \kappa\gamma_{\rm crit}^2$, since in this situation, acceleration dominates over cooling. In the energy bands above the critical energy (but still within the synchrotron bump of the SED), the cooling losses dominate, so the flux increase will be caused by the cooling of higher-energy electrons. For this case, we rewrite the Eq. 6.19 for the time during which an electron initially having the maximal Lorentz factor $\gamma_0 = \gamma_{\rm max}$ will cool down to a lower Lorentz factor γ:

$$t = t_{\rm FI} \cdot \ln\left(\frac{\gamma}{\gamma_{\rm max}} \cdot \frac{\gamma_{\rm max}-\gamma_{\rm crit}}{\gamma-\gamma_{\rm crit}}\right) \tag{6.21}$$

For $\gamma_{\rm max} = 10\,\gamma_{\rm crit}$, and $\gamma = 1.1\,\gamma_{\rm crit}$ (cooling almost down to the position of the cooling break), we obtain that the most energetic electron will arrive to the energy bin slightly higher than $\gamma_{\rm crit}$ during $\simeq 2\,t_{\rm FI}$. That means that the light curve at energy $E > E_{\rm c} = \kappa\gamma_{\rm crit}^2$ (in the cooling-dominated part of the synchrotron hump) will show a very sharp rise during only $\sim 2\,t_{\rm FI}/\delta_{\rm b}$, and then, since the cooling inhibits the acceleration process and prevails, the light curve should experience an immediate rapid drop at a time-scale of the order of $\sim t_{\rm cool}$. The overall shape of this light curve will represent a rather narrow sharp peak.

We can now clarify the interpretation of $t_{\rm rise}$. In general, one should infer the transit time of the shock from the observed time interval $t_{\text{s-f}}$, which generalizes the concept of $t_{\rm rise}$. The quantity $t_{\text{s-f}}$ is the time interval between the start of the flare and the moment when the flux starts to fall, measured in the energy bands where the acceleration dominates over the cooling process, $E < E_{\rm c} = \kappa\gamma_{\rm crit}^2$:

$$t_{\rm cs} \approx t_{\text{s-f}} \cdot \delta_{\rm b} \tag{6.22}$$

The value of $\gamma_{\rm crit}$ is not always obvious, since we do not know $t_{\rm FI}$ beforehand, but we can select the light curve showing the longest $t_{\text{s-f}}$ for its measurement, as such long flux rise (or rise and then stagnation) can be only attributed to the acceleration process within the considered scenario. The longest $t_{\text{s-f}}$ is then directly related to the duration of the activity of the Fermi-I acceleration inside the emitting zone via the Eq. 6.22.

Then, the Fermi-I acceleration time-scale $t_{\rm FI}$ is determined from the longest observed flux increase with the help of the Eq. 6.11 if we use the light curve, and using the Eq. 6.12 if one prefers to use spectral measurements. The electron spectrum at the peak of the flare is the time-dependent electron spectrum $N_{\rm e,FI}(\gamma, t)$ at the moment $t = t_{\rm cs}$: $N_{\rm e,peak}(\gamma; t_{\rm FI}) = N_{\rm e,FI}(\gamma, t = t_{\rm cs}; t_{\rm FI})$, where the semicolon separates function arguments from parameters. The peak particle distribution depends on the parameter $t_{\rm FI}$, the value of which regulates the observed flux enhancement in X-ray band. We solve numerically the Eq. 6.11 for $t_{\rm FI}$, and retrieve its value $t^*_{\rm FI}$ needed to yield the increase of the flux observed in the X-ray light curve.

Finally, with the recovered shock acceleration time-scale $t^*_{\rm FI}$, one can predict the expected flux enhancement in the optical band from the Eq. 6.11 (optical light curve) or Eq. 6.12 (optical spectral measurements), by applying the $t^*_{\rm FI}$ to the peak electron spectrum, $N_{\rm e,peak}(\gamma) = N_{\rm e,peak}(\gamma; t_{\rm FI} = t^*_{\rm FI})$. Then we compare the predicted value of the flux increase to the one observed in the optical data. In case there is a significant discrepancy between the two values, this implies that the one-zone model with a transient shock cannot explain the observed MWL flaring behavior, and we can discard this scenario. If the values appear in good agreement, this supports the one-zone model, however it is clear that further modeling is required to make definitive conclusions.

In case the timing properties of the emission are not very well constrained, and one does not have a precise information of the time interval $t_{\text{s-f}}$, this parameter can be reconstructed adding a light curve in a third energy range, e.g. hard X-rays (one has to always make sure that the energy bands have to be within the synchrotron part of the SED). Instead of solving one equation for the X-ray band to deduce $t_{\rm FI}$, we

solve a system of two equations for soft X-ray and hard X-ray flux increase factors, to derive at the same time $t_{\rm FI}$ and $t_{\rm s\text{-}f}$. In case such system has no solution, we can already reject the scenario on the grounds of its failure to describe together soft X-ray and hard X-ray flux enhancements. In case the solution exists, we apply the retrieved $t^*_{\rm FI}$ and $t^*_{\rm s\text{-}f}$ to the general form of the time-dependent electron spectrum, and predict with the recovered peak electron spectrum the factor by which the optical flux has to augment at the moment of the flare peak. Again, comparison between the expected value and the one derived from the data, allows us to judge whether the one-zone scenario could be a plausible option to explain the flaring event or the model should be rejected.

6.5 Physical modeling of Mrk 421 February 2010 flare

In this section, we perform a detailed physical modeling of the Mrk 421 February 2010 flare, using the time-dependent EMBLEM code we have developed. We investigate whether the flare can be caused by a moderate, stable (growing and then fading) perturbation of the quiescent emission region, and endeavor to connect the low-state emission (sub-section 6.2.1) to the one during the outburst (Section 6.3), within the framework of the one-zone and two-zone scenario.

Inspecting the set of MWL light curves of the outburst, we establish that during the flare the X-ray flux experienced a substantial increase, by a factor $\sim 4-7$), whereas in the optical V-band the source displayed very modest flux change (by 20-30% at most). This implies that during the high state, the photon spectrum in the optical-to-X-rays range was harder than in the quiescent state. As was already discussed, such behavior might be caused by particle acceleration inside the source. Based on that, we consider, that in both one-zone and two-zone models, the flaring activity is arising due to an intervening electron acceleration process.

We first test one-zone scenarios applying the criterion established in the previous section, as well as the EMBLEM code, to explore whether it is possible to connect the steady-state emission to the flaring one in a simplest way with a minimal number of free parameters. Next, after having concluded the inability of the one-zone models to fit the flare, we gradually increase the level of complexity and number of free parameters, by trying to achieve the flaring state described by the MWL data set, from the low state, within a two-zone model. Finally, we discuss the best-fit scenario and possible implications for the origin of the remarkable outburst of Mrk 421.

6.5.1 One-zone model

We first try to explain the observed flaring behavior by disturbing the steady-state configuration of the source described in sub-section 6.2.2 with re-acceleration of particle population in the emission blob by (1) a traversing shock wave, and (2) by Fermi-II process due to spontaneous turbulence confined to the emitting zone. We assume that the macroscopic physical parameters of the blob do not vary to the first order, and remain equal to the ones of the low state.

Analytical results: a passing shock

We first check whether the flare could be triggered by a shock passing through the emission region and disturbing its electron population and apply the general validity criterion derived in Section 6.4. The following study both serves as a verification of the one-zone model, and as an example of the application of the general criterion. As input data for the test procedure, we use the X-ray light curve in the spectral range from 0.5 to 2 keV by *Swift*-XRT, and the optical V-band "light curve".

We first constrain the transit time of the crossing shock wave. Most of the available light curves related to the synchrotron emission (for Mrk 421 it spans from radio band to hard X-ray energies), exhibit a rise time of 3 to 4 days in the observer's frame, only the hard X-ray flux enhances within 1.5 – 2 days. Also, the 0.5 - 2 keV light curve does not feature any flux saturation features, with a rather cuspy peak, which supports the assumption that the flux in the range 0.5 to 2 keV started to drop abruptly due to the shock exit from the blob. We thus adopt an average value of $t_{\text{s-f}} = 3.5$ d seen in the X-ray light curve, for the duration over which the flux augments to its peak value, which translates into $t_{\text{cs}} \approx 101.5$ d in the emitting zone frame. From this we also can estimate the velocity of the shock speed relative to the blob, $\beta_{\text{sh}} = \frac{R_{\text{b}}}{c\, t_{\text{cs}}} \approx 0.1$.

Next, from the X-ray light curve we directly measure the flux enhancement factor $\xi_{\text{X}} = \frac{F_{\text{x,peak}}}{F_{\text{x,q}}} \approx 3.7$. Then, we use the Eq. 6.11 and search for t_{FI} to reproduce the observed value of flux growth ξ_{X}. We numerically solve the equation of a form

$$1 + \frac{\int_{E_{\min}}^{E_{\max}} \int_0^{t_{\text{cs}}} G(E, t', \mathbf{t}_{\text{FI}})\, dt'\, dE}{\int_{E_{\min}}^{E_{\max}} F(E)\, dE} = \xi_{\text{X}} \tag{6.23}$$

where $G(E, t', t_{\text{FI}}) = E^s\, K(\bar{\gamma}(E), t', t_{\text{FI}})$, $K(\bar{\gamma}(E), t', t_{\text{FI}})$ is the function under integral in the expression for $N_{\text{e,FI}}(\gamma, t)$ given by the Eq. 6.14, $F(E) = E^s\, N_{\text{e,0}}(\bar{\gamma}(E))$, and the sought parameter t_{FI} is presented in bold.

We evaluate numerically the steady-state spectrum $N_{\rm e,0}(\gamma)$ using the Eq. 6.15 on a Lorentz factor grid. The result agrees very well with the solution for the low-state electron spectrum, obtained with the EMBLEM code in the sub-section 6.2.2, which serves as an additional validation of the code. We fit the samples obtained from numerical integration with a broken power law with an exponential cutoff, which well approximates the steady-state electron distribution in our case. We call this best-fit function in further numerical computations to avoid evaluating the integral 6.15 each time, and calculate the constant low-state X-ray flux, $\int_{E_{\rm min}}^{E_{\rm max}} F(E)\,dE$ appearing in the Eq. 6.23. We then vary the parameter $t_{\rm FI}$ in the Eq. 6.23 until we reach equality between the LHS and the RHS. For all numerical integration computations, we use Python sub-package Integrate in the package SciPy (`scipy.integrate`).

As the result, we retrieve a value of $t_{\rm FI} \approx 1.65\,R_{\rm b}/c \approx 17.8$ d in the blob frame. A very similar value is obtained if we use peak spectral measurement at ~ 3 keV (by XRT), compared to the quiescent one, to which we apply Eq. 6.12. In this case we have to numerically solve a simpler equation of a form

$$1 + \frac{\int_0^{t_{\rm cs}} K(\bar{\gamma}(E), t', \mathbf{t}_{\rm FI})\,dt'}{N_{\rm e,0}(\gamma(E))} = \xi_{\rm X,spec} \tag{6.24}$$

We also find that the inferred value of $t_{\rm FI}$ very weakly depends on the shock transit time $t_{\rm cs}$, as long as it is at the order of $\sim 10^2$ d. In addition, using the Eq. 6.20, one could see that the X-ray flux stagnation should occur after $\simeq 8\,t_{\rm FI} \approx 13\,R_{\rm b}/c$, which is somewhat more than the duration of the shock transit, $t_{\rm cs} = 101.5$ d $\approx 10 R_{\rm b}/c$. This means that the shock wave leaves the emitting region before the X-ray flux would start saturating, and thus the flux increase cessation and the subsequent light curve fall is indeed caused by the shock leaving the blob.

Finally, we use the Eq. 6.12 with the recovered Fermi-I time-scale $t_{\rm FI}$ to predict the spectral flux enhancement at optical V-band frequency, and compare it to the one measured from the optical data. We find that a transient shock, that causes the observed peak flux at X-rays, will produce a flux increase factor $\xi_{\rm opt} \approx 3.3$ at the optical wavelengths, which is much higher than the observed value of ≈ 1.26. Thus, the shock wave that yields the observed amount of X-ray emission at the flare peak, renders a too high optical flux. Thus we conclude that the *one-zone scenario with a passing shock* **does not provide a satisfactory description** of the MWL flare data set. This divergence is demonstrated in Fig. 6.7. The top panel displays the particle spectrum, perturbed by the shock wave, having Fermi-I acceleration time-scale and the transit time derived from the X-ray light curve. The bottom panel of Fig. 6.7 illustrates the SED associated with the electron spectrum perturbed by the

shock with these parameters (red curve). One can clearly see that within the one-zone model with a passing shock, the optical emission appears to be highly overproduced.

We verify that our analytical calculation of the electron spectrum perturbed by a shock (Eq. 6.14) is fully consistent with the numerical solutions obtained with the EMBLEM code.

We also check the admissibility of the δ-approximation by calculating numerically the flux increase factors with our EMBLEM code from the simulated light curves for the X-ray and optical bands ("exact" computation), and comparing them to the ones obtained via the analytical approach (involving the δ-approximation). The numerical value of the X-ray flux increase ratio appears to be ≈ 5.4, and the optical one ≈ 3.1, while the theoretical one's are ≈ 3.7 and ≈ 3.3 respectively. Therefore, we conclude that, while the δ-approximation does not yield extremely accurate values, still it gives quite reasonable results.

Finally, we also point out that the expected optical flux predicted with the Eq. 6.12 or Eq. 6.11, is in fact the lower limit on the flux increase factor. With the full numerical code we find that in case one assumes an escape time-scale longer than $1\,R_{\rm b}/c$ during the shock passage, or takes into account the inverse Compton cooling process, the optical flux will enhance even more, if the X-ray flare will be still reproduced. Therefore the general criterion we developed, allows to estimate the lowest possible flux increase at the optical wavelengths at the flare peak.

Numerical simulations results: shock and turbulence

Following the analytical approach we have already ruled out the one-zone model in which the flare is initiated by a passing shock. Another perturbing process that might be responsible for the production of the outburst within the one-zone scenario, is a stochastic acceleration of the electron population within the emitting blob. The Fermi-II acceleration could be triggered, for instance, by a spontaneously arising turbulence within the emission region due to various physical processes, e.g. hydrodynamical instabilities, streaming instability, plasma waves, etc. This scenario has to be tested with the full numerical EMBLEM code. We also explore a one-zone scenario in which both shock and stochastic acceleration processes are acting on the blob population at the same time. This scenario corresponds to a quite common physical situation where a shock traversing the emitting zone also induces turbulence. Again, we keep the parameters of the low state of the source unaltered during the acceleration phase, including the escape time-scale of $1\,R_{\rm b}/c$.

We vary the Fermi-II time-scale for the scenario with only turbulent acceleration,

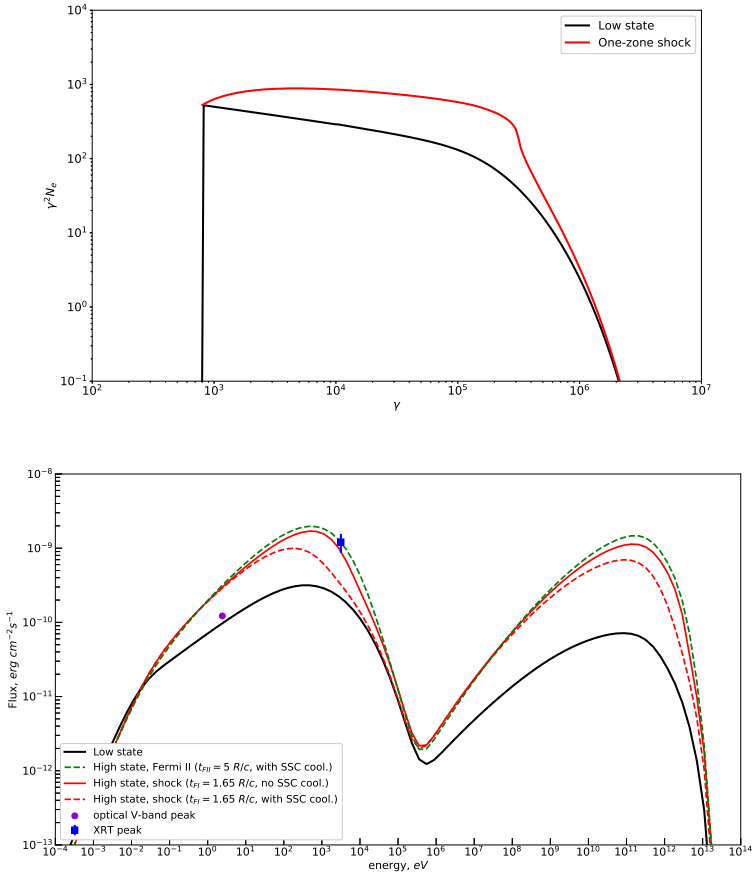


Figure 6.7: Simulated electron spectrum and SEDs for one-zone scenario of the outburst. *Top panel*: electron spectrum perturbed by a shock with $t_{\rm FI} = 1.65\,R_{\rm b}/c$ at the moment of the flare peak, calculated analytically (using the Eq. 6.14), compared to the low-state particle spectrum. Both these spectra do not include the inverse Compton cooling effect. *Bottom panel*: SEDs at the flare peak for the scenario with a transient shock (dashed red curve) and turbulence-induced Fermi-II acceleration (green curve) perturbing the emission region, simulated with the EMBLEM code, together with superimposed optical and X-ray spectral data at the peak of the outburst. Solid red curve represents the SED corresponding to the analytical electron spectrum illustrated in the left panel (full SED computation, inverse Compton cooling neglected), dashed red curve indicates the same model but with full radiative losses including synchrotron and inverse Compton cooling. Green curve shows the SED for the scenario in which stochastic particle acceleration with the time-scale $t_{\rm FII} = 5\,R_{\rm b}/c$ is disturbing the electron population in the emitting zone (inverse Compton cooling is included). The Fermi-II time-scale is adjusted in a way to reproduce the X-ray data at the peak. The black curve displays the low-state SED of Mrk 421. For all the scenarios, the acceleration process in the blob is activated for 101.5 days in the blob frame, which corresponds to 3.5 days in the observer's frame. One can see that all one-zone scenarios reproducing the observed X-ray flare overshoot significantly the optical measurements.

and both Fermi-I and Fermi-II time-scales for the model with a combination of shock and turbulent acceleration. The duration of the acceleration episode $t_{\rm dur,acc}$ is imprinted on the light curve profile and is commensurate with the time-scale of the light curve rise (or time interval before the flux starts to drop) in the source frame. We set the $t_{\rm dur,acc} = 101.5$ d, corresponding to the measured average X-ray light curve rise of 3.5 days (in the frame of the observer), corrected for the relativistic effects. For both models, no set of the two parameters lead to a reasonable fit of the data. For the one-zone scenario in which only stochastic acceleration disturbs the emitting zone, the model describing well the X-ray flux in the high state, overproduces again the optical emission during the peak of the outburst. This mismatch is well visible in the bottom panel of Fig. 6.7 (green curve). On the other hand, we find that the model reproducing the maximal optical flux during the flare, substantially undershoots the X-ray data. If we further consider that the escape time of particles during the Fermi-II acceleration activity can be longer than $1\,R/c$ (according to Eq. 5.34 and Eq. 5.56), it becomes even more difficult to achieve a satisfactory fit: the optical flux excess appears to be even higher, making the discrepancy even worse. The same optical emission overproduction problem remains when combining the Fermi-I and Fermi-II particle acceleration: we did not succeed to find any combination of the two acceleration time-scales $t_{\rm FI}$ and $t_{\rm FII}$ which could reduce the excess of the emission in the optical band, and yield a satisfactory description of the MWL data set in the high flux state.

Therefore, we conclude that *one-zone scenarios* with a moderate perturbation of the steady state of the source by Fermi-I/Fermi-II processes, **are not able to explain** the observed MWL emission characteristics during the outburst, and appear too much constrained to reproduce the observed variability.

6.5.2 Two-zone model

As it appears to be impossible to achieve a reasonable fit of the MWL data within the framework of the one-zone scenarios, we suppose a two-zone configuration. In this model, the low-state and flaring emission are produced in two different, however *physically connected* regions, having some difference in their physical conditions. The issue with the excess of the optical flux during the flare peak, arising in one-zone scenarios, can be resolved if the optical emission during the outburst is dominated by the quiescent emission zone (the emitting blob), and the emission in the X-ray spectral range by a second zone (flaring region). Since the conditions required to ensure production of the radiation up to VHE γ-ray band have to be satisfied in both the low-state and flaring emission regions, their physical characteristics cannot be too

different. We suppose that the quiescent emission originates from a relatively large zone inside the extended jet, for which we assume spherical geometry and homogeneity ("quiescent blob"). The flaring emission is formed in a smaller region, adjacent to the quiescent blob. Both the quiescent and flaring emission zones are traveling along the jet axis at a speed close to the speed of light, and comoving, i.e. having the same Doppler factor.

Yan et al. (2013), studying the Mrk 421 February 2010 flare, reveal that a stationary one-zone SSC scenario with an underlying particle spectrum having a log-parabola form, fits the high-state MWL data set better, than a power law with a cutoff model. In addition, authors find that the observed variability in the VHE γ-ray regime can be only accommodated with a log-parabolic underlying electron spectrum. This means that the turbulent acceleration mechanism, generating electron spectra with a log-parabola shape, appears to be preferable to produce the flaring state, over the shock acceleration process producing a power law with a cutoff particle distribution. Thus, we consider that the Fermi-II acceleration process arising due to turbulence is responsible for launching the flaring activity in the flaring region. Besides, in the study of the outburst by Zheng et al. (2014), the authors ascribed the origin of the flare to a phase of intensified electron injection, supposing also that the particles in the source are accelerated to high energies via stochastic acceleration mechanism. Trying to reproduce the observed data with that model, authors find that a better fit of the spectral and timing characteristics of the flare is obtained assuming the "hard-sphere" turbulence, compared to when the other turbulence types are invoked. Based on that result, we adopt a description of turbulence with a "hard-sphere" ($q = 2$) spectrum.

In a general setting of the considered two-zone flare scenario, a turbulent region suddenly emerges at the interface between the quiescent blob and the surrounding jet medium. The two zones produce emission and may exchange particles due to electrons escaping from one region to another one. We can identify two limiting cases of such a system: a quiescent emission region coupled with either (A) a non-radiative turbulent zone having a considerable flux of escaping electrons into the emitting region, or (B) a radiative turbulent zone with negligible electron escape (see Fig. 6.8). In the first configuration, the radius of the turbulent region has to be similar to the one of the quiescent blob, and the magnetic field has to be much weaker so that only insignificant amount of emission is produced. In the second configuration, on the contrary, the magnetic field has to be comparable to the one in the quiescent emission region, and the size should be much smaller.

We also suppose that above radio frequencies the emission of the large-scale jet is negligible. This could be the case if the jet medium has a substantially weaker magnetic field or a lower density of particles than the blob.

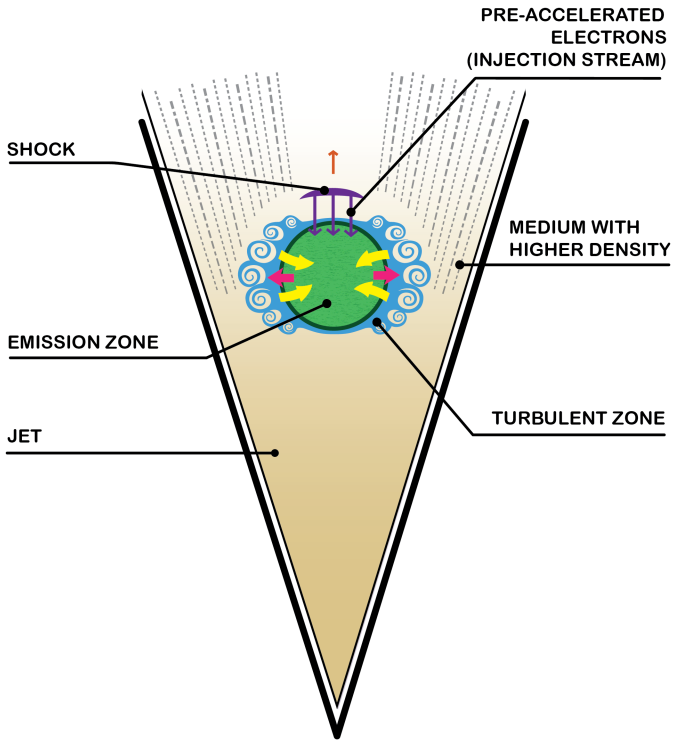


Figure 6.8: Sketch representing a generic two-zone flare model, in which a turbulent region appears around the emitting zone. The gray dashed lines indicate a material with higher density or different speed, disturbing the medium in the vicinity of the quiescent blob and causing the formation of turbulence. The violet curve above the blob illustrates the shock accelerating particles from the up-stream plasma and injecting them into the quiescent emission zone (flux of the injected particles is shown by violet arrows). The quiescent blob and the turbulent zone exchange electrons: ruby-colored arrows depict the injection of particles escaping from the emitting blob to the flaring region, whereas the yellow arrows display the flow of electrons escaping from the turbulent region to the quiescent emission zone. The flux indicated in yellow may be either significant or not, depending on the sizes of the zones and the time-scales of particle escape in each of them. In the best-fit scenario presented in Fig. 6.12 and 6.13, the injection of electrons from the turbulent zone into the blob is negligible.

Scenario A: emitting zone with a non-radiative turbulent acceleration region

Let us first focus on a two-zone scenario A, in which the turbulent region produces only negligible amount of radiation, however has an important flux of escaping particles. The spectral hardening (or equivalently, higher variability amplitude with increasing photon energy) during the flare, discussed at the beginning of this section, appears in this model due to an inflow of electrons with a hard spectrum into the quiescent emission region. This population of particles having a hard spectrum is produced in the turbulent accelerating zone via Fermi-II mechanism.

We consider the following scenario for the flaring behavior. Once the turbulent zone is formed, electrons escaping from the quiescent blob to the turbulent region (represented by red-pink arrows in Fig. 6.8) undergo stochastic re-acceleration. This process hardens the spectrum of the particle population, as well as modifies the position of the cutoff in the spectrum. Re-accelerated electrons escape the turbulent zone, and part of the output particle flux is injected into the emitting region (shown by yellow arrows in Fig. 6.8). In this scenario flaring activity is arising from additional injection of particles in the emitting zone on top of the quiescent injection flux from the upstream shock.

Such scenario is similar to the one proposed by Kirk et al. (1998), but instead of one non-radiative accelerating region, we have two: a shock leading the blob inducing the quiescent emission, and the transient turbulent region inducing the flare.

We model electron acceleration in the turbulent zone and the varying emission of the blob with the EMBLEM code. We assume that only a fraction $f_{\rm inj}$ of particles escaping the turbulent region reaches the blob. The exact value of this fraction is determined by the geometrical configuration of the system "blob with a turbulent zone". Here leave $f_{\rm inj}$ as a free parameter and vary it together with other parameters of the turbulent zone, namely its magnetic field $B_{\rm az}$, size $R_{\rm az}$, Fermi-II and escape time-scales $t_{\rm FII,az}$ and $t_{\rm esc,az}$ respectively, and its life time $t_{\rm l,az}$.

The ranges over which the parameters were allowed to vary are the following:

- $10^{-3} \leq B_{\rm az} \leq 0.04$ G
- $0.1\,R_{\rm b} \leq R_{\rm az} \leq 1\,R_{\rm b}$
- $1\,R_{\rm az}/c \leq t_{\rm FII,az} \leq 100\,R_{\rm az}/c$
- $3 \leq t_{\rm l,az} \leq 5$ d (observ. frame)

Parameter	Symbol	Value
Magnetic Field [G]	$B_{\rm az}$	0.027
Comoving effective size [cm]	$R_{\rm az}$	5.5×10^{15}
Time-scale of the stochastic acceleration	$t_{\rm FII,az}$	$43\,R_{\rm az}/c \approx 91.5$ d (source frame)
Time-scale of the particle escape	$t_{\rm esc,az}$	$18\,R_{\rm az}/c \approx 38.3$ d (source frame)
Life time [d]	$t_{\rm l,az}$	4.65 (observ. frame)
Fraction of particles injected in the blob []	$f_{\rm inj}$	3 %

Table 6.2:: Physical parameters of the non-radiative turbulent region (Two-zone scenario A)

- $1 \leq f_{\rm inj} \leq 100$ %

The resulting best-fit parameters are summarized in Tab. 6.2.

The time evolution of the SED during the flare, with superimposed spectral measurements by VERITAS and H.E.S.S. is depicted in Fig. 6.9. The comparison of the model representations of the light curves and a subset[5] of the MWL flare data is presented in Fig. 6.10. One can see that the *two-zone scenario A* describes well the varying synchrotron emission, however **underpredicts** the γ-ray emission by a factor of ~ 3. Therefore, the two-zone scenario A, although resolving the problem of optical flux excess at the flare peak, and being able to explain the synchrotron flux behavior, appears to be imperfect.

Scenario B: emitting zone with a radiative turbulent acceleration region

Finally, we consider the two-zone scenario B, in which the turbulent region is producing important radiation, but the particle flux from it to the quiescent emitting zone (which we will here call the "*blob*") is subdominant. Electrons escaping from the quiescent blob are injected into the turbulent zone (indicated by red-pink arrows in Fig. 6.8), and subsequently re-accelerated via the stochastic mechanism. The particles in the turbulent region emit synchrotron and IC emission, and the observed flux from the source increases. In this scenario, the hardening of the photon spectrum in the high state happens because the turbulent zone radiates emission with a harder spectrum than the one of the low-state emission, due to the Fermi-II re-acceleration

[5]This modeling was performed in mid-2019 and published in the ICRC proceedings (Dmytriiev et al. 2019a). At that time not all data collected during the flare were reduced or available. A more complete MWL data set is presented in Fig. 6.12 and Fig. 6.13 for the two-zone scenario B.

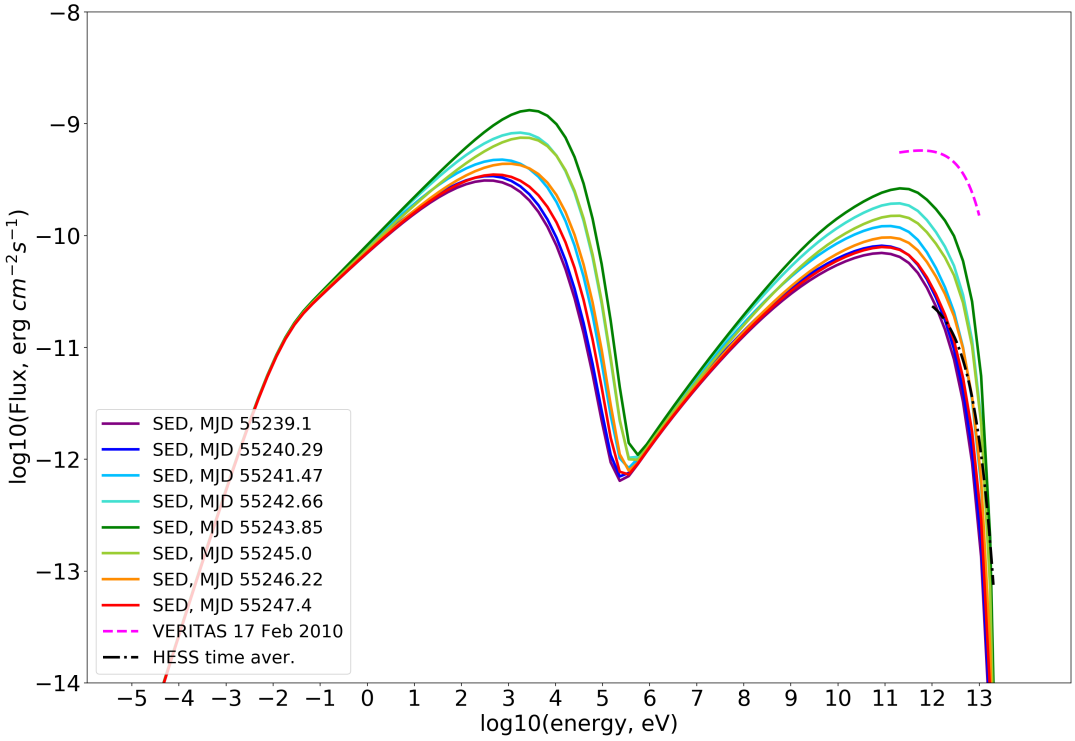


Figure 6.9: Solid lines: modeled time evolution of the SED (advancing from violet to red) during the outburst (two-zone scenario A). Magenta dashed line: spectral measurement by VERITAS during 17 February 2010 (1 day after the flare peak) (Abeysekara et al. 2020). Black dash-dotted line: spectral measurement by H.E.S.S. time-averaged over the period of the flare decay (Tluczykont 2011). One can notice that the model undershoots the data in the γ-ray band.

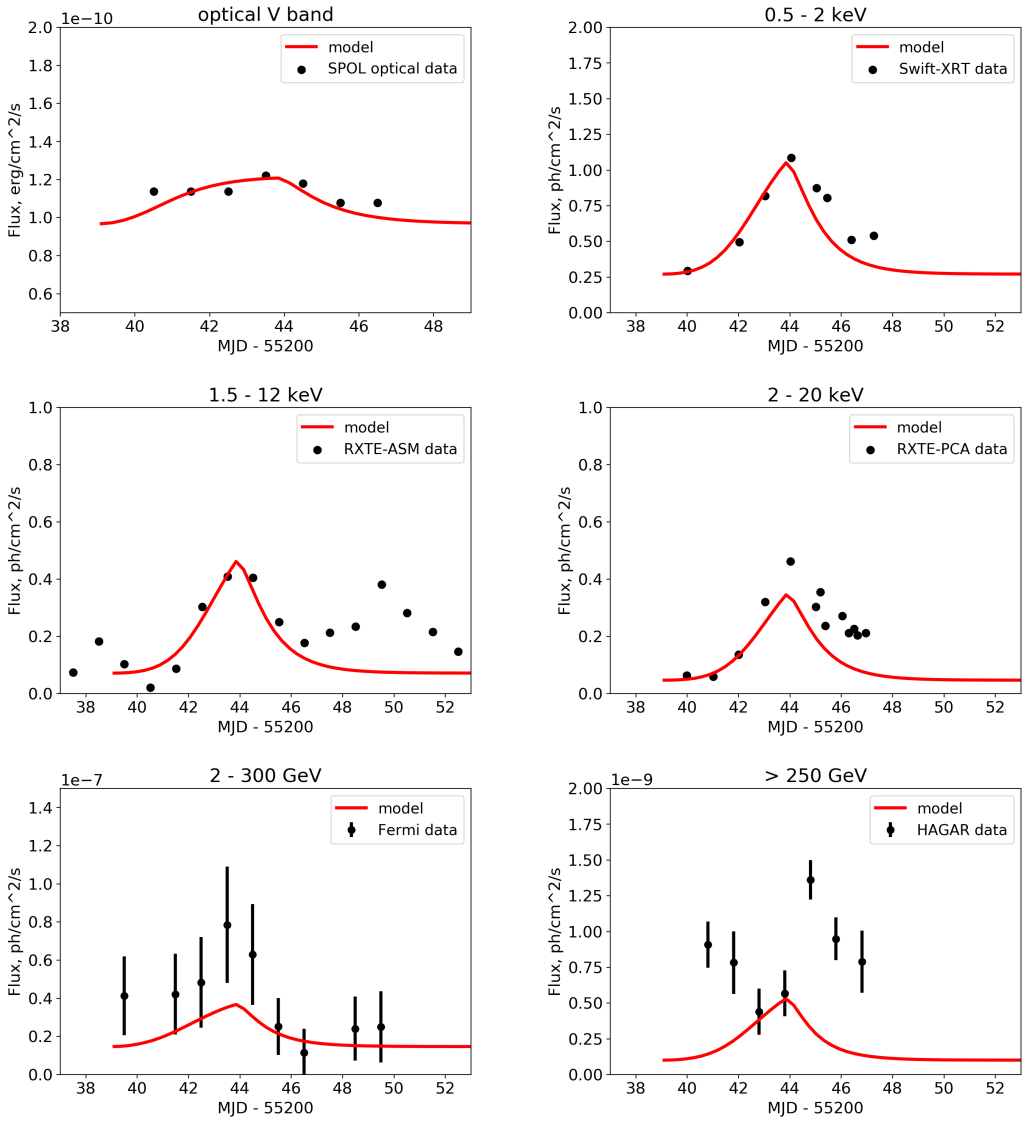


Figure 6.10: Comparison of model multi-band light curves representing the two-zone scenario A and the subset of the flare data. One can clearly see that the model underpredicts the γ-ray flux.

process.

1. Basic assumptions, injection and escape

The emission of the quiescent blob is described with our modeling of the steady state presented in Section 6.2.2 keeping the same physical parameters listed in the Table 6.1. As the blob and the turbulent region are physically connected, several constraints on the physical parameters of the turbulent region exist. As already mentioned, the Doppler factor of the turbulent zone will be equal to the one of the blob, since we suppose that they are moving together as one unit. Since we consider the limiting case where the flux of particles from the turbulent region to the quiescent blob is minor with respect to the injection flux of the steady state, the radius of the turbulent region $R_{\rm tr}$ has to be inferior to the one of the blob $R_{\rm b}$. Next, as the turbulent zone has to provide important contribution to the total emission during the flare, its magnetic field $B_{\rm tr}$ should be strong enough, which means it is expected to be commensurate with the strength of the field inside the blob, or higher: $B_{\rm tr} \sim B \sim 0.04$ G. Also, to be sufficiently bright, the turbulent zone is expected to have a relatively high density, which can be achieved if the escape time of particles from it is much longer than $1\,R_{\rm tr}/c$. This condition will be fulfilled automatically, as the turbulence inside the region impedes fast leaking of particles and confines them stronger than inside the blob (see sub-section 5.1.4 and Eq. 5.56 and 5.58). Finally, the spectrum of electrons injected in the turbulent zone is constant in time and has the same shape as the low-state electron spectrum in the quiescent blob:

$$Q_{\rm inj,qr\text{-}tr}(\gamma) \simeq \frac{N_{\rm e,0}(\gamma)}{t_{\rm esc}} \left(\frac{R_{\rm b}}{R_{\rm tr}}\right)^3 \cdot Y_{\rm qr\text{-}tr} \tag{6.25}$$

where $N_{\rm e,0}(\gamma)$ is the steady-state electron spectrum of the quiescent blob (sub-section 6.2.2, and Eq. 6.15), $t_{\rm esc} = 1\,R_{\rm b}/c$ is the time-scale of particle escape from the blob (Table 6.1), and $Y_{\rm qr\text{-}tr}$ is the share of particles that after escaping from the blob are injected into the turbulent zone.

Assuming homogeneity of the escaping and injected particle fluxes, the fraction $Y_{\rm qr\text{-}tr}$ is defined by the geometrical configuration of the two regions. Here we do not leave this parameter free (contrary to the modeling presented for the two-zone scenario A), and estimate it as follows. The turbulent region is assumed to have a form of a thick turbulent torus around the blob. For practical reasons we adopt here a simplified layout of the system in which the turbulent zone is composed of a few identical small spherical zones or "*eddies*" at the lateral edge of the quiescent blob. The radius of one eddy $R_{\rm ed}$ is related to the effective size of the turbulent region $R_{\rm tr}$

via $R_{\rm tr}^3 = 4\,R_{\rm ed}^3$. The contribution of each small eddy to $Y_{\text{qr-tr}}$ is determined by the solid angle Ω enclosing it:

$$Y_{\text{qr-tr}} = \frac{4\times\Omega}{4\pi} \simeq \frac{1}{4\pi}\,4\,\frac{\pi R_{\rm ed}^2}{R_{\rm b}^2} = \left(\frac{R_{\rm ed}}{R_{\rm b}}\right)^2 \tag{6.26}$$

Substituting this expression into the Eq. 6.25, one gets:

$$Q_{\text{inj,qr-tr}}(\gamma) = \frac{N_{\rm e,0}(\gamma)}{4^{2/3}\,t_{\rm esc}}\,\frac{R_{\rm b}}{R_{\rm tr}} = \frac{N_{\rm e,0}(\gamma)}{4\,t_{\rm esc}}\,\frac{R_{\rm b}}{R_{\rm ed}} \tag{6.27}$$

As the electron spectrum inside the quiescent blob remains unchanged during the flaring event (we assumed negligible particle flux flowing from the turbulent zone towards the blob), the injection of escaping electrons in the flaring region happens at a constant rate.

2. Turbulent acceleration and its evolution

Let us now consider Fermi-II acceleration of particles in the small turbulent eddy. From the Eq. 5.58 the time-scale of electron escape from a turbulent eddy linearly increases with an increasing energy density of magnetic field fluctuations $t_{\rm esc,ed} \propto \frac{\delta B^2}{2\mu_0}$. The stochastic acceleration time-scale is inversely proportional to the fluctuations energy density and to the square of the Alfvén speed: $t_{\rm FII} \propto \frac{1}{\beta_{\rm A}^2}\,\frac{2\mu_0}{\delta B^2}$ (see Eq. 5.32). The Alfvén speed is not constant in time and varies due to time-dependent electron density in the turbulent eddy. Therefore, in order to complete the time-dependent description of the turbulent acceleration and particle escape, one needs a time profile for the evolution of the turbulence, $\delta B^2 = \delta B^2(t)$.

During the stochastic acceleration process, particles gain energy by extracting it from the turbulent motions of the plasma. The energy density stored in magnetic field fluctuations at a certain moment of time is thus governed by the equilibrium between the injection of the turbulent energy in the system and its dissipation due to work it has done on the acceleration of electrons. In case one neglects other losses (e.g. Alfvén wave damping, etc.), as well as assumes that losses do not depend on wavenumber, the equation describing the time evolution of the energy contained in turbulence, $U_{\rm turb}(t) = \delta B^2(t)/(2\mu_0)$ has the following form (Burn 1975):

$$\frac{dU_{\rm turb}}{dt} \;=\; Q_{\rm turb}(t) - \int_{\gamma_{\rm min}}^{\gamma_{\rm max}} \frac{2\gamma m_{\rm e}c^2}{t_{\rm FII}(t)} N_{\rm e,tr}(\gamma,t)\,d\gamma \;=\; Q_{\rm turb}(t) - \frac{2\,\varepsilon(t)}{t_{\rm FII}(t)} \tag{6.28}$$

The integral on the RHS on this equation represents the energy gain by electrons being accelerated by the Fermi-II mechanism. It is obtained by integration of the term in the kinetic equation (Eq. 5.2), describing the systematic energy gain due to stochastic acceleration. $Q_{\rm turb}(t)$ is the time-dependent rate of injection of turbulent energy in the system, $N_{\rm e,tr}(\gamma, t)$ and ε are the electron spectrum and energy density of electrons in the turbulent region (which are the same as in one eddy). Plugging the expression for the Alfvén speed depending on the electron energy density (Eq. 5.34), and the relation for the $t_{\rm FII}$ depending on the parameters of the medium (Eq. 5.32), we recover for the loss term:

$$\left(\frac{dU_{\rm turb}}{dt}\right)_{\rm loss} \approx \frac{U_{\rm turb}}{\lambda_{\rm max}/c} \tag{6.29}$$

The form of the term describing the injection of the turbulent energy $Q_{\rm turb}(t)$ depends on the physical model of turbulence development. Detailed treatment of turbulence generation is beyond the scope of our modeling. For this reason, we try to parametrize this term in a simple way with a minimal number of free parameters. The simplest form we first consider is a step function, i.e continuous injection of turbulent energy with a constant rate during a certain time interval. In this case, after the injection has started, the density of turbulent energy experiences an increase from zero to a constant maximum level over a time-scale of $t_{\rm turb} \approx \lambda_{\rm max}/c$. The turbulent energy density in the region remains constant as long as the turbulent energy is supplied (at a constant rate). Once the injection stops, electrons extract all available turbulent energy contained in the region at the same time-scale $t_{\rm turb}$, and the turbulent energy density decays exponentially. The build-up and dissipation time-scale $t_{\rm turb}$ corresponds to the formation/decay time of the longest mode in the turbulence spectrum $\lambda_{\rm max}$, and represents the shortest possible time for the turbulence to develop or to decay. However such simple temporal profile for injection cannot be used to reproduce the observed light curves: we find that the flares simulated with such constant turbulent injection exhibit a protracted plateau instead of a sharp peak. Therefore we conclude that such time profiles are not suitable to describe the observed shape of the light curves, and consider a more complex parametrization.

We suppose a more realistic time-dependent profile of the turbulent energy injection with a linear rise and decay, occurring gradually, i.e. on time-scales $t_{\rm turb,r}$ and $t_{\rm turb,d}$ respectively, which are longer than $\lambda_{\rm max}/c$. Such injection rate evolution may be caused by e.g. the emitting region passing through a dense zone with a spatial gradient of density. The rise and decay times of the turbulence injection rate are expected to be of the order of the rise and fall time-scales seen in the flare light curves, corrected for the relativistic effects. As the time-scales of the rise and decline of the

injection rate are significantly longer than $\lambda_{\rm max}/c$, we can neglect the boundary effects (build-up and decay of the turbulence over time $\sim \lambda_{\rm max}/c$), and so the profile of time evolution of the turbulent energy density $\delta B^2(t)/(2\mu_0)$ approximately mimics the temporal profile of the injection function $Q_{\rm turb}(t)$. We also assume that the maximum value of the turbulent energy density, achieved at the moment of peak injection rate, is approximately equal to the energy density of the non-turbulent ordered component of the magnetic field in the eddies, $\delta B^2|_{\rm peak} \sim B_{\rm tr}^2$.

3. SSC emission

We now model the varying emission of the turbulent eddies using the temporal profile of the injection of turbulence discussed above, and the Eq. 6.27 as the electron injection spectrum during the turbulent phase. After the turbulence has completely decayed (at the moment when δB^2 falls to zero), the region dissolves in the ambient jet medium and its magnetic field diffuses. At this point we can consider that the particles escaping from the blob are no longer injected into the leftover, as they are simply crossing it. Therefore we simulate the disintegration of the turbulent region by simply stopping the injection of electrons into it. After that, the particles in the region cool and escape from it, and its observed emission fades relatively quickly (over $min(t_{\rm cool}(\gamma), t_{\rm esc,tr})$), so we do not model the dissipation of its magnetic field, keeping it constant throughout the entire evolution, as after the turbulent energy fell to zero, the radiation of the turbulent region becomes already subdominant, and the exact way its flux falls after that is not important. In order to have a self-consistent description, when solving the kinetic equation (Eq. 5.2) with the EMBLEM code, we calculate at each time step the time-dependent particle energy density and the Alfvén speed with the help of the Eq. 5.34, as well as the escape and stochastic acceleration time-scales using the Eq. 5.58 and Eq. 5.32 respectively. We consider that the electrons escape from a turbulent eddy in a free streaming mode at the very beginning and the very end of the turbulence injection, at a time-scale $t_{\rm esc,ed,0} = 1\, R_{\rm ed}/c$, while during the time the eddy is turbulent ($\delta B^2 \neq 0$) the time evolution of the escape time-scale $t_{\rm esc,ed}(t)$ replicates the behavior of the turbulence level. The evolution of the Fermi-II time-scale in time has a more complex profile than the escape time-scale due to the dependency on the inverse square of the time-dependent Alfvén speed.

4. External Compton contribution

When modeling several emission regions located nearby, it is necessary to take into account the external Compton (EC) effects. In our particular configuration, the flaring emission is scattered off the particles in the quiescent blob, and the quiescent

emission interacts with the electron population in the turbulent zone. Estimating contributions of each process, we find that the latter effect is rather significant for our conditions, while the former one can be neglected. Qualitatively, the first effect is small because the emission of the turbulent zone is significantly diluted when seen in the quiescent blob by a factor $\sim 1/20$ (post-verification after the best-fit model was found). Therefore, we take into account only the IC scattering of the quiescent blob synchrotron radiation off the electron population in the turbulent zone.

We thus treat the synchrotron emission of the quiescent blob as an external radiation field for the turbulent region, and use the sum of the synchrotron emission of a turbulent eddy and of the steady-state emission region as the seed photon field when computing the IC emission produced by the turbulent eddy, and the IC cooling rate, which is enhanced due to the presence of the external photon field. Compared to the case where this cross-scattering is not taken into account, inclusion of this effect causes an average amplification of the total GeV-to-TeV γ-ray flux level by $\sim 40\%$ and a drop of the total flux in soft-to-hard X-ray range by a comparable value. The effect is therefore quite significant.

5. Final modeling of the flare

The free parameters of the model are the *parameters of the turbulent region*:

1. Radius $R_{\rm tr}$
2. Magnetic field strength $B_{\rm tr}$
3. The longest wavelength in the turbulence spectrum $\lambda_{\rm max}$ (controlling the time-scale of electron escape at the moment of the turbulence peak)
4. Rise time of the profile of the turbulent energy injection rate $t_{\rm turb,r}$
5. Decay time of the profile of the turbulent energy injection rate $t_{\rm turb,d}$

We tune these five parameters in a way, that the total emission, which is the sum of the constant emission of the quiescent blob and the variable emission of the turbulent region matches the observed MWL emission given by the data set.

The ranges over which the parameters were allowed to vary are the following:

- $0.03\,R_{\rm b} \leq R_{\rm tr} \leq 0.3\,R_{\rm b}$
- $0.04 \leq B_{\rm tr} \leq 0.2$ G

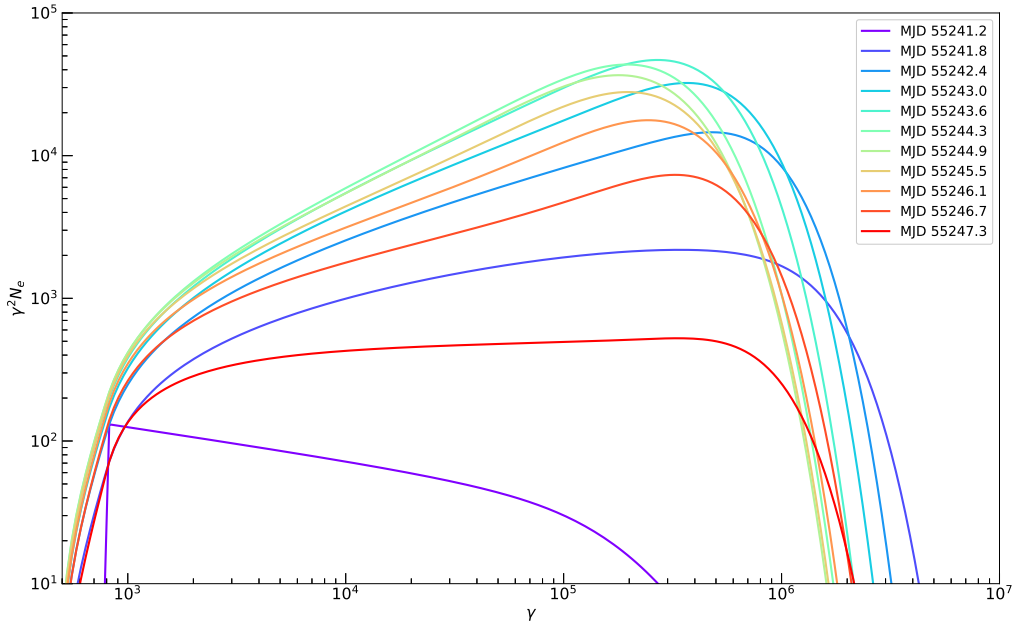


Figure 6.11: Simulated time evolution of the electron spectrum in the turbulent region during the Mrk 421 February 2010 flaring event. The electron spectrum is evolving from violet to red curves. The evolution is presented with a time step of ∼0.6 d.

- $0.01\,R_{\rm tr} \leq \lambda_{\rm max} \leq 1.0\,R_{\rm tr}$
- $2.5 \times 10^6 \leq t_{\rm turb,r} \leq 10^7$ s (source frame)
- $7.5 \times 10^6 \leq t_{\rm turb,d} \leq 1.5 \times 10^7$ s (source frame)

The evolution of the electron spectrum in time in the turbulent region is shown in Fig. 6.11. The process of electron acceleration leaves a characteristic signature on the temporal behavior of the particle distribution in a form of a pronounced spectral hardening, as well as in a noticeable growth of the maximum electron energy with time. One can also notice a steep fall of the particle spectra at the low Lorentz factors, near the minimal value. This effect occurs due to migration of low-energy electrons to high Lorentz factors during the acceleration phase. As the supply of the turbulent energy starts to decline, the efficiency of the acceleration process drops, leading to softening of the spectral slope and a fall in maximum electron energy due to prevailing electron escape and cooling losses.

Parameters of the February 2010 flare	Symbol	Value
Magnetic Field [G]	$B_{\rm tr}$	0.05
Comoving effective size of the turbulent region [cm]	$R_{\rm tr}$	3.65×10^{15}
Maximal wavelength in the turbulence power spectrum	$\lambda_{\rm max}$	$0.023\, R_{\rm tr}$
Duration of the rise of the turbul. energy inj. rate (source frame) [s]	$t_{\rm turb,r}$	5×10^6
Duration of the decay of the turbul. energy inj. rate (source frame) [s]	$t_{\rm turb,d}$	10^7

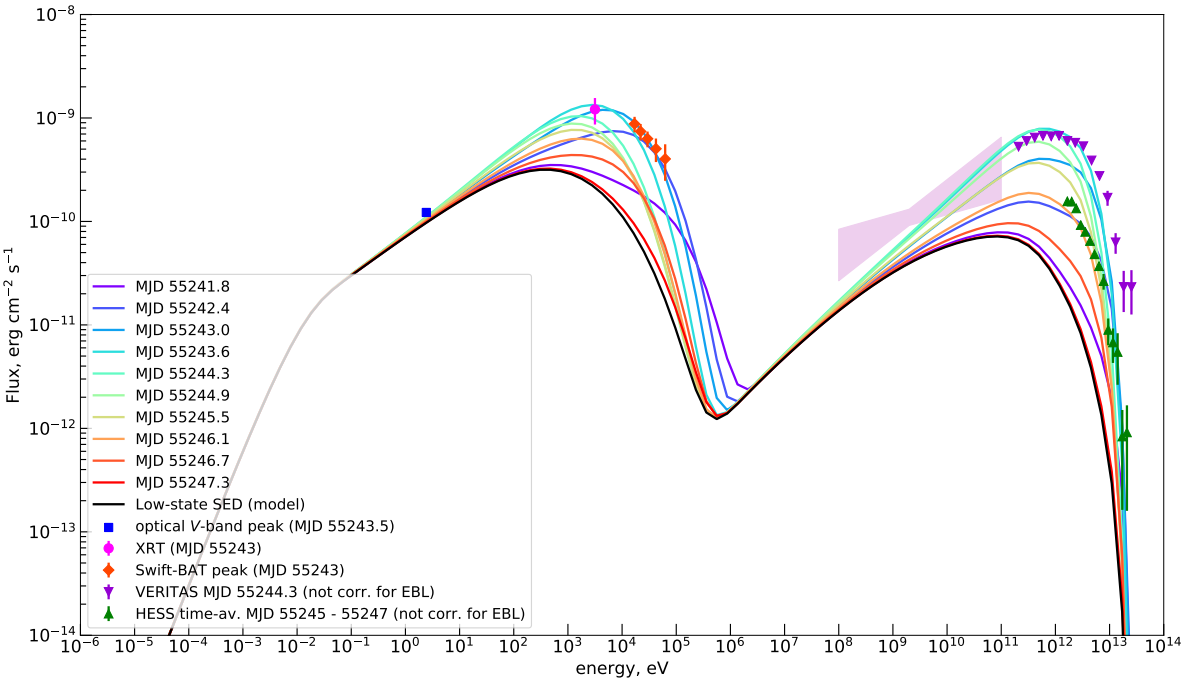
Table 6.3:: Physical parameters of the turbulent region and of the flaring state in the two-zone scenario B.

For the best-fit model, the simulated time evolution of the broad-band SED of the total emission with superimposed spectral measurements from the MWL data set, is illustrated in Fig. 6.12. The overall description of the available spectral information is very reasonable. As one can see, the two-zone model with a radiative turbulent zone is able to reproduce quite well the observed flux enhancement depending on the photon energy during the flare, giving a correct prediction of a strong flare in X-ray and γ-ray spectral domains, and very slight flux magnification at the optical wavelengths. Only in the range 0.1-1 GeV the model somewhat underestimates the variability level.

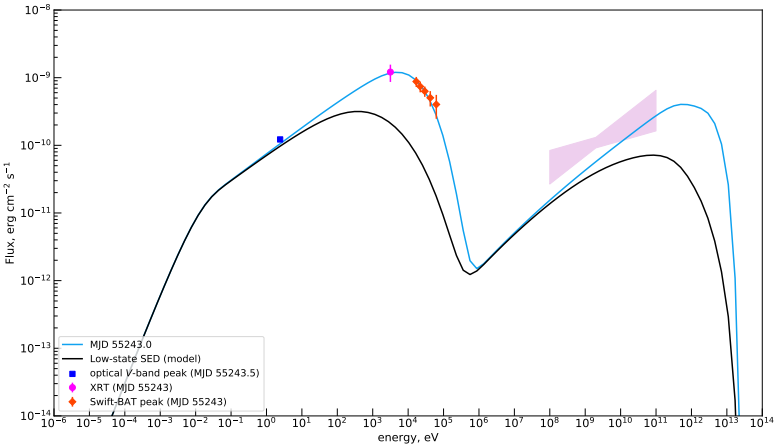
Next, we compare the model representations of the light curves to the MWL flux variability measurements (see Fig. 6.13). The model reproduces quite well the observed behavior of the varying multi-band flux during the flare, when considering the major (first) flaring event and disregarding the weaker second one. In particular, a "shoulder" seen in the X-ray data between the two events, arises naturally in the turbulent acceleration model that we discuss in the Section 6.6. Our scenario also predicts rather accurately the position of the peak of the outburst in different energy bands, including the appearance of the soft X-ray flux maximum ~ 1 day later than the one in the hard X-ray range. The simulated optical flux variability shows a correct amplitude, however a slight offset is present in the pre-burst (benchmark) flux level.

Table 6.3 summarizes the parameters of the turbulent region. The effective radius of the turbulent region is found to be an order of magnitude smaller than the one of the steady-state emission region. The magnetic field in the turbulent zone appeared to be comparable to the one in the quiescent blob.

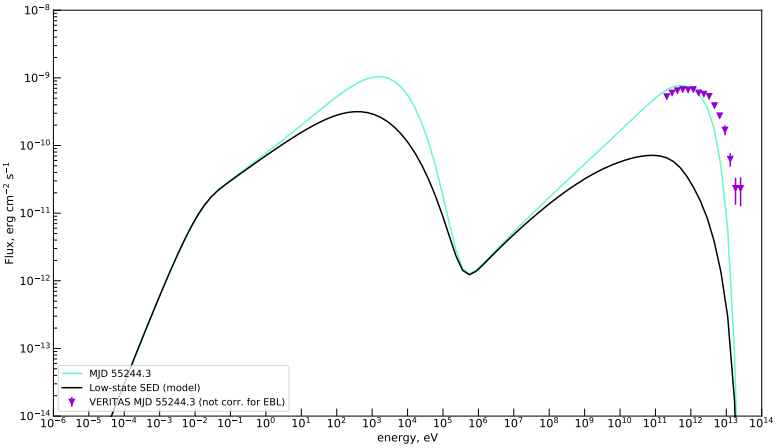
We need now to verify whether the best-fit parameters are realistic. The maximal wavelength in the turbulence spectrum should not be longer than the spatial extent of the turbulent region, $\lambda_{\rm max} \leq R_{\rm tr}$, since the modes of the turbulent cascade are limited by the boundaries of the zone permeated by turbulence, i.e. the turbulent zone. In addition, from the Eq. 5.58 one can see that $\lambda_{\rm max} > R_{\rm tr}$ will lead to superluminal



MJD 55243.0



MJD 55244.3



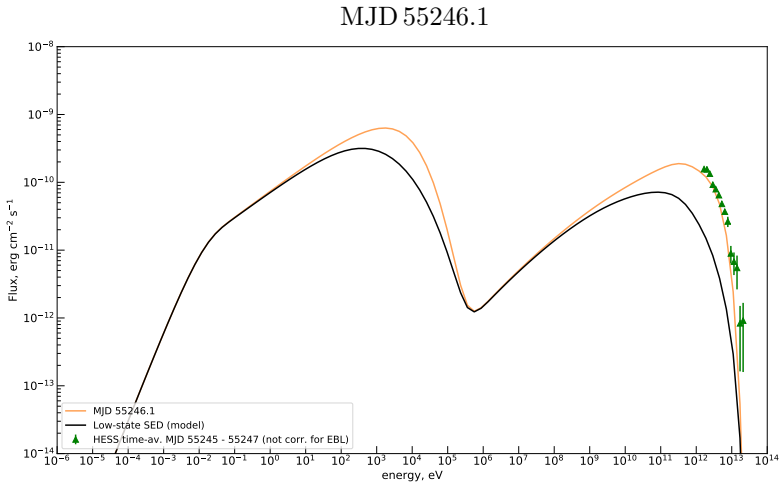
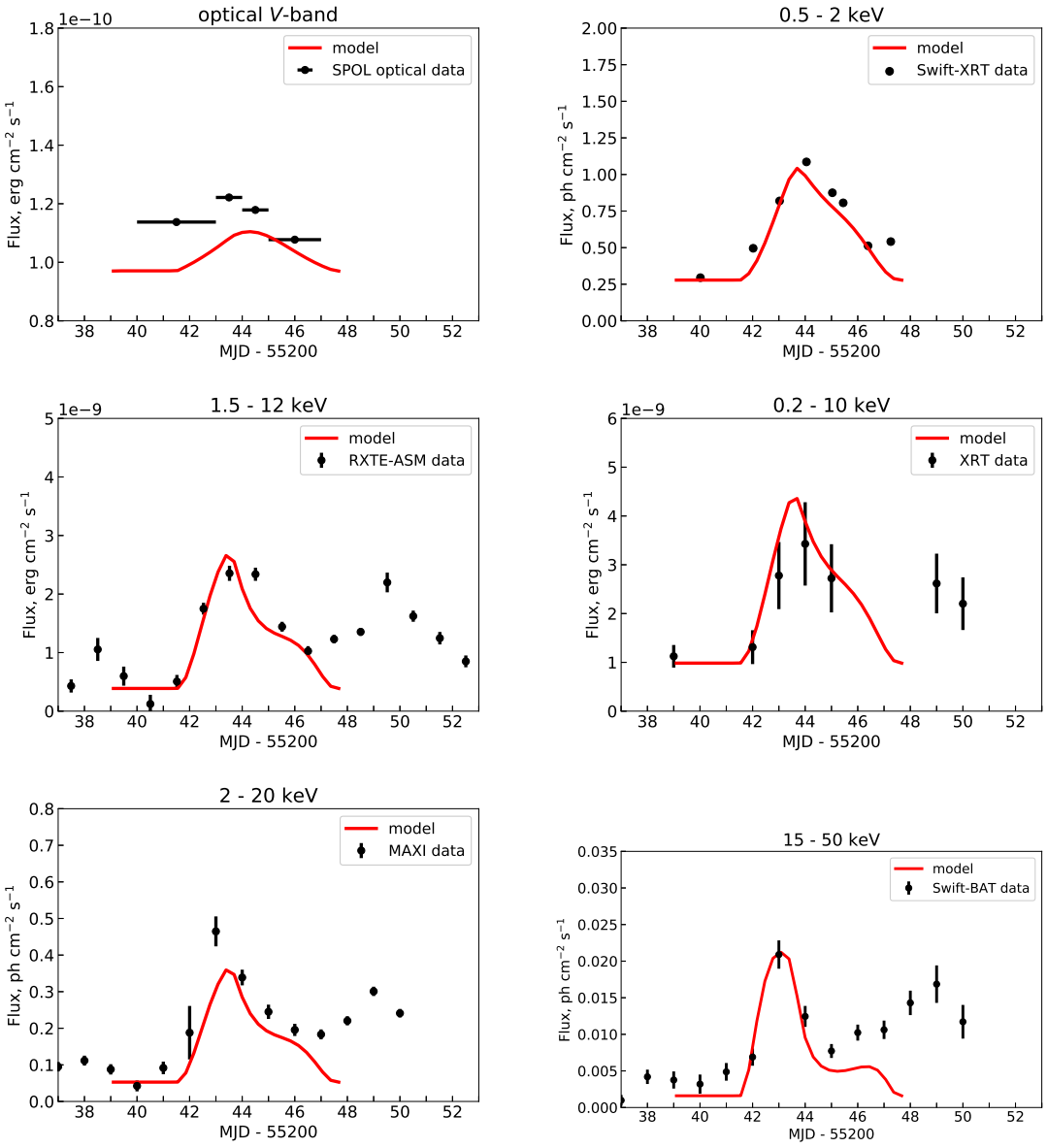


Figure 6.12: Simulated time evolution (from violet to red curves) of the broad-band SED of Mrk 421 during its February 2010 flare, with the spectral measurements from the data set of the flare superimposed for comparison (4 panels on 2 pages). *Top panel*: full SED evolution, illustrated with a time step of ∼0.6 d. *Bottom panel*: comparison of the model SED with spectral data for MJD 55243.0, MJD 55244.3 and MJD 55246.1. The model SEDs are absorbed on the EBL using the model by Domínguez et al. (2011). The black curve indicates the SED model of the low-state of the source. The blue square point displays the optical flux during the peak of the flare, the magenta circular point – the XRT flux at ∼3 keV (16 February 2010, MJD 55243), the red diamond points – the *Swift*-BAT spectral data during the flare peak (16 February 2010, MJD 55243), the violet down-pointing triangle points – the VERITAS spectral measurement (17 February 2010, MJD 55244.3) (not corrected for EBL), the green up-pointing triangle points – the H.E.S.S. SED during the fall of the flare, time-averaged over the period 17-20 February 2010 (MJD 55245.0 – 55247.0) (not corrected for EBL). The pink butterfly corresponds to the *Fermi*-LAT uncertainty band for the SED at the flare peak (16 February 2010, MJD 55243). Optical data (host galaxy subtracted) is derived from Shukla et al. (2012), VERITAS spectral measurement from Fortson et al. (2012), H.E.S.S. data from Tluczykont (2011). The spectral data of XRT, *Swift*-BAT and *Fermi*-LAT are taken from Singh et al. (2015).


optical V-band
1e−10
model
SPOL optical data
Flux, erg cm^{-2} s^{-1}
MJD - 55200
0.5 - 2 keV
model
Swift-XRT data
Flux, ph cm^{-2} s^{-1}
MJD - 55200
1.5 - 12 keV
1e−9
model
RXTE-ASM data
Flux, erg cm^{-2} s^{-1}
MJD - 55200
0.2 - 10 keV
1e−9
model
XRT data
Flux, erg cm^{-2} s^{-1}
MJD - 55200
2 - 20 keV
model
MAXI data
Flux, ph cm^{-2} s^{-1}
MJD - 55200
15 - 50 keV
model
Swift-BAT data
Flux, ph cm^{-2} s^{-1}
MJD - 55200

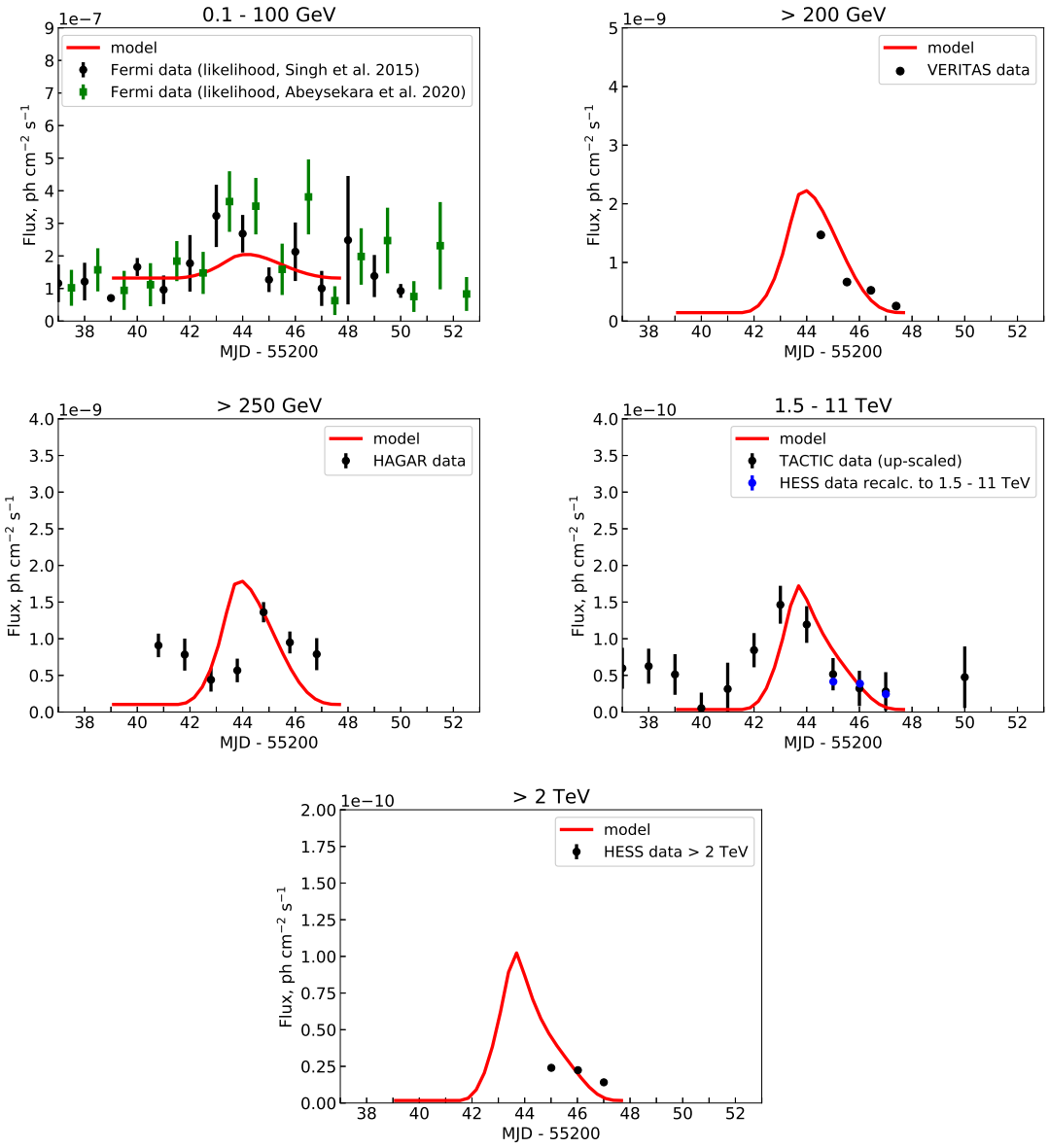


Figure 6.13: Comparison of the simulated light curves representing the two-zone model (B) to the observational data (11 panels on two pages). The set of multi-band light curves includes X-ray light curves by XRT and MAXI (Singh et al. 2015), *Swift*-XRT and RXTE-ASM (Shukla et al. 2012), the *Fermi*-LAT light curve (Singh et al. (2015) ; Abeysekara et al. (2020)), and light curves in the VHE regime by H.E.S.S. (Tluczykont 2011), HAGAR (Shukla et al. 2012), TACTIC (Singh et al. 2015), and VERITAS (Abeysekara et al. 2020). The optical flux time evolution (host galaxy subtracted) is derived from Shukla et al. (2012).

escape of particles. In our best solution, the maximal Alfvén wavelength $\lambda_{\rm max} \approx 0.023 R_{\rm tr}$, which is much smaller than the size of the turbulent region. Moreover, this value represents an upper limit on the maximal wavelength. As we supposed the strongest turbulent level is reached at the peak during the turbulence, $\left(\frac{\delta B}{B_{\rm tr}}\right)^2 = 1$, from the Eq. 5.58 one can see that with lower turbulence levels the $\lambda_{\rm max}$ would be even smaller than the value we obtained under our assumption.

As a final step, we can check whether our initial assumption that the flow of electrons escaping from the turbulent zone to the emitting blob is indeed negligible in the conditions we derived as the result of the fit. The corresponding volumetric electron injection rate under assumption that half of particles leaving the turbulent zone will arrive in the quiescent blob is:

$$Q_{\rm inj,tr\text{-}qr}(\gamma, t) \simeq \frac{N_{\rm e,tr}(\gamma, t)}{2 t_{\rm esc,ed}(t)} \left(\frac{R_{\rm tr}}{R_{\rm b}}\right)^3 \tag{6.30}$$

Using this relation, we establish that around the peak of the outburst, the spectrum of additionally inflowing electrons is an order of magnitude lower than the one producing the low-state emission (generated by the shock leading the blob). Thus, the particle injection into the blob, driven by the flaring zone, indeed has an insignificant contribution in the overall injection rate, which is coherent with the two-zone scenario we developed.

6.6 Discussion and perspective

In this chapter, we have developed a general analytical approach to establish feasibility or dismiss a one-zone model with a passing shock to launch flaring events in blazars. It was demonstrated that the MWL data set of the February 2010 flare of Mrk 421 cannot be reproduced with such one-zone shock scenario, and neither with turbulent re-acceleration in a one-zone model.

A self-consistent *two-zone model*, with a large emission zone responsible for the quiescent emission and a smaller, physically connected turbulent region producing the flaring emission, with the flare driven by Fermi-II acceleration process, yields a very reasonable description of the available MWL spectra and light curves for this event. The observed spectral hardening in the high state and asymmetric light curve profiles appear naturally in the simulation of the processes during the outburst, including electron acceleration, radiative cooling and particle escape.

In this model, the low-state emission and flare emission can be connected with a minimum number of free parameters, and the flare arises naturally as a perturbation on the edge of the quiescent emission region.

Although the best fit of the February 2010 MWL data set is still not perfect in all energy bands, the overall description of the flare is very reasonable. Let us consider different energy ranges in more detail.

X-ray band

The proposed two-zone scenario provides a very accurate representation of all the X-ray data of the main flare. The model reproduces well the detailed shapes of the X-ray light curves, as well as the ∼1 day time lag seen between the hard and soft X-ray flares (see Fig. 6.13). The varying turbulence injection and in particular the gradual fading of the supplied energy cause a remarkable feature in the light curves: a flux stagnation occurring during the fall of the flare before its end. This effect is clearly visible in the X-ray range and also manifests, to a lesser extent, at higher energies. The cause of such effect is a more rapid electron escape once the turbulence level drops substantially. Due to the faster particle leak, the energy density (contained in particles) in the turbulent region declines, and the Alfvén speed grows and balances the diminution of the turbulent energy density in the expression for the Fermi-II acceleration time-scale (see the Eq. 5.32). This stabilizes the stochastic acceleration time-scale temporarily, the mild particle acceleration balances cooling and escape, and the flux "freezes" at a nearly constant level for a while. This behavior goes on until the decay of the turbulence starts to dominate and the Fermi-II acceleration starts to sharply lose its efficiency. In the data set, three X-ray light curves show signs of such a feature, appearing just before the secondary flaring event (starting at MJD 55246 – 55247), the presence of which, however, prevents us to reliably detect this flux stagnation effect in the data. The evolution of the X-ray flux attaining the shoulder, is quite well described by the model, but the Alfvén speed growth appears to be insufficient to trigger the secondary flaring event. Hence, we assume that the second flare might be initiated by a "second wave" of turbulent energy input, more moderate in the total energy budget. The supposed "second wave" of turbulence injection could be induced by e.g. passage through another cloud of dense material, or by growth of a second, "echo" instability mode. We, however do not model the appearance of the secondary flare here.

Soft γ-ray band

The best-fit model slightly undershoots the *Fermi*-LAT spectral measurements. A possible way to achieve a better description of the measured 0.1-1 GeV flux (Fig. 6.12) without modifying its synchrotron counterpart (optical flux) is to assume

a presence of an additional external Compton component. In our case, the effect of the interaction of flaring synchrotron emission with the particles residing in the quiescent blob appears to be negligible (for the best-fit physical parameters). However, for a different set of the parameters, this process could become quite important and contribute significantly to the total flux in the GeV band. For example, in the case of a larger size of the turbulent region, the density of flaring radiation seen in the quiescent blob will be higher. Also, in this situation, the escaping particle flux from the turbulent region to the quiescent blob could become non-negligible and has to be properly taken into account. However, a full new modeling is required to test our assumptions and to check whether these additional effects can indeed foster an increased GeV emission production. For that, one has to include these processes into the generic two-zone model, find a new solution with a different set of physical parameters, and examine whether a better representation of the GeV data has been achieved (together with an adequate fit in the other energy bands).

VHE γ-ray band

In the VHE γ-ray band, the model appears to be consistent with the data, although does not describe in a very accurate manner all the details of the light curves. We also did not attempt to model the intra-night variability, reported by VERITAS. One can reasonably assume that the flux variations proceeding at the 1 h time-scale might originate from compact flaring subregions in the composite turbulent zone. In our approach, for simplicity, we modeled it with four independent spherical eddies, while in reality the turbulent region could consist of a large number of individual small cells with a random magnetic field direction and constantly fluctuating particle density and velocity fields (Marscher 2014). Such more complex structure of the turbulent zone could explain the stochastic properties of the VHE flux.

Optical band

The model shows a slight discrepancy with the optical light curve at the level of $\sim 10\%$. However, there might be a systematic shift affecting (1) the optical data and caused by uncertainties in the subtraction of the flux of the host galaxy, and/or (2) the modeling of the quiescent net optical flux, arising due to non-negligible scatter of the low-state optical measurements (see Fig. 6.3). We avoid taking the values of the optical flux at the pre-burst and post-burst stage (Fig. 6.13) as the quiescent flux estimate, as it is not clear from the light curve whether the flux starts or drops from/to the quiescent level, as the light curve shows too few points and large horizontal error bars. Thus, as we reproduce the observed variability amplitude and the overall offset is not significant, our description of the optical measurements can be considered as reasonable.

Complexity of the physics involved in the scenario

The turbulent particle acceleration was previously considered by different authors as the mechanism launching HBLs flares. Within this scenario, one can adopt different turbulence types. As an example, Tramacere et al. (2011) study six HBLs flares and reproduce various tendencies observed in the data invoking Fermi-II particle acceleration occurring on relatively short time-scales. The authors also find that, in fact, two acceleration scenarios are able to reasonably describe the available X-ray data: (*i*) the stochastic acceleration time-scale is varying due to time-dependent δB^2 or $\beta_{\rm A}$, together with the constant hard-sphere spectrum of the turbulence, or (*ii*) the index of the turbulence spectrum is changing with time. In our modeling, we adhere to the first option, using however a more complex coherent description of the evolution of the energy density of magnetic field fluctuations δB^2 and of the Alfvén speed $\beta_{\rm A}$, and also considering a two-zone configuration.

There are several processes that can induce turbulence near the quiescent blob, that is considered to move relativistically along the jet axis. One possible mechanism is a spontaneous excitation of Kelvin-Helmholtz or rotationally-induced Rayleigh-Taylor instability (Meliani & Keppens 2009) at the interface between the faster inner spine represented by the quiescent blob, and the slower outer sheath of the jet (a so-called spine-sheath structure of the jet (Sol et al. 1989)). Another scenario in which the turbulence could be triggered, is when the blob is crossing a dense gas cloud, which may be formed e.g. as a result of an interaction between a red giant star and the jet (Barkov et al. 2012). During the passage, the density of the plasma flowing past the blob increases, and the Reynolds number could surpass its critical value, leading to generation of a transient turbulent zone around the blob. The plasma instabilities and various complex physical processes that lead to a sudden excitation of the transient turbulence in the vicinity of the blob, are impossible to simulate in relatively simple radiative models as the one we have developed. As already mentioned, we adopt a very simplified, but at the same time seeming to be fairly realistic profile for the evolution of the turbulent energy injection rate, describing it with an ordinary linear rise and fall. Nonetheless, by performing different trials, we find that the exact form of the turbulence injection profile does not have a strong effect on the shape of the simulated light curves, as long as the time-dependent injection function represents a kind of a bump, having appropriate rise and decay time-scales. So, at the end, a very simplistic modeling of the turbulence is still able to satisfactorily reproduce the observed MWL variability pattern during the flare.

Alternative two-zones models

As one-zone scenarios have been firmly ruled out for the February 2010 flare

of Mrk 421, a two-zone model was adopted in this work. We constructed here the two-zone scenario which seemed to be the most promising to us, taking into account previous results on this outburst available in the literature. Nevertheless, several alternative two-zone models could be analyzed as well.

For instance, Cao & Wang (2013) in their attempt to explain the weak variability of Mrk 421 in the optical band (compared to X-rays and VHE γ-rays) during its June 2008 outburst, consider an interesting two-zone model featuring a quiescent and a variable emission component. In their scenario, the low-state emission, dominating the total flux in the radio-to-optical range, is produced in the outer jet by particles accelerated via the Fermi-I mechanism. The variable emission component originates from a much smaller zone in the inner jet, with low-energy particles being continuously injected to this zone and accelerated via Fermi-II process. A change in the Fermi-II acceleration time-scale is then producing a flare. The parameters of this two-zone model are adjusted in a way to achieve a reasonable description of instantaneous SED at low and high states. The geometrical configuration of the two zones invoked by the authors seems to be quite attractive, but a full time-dependent modeling is required in order to explore such a model in detail.

Obviously, at this point, any flare scenario is developed with many simplifying assumptions, as present knowledge on the VHE γ-ray emitting regions is limited, as well as the understanding of detailed physical processes in play. For example, in the two-zone scenario constructed here, one would potentially expect some distortion of the "stationary" leading shock as a back reaction to the external perturbation that is responsible to trigger the turbulence around the quiescent blob, as well as a possible shock-turbulence interaction (Andreopoulos et al. 2000). A variety of linear and non-linear phenomena that can deform the stationary shock may arise, affecting the low-state emission, and altering the properties of the turbulence. In our modeling however, we have completely neglected these possible effects. Another option that we did not consider here is that, instead of the turbulence being directly injected at the edges of the blob, the external perturbation may at first disturb the "quiescent" leading shock, which can in turn amplify the turbulence in the downstream around the blob. Another physical process expected to play an important role is such collisionless plasmas with the presence of shocks and turbulence, is the magnetic reconnection (e.g. Karimabadi et al. (2014)), that we did not include in our model, although it could provide a considerable contribution to particle acceleration, described with a term in the kinetic equation formally similar to that of shock acceleration.

The assumption of flare as a weak perturbation

The fundamental hypothesis in our generic model is that some VHE flares can

be considered as a moderate perturbation of the steady state. This, in fact, non-trivial assumption proved to be quite powerful, allowing us to significantly reduce the number of free parameters, and to constrain the global flare scenario quite tightly.

A preliminary examination shows that the 2008 flare of Mrk 421 could be a good example of an event in which the weak perturbation hypothesis can be valid. It will be interesting to explore the domain of applicability of such an assumption, by exploiting it in the modeling of other AGN flares, once MWL data sets with a better quality and time coverage will become available.

Furthermore, it is evident that AGN flares are merely non-destructive transient phenomena, and that their total energy budget, although quite remarkable during the flaring episodes, still remains minor in comparison to the energy emitted over years in the quiescent states, as observed so far in bright blazars. However, a future detection of extended VHE γ-ray emission from blazars might impose a strong limitation on the considered approach, except for the case where the bulk of the quiescent emission is produced in a radiatively dominant compact region. Up to now, in AGN, extended VHE γ-ray emission from a jet was observed only from a non-blazar object – the radio galaxy Cen A (H. E. S. S. Collaboration et al. 2020a). This result casts doubt on the validity of the proposed approach for radio galaxies, which could be explored in more detail if there were VHE flares observed from Cen A, and this is not yet the case so far.

Lognormality and noise

Another method to investigate the link between quiescent and flaring states of blazars, is to characterize the temporal variability of their emission. The power spectral density (PSD) of VHE γ-ray light curves of bright blazars typically represents a power law of a form $P(\nu_t) \propto \nu_t^{-\beta}$, where ν_t is the temporal frequency, and β is a variable index ranging from 1 to 2. The PSD shows how the variability amplitudes are distributed over the different time-scales and hints that underlying stochastic processes are operating in the emitting regions, with correlated colored noises usually of the two types: (1) flicker or pink type ($\beta = 1$), or (2) the random walk or red type ($\beta = 2$).

A thorough analysis conducted for the blazar PKS 2155-304 (Albert et al. (2007b) ; H. E. S. S. Collaboration et al. (2010) ; H. E. S. S. Collaboration et al. (2017)) showed that the flux distribution obeys a lognormal distribution in a steady state with very low activity over several years, as well as in a prominent flaring state in 2006.

Lognormality is a quite common situation in nature and technology. In the context of AGN, it can manifest, for example, in accretion disks, and could possibly also

appear in turbulent jets due to underlying physical processes. Indeed, various laboratory studies of turbulent flows, including boundary layers and jets, have actually revealed that fluctuations of mean energy and mean rate of energy dissipation appear to be lognormal on the large-scale fluctuations. It is expected that such lognormality, manifesting independently of the Reynolds number and configuration in which the turbulence is generated, will be universal and, perhaps due to certain common multiplicative stochastic process linked to interactions between many scales via the transfer of energy, the product of a large number of independent stochastic variables (Mouri et al. 2009). In the two-zone scenario presented in this work, the turbulent processes underlay both the low-state emission and the flaring one, in a somewhat different manner via shock and stochastic acceleration mechanisms respectively. Despite the fact, that AGN jets are rather far from the laboratory ones, it would be an attractive possibility to ascribe the lognormality seen in VHE light curves to a turbulent process that, to some extent, shares similarities with the universal one supposed to manifest in laboratory turbulent flows, and to which one can apply the multiplicative central-limit theorem. Indeed, particle-in-cell simulations have also demonstrated a log-normal distribution of the number density and internal energy density in turbulent collisionless magnetized relativistic electron-positron plasmas, resembling to those believed to be present in AGN jets (Zhdankin et al. 2018). Adhering to this interpretation, comprehensive PSD studies could allow to probe the characteristics of the turbulence in the VHE γ-ray emitting zones.

Actually, for the two blazars, PKS 2155-304 and Mrk 421, different values of the β index were obtained for the low and high states, with the same tendencies in both sources. The results point out flicker/pink noise for the states of steady emission and a random walk / red noise for the states of VHE flaring activity, with $\beta = 1.1^{+0.10}_{-0.13}$ for PKS 2155-304 on time-scales from 1 day to several years (H. E. S. S. Collaboration et al. 2017) and $\beta = 1.1^{+0.5}_{-0.5}$ for Mrk 421 on time-scales ranging from months to years (Goyal 2020), and with $\beta \simeq 2$ for PKS 2155-304 on time-scales in the range from a few minutes to a few hours (Aharonian et al. 2007) and $\beta \simeq 1.75$ for Mrk 421 on time-scales from seconds to hours during its February 2010 flare, measured on February 17 (Abeysekara et al. 2020). As one can see, in both objects, the PSD is flatter for low activity states, and steeper for high activity states. A tentative interpretation of this trend, extrapolating from the two-zone model developed in this thesis, is that the different types of noises seen in blazars originate from the different dominant acceleration processes, associated with perturbative events and with respective turbulences underlying the steady state or flaring activity states. The slowly varying Fermi-I acceleration and particle injection by the shock ahead of the blob, together with prolonged slowly variable turbulence within the blob (with $q \leq 2$), may allow a better distribution of the variability amplitudes across the different related

time-scales above the 1-day time-scale, given also that shock front perturbations can happen over a broad range of time-scales, including days, months and years. Conversely, transient stochastic acceleration induced by the short-term turbulence with a hard-sphere spectrum ($q = 2$) adopted in our modeling, can allow keeping larger variability amplitudes at $\sim$1 hour time-scales matching the largest spatial scales of the turbulent region, in comparison to the shortest time-scales (minutes or seconds), due to cascade processes in the turbulent plasma, which leads to a more red noise than in the case of the quasi-stationary long-term emission.

Analogy with hotspots of extragalactic radio sources

When constructing the two-zone scenario presented in this work, first the simplest scenarios were taken as a starting point, and then meaningful complexity was added step by step as dictated by the constraints provided by the MWL data set. The final two-zone model developed for the February 2010 flare of Mrk 421, includes a stationary shock located upstream of a central blob producing the steady VHE γ-ray emission, and a smaller transient turbulent region which generates the flaring emission, and is situated at the edges around the blob. In this physical picture, the low-state emission arises due to electrons accelerated via the shock acceleration mechanism, while the flaring emission is produced by the electrons accelerated via Fermi-II mechanism. Such a setting, where both Fermi-I and Fermi-II processes are at work, resembles to a situation which is anticipated and observed at much larger spatial scales in some of the hotspots at the ends of extragalactic jets (e.g. Kruells (1992)). Indeed, based on detailed analysis of maps, spectra, or polarization characteristics of hotspots observed in low-redshift sources, various authors argue that both Fermi-I and Fermi-II processes are needed to describe the observational data of the hotspots, and develop multiple-zone particle acceleration models combining the two acceleration mechanisms. The spatially resolved hotspots they present, show some resemblance to the configuration of the two-zone model developed in this thesis, and also feature a compact front shock with a diffuse turbulent zone in its wake (e.g. Isobe et al. (2017)). Such high-resolution hotspot maps were, for example, derived for the radio galaxies 3C 105, 3C 195, 3C 227 and 3C 445 (Orienti et al. (2012) ; Orienti et al. (2017) ; Migliori et al. (2020)). Thus, the model constructed independently in this thesis for the compound steady and flaring VHE γ-ray emitting region of Mrk 421, characterizes it as a kind of "*mini-hotspot*", emerging in the jet much closer to the central engine at distances below 1 parsec. Despite a large difference in the related spatial, energy and temporal scales, such an analogy could be fruitful, for example, to improve the description of the MHD characteristics and macrophysics of the VHE emitting region

based on various data collected for the spatially-resolved hotspots.

Conclusion

The physical scenario we propose in this thesis seems to be a reasonably realistic and totally viable model to initiate blazar flaring activity proceeding over time-scales of $\sim$1 day. Further application to other flare data sets (both archival ones and those resulting from future MWL campaigns) will allow to test this model and verify whether the intermittent turbulent re-acceleration is indeed at the origin of the majority of blazar flares.

Chapter 7

Preparation of Cherenkov Telescope Array

In this chapter we present the preparatory studies conducted for the future Cherenkov Telescope Array (CTA). The future instrument and its key science projects are presented in Section 7.1. In Section 7.2 we present the Gamma-Ray Cherenkov Telescope (GCT), one of the designs for the Small-Sized Telescopes (SST) sub-array of CTA, and in Section 7.3 we perform simulations of the ideal and non-ideal optical performance of the GCT. Finally, in Section 7.4 perspectives are discussed.

7.1 CTA project

7.1.1 Overview

Current major IACT systems (H.E.S.S., MAGIC and VERITAS) have provided many important and sometimes unexpected scientific results, uncovering violent astrophysical processes at work in different γ-ray sources at extreme energies[1]. A step-change in our understanding of high-energy universe is expected with the start of operations of the Cherenkov Telescope Array (CTA, see Fig. 7.1). It is the next-generation ground-based γ-ray instrument, expected to start data taking in 2022, and is presently under

[1] e.g. H.E.S.S.: H. E. S. S. Collaboration et al. (2013) ; H. E. S. S. Collaboration et al. (2015) ; HESS Collaboration et al. (2016) ; H. E. S. S. Collaboration et al. (2020b), VERITAS: VERITAS Collaboration et al. (2011) ; Acciari et al. (2011a) ; Abeysekara et al. (2015), MAGIC: MAGIC Collaboration et al. (2008) ; MAGIC Collaboration et al. (2016) ; MAGIC Collaboration et al. (2019).

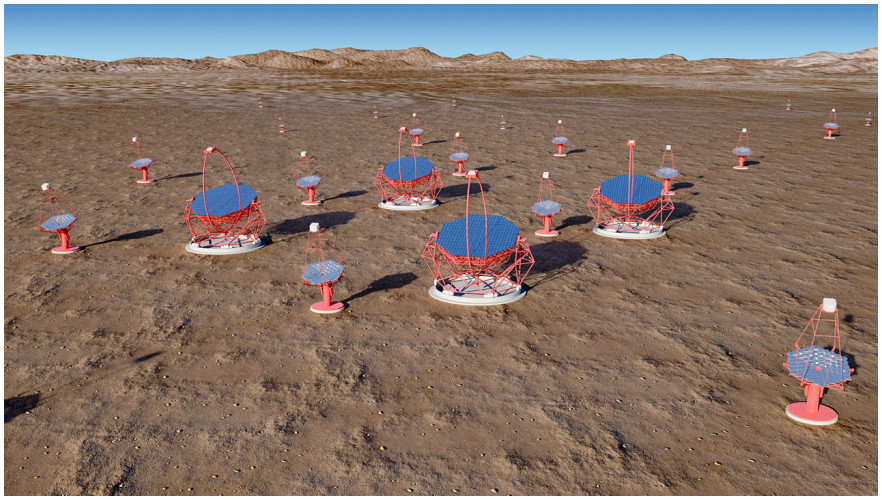


Figure 7.1: Computer-generated image of the future Cherenkov Telescope Array. (image source: eso.org)

development. The CTA project includes about 1500 participants from 31 countries worldwide. The full array will be composed of northern and southern arrays to ensure coverage of the entire γ-ray sky. The Northern Hemisphere array will be hosted by Roque de los Muchachos Observatory in La Palma, Spain, and the Southern Hemisphere array will be located near the European Southern Observatory's (ESO) Paranal Observatory in Chile. CTA will be the largest γ-ray observatory for the coming decades. The instrument will have greatly improved performance compared to current-generation IACTs:

- *order of magnitude higher flux sensitivity* (higher photon rate for faint sources, and access to short time-scale phenomena) (see Fig. 7.2)
- Substantially better
  - *angular resolution* (improvement of extended sources imaging)
  - *spectral resolution* (higher quality spectra)
  - *timing resolution* (resolving time delays and short time-scale flux variations)
- *an extended energy range*, yielding unprecedented spectral coverage from **a few tens of GeV to ∼300 TeV** (low energy end: outperforming *Fermi*-LAT at GeV energies, high-energy end: opening a new window at high-energies and probing extreme particle accelerators)

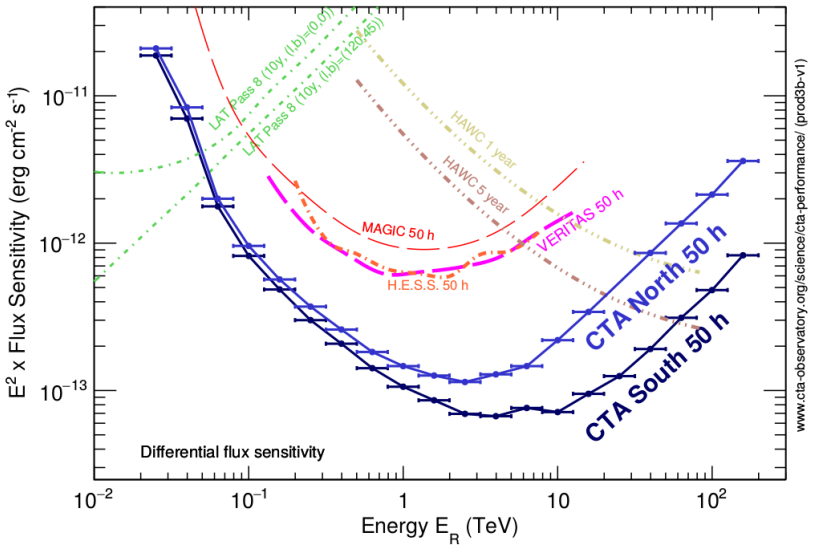


Figure 7.2: Energy flux sensitivity of CTA (North and South sites). The sensitivity threshold is defined as detection of a source at the level of five standard deviations with an energy binning of five independent logarithmic bins per decade of energy. (image source: Cherenkov Telescope Array Consortium et al. (2019))

Owing to these advanced characteristics, CTA will play a decisive role in addressing various astrophysical problems of fundamental importance, which are discussed in the next sub-section (7.1.2).

In order to have drastically enhanced sensitivity, as compared to current-generation instruments, CTA will use more than 100 individual telescopes distributed across wide surface on the ground on both sites. The northern site will host 19 telescopes, while the southern one 99. The full array will be composed of three classes of telescopes based on their sensitivity: Large-Sized Telescopes (LST), Medium-Sized Telescopes (MST) and Small-Sized Telescopes (SST). The northern site will use 4 LSTs and 15 MSTs, and the southern site will include 4 LSTs, 25 MSTs and 70 SSTs. The LSTs provide very large collective mirror areas, allowing to "see" very well faint Cherenkov light flashes from low-energy γ-rays on top of the fluctuating NSB and hence be highly sensitive in the $\sim$10 GeV domain. The LST mirror will be 23 meters in diameter, and its camera will be equipped with photomultiplier tubes (PMTs). The SSTs, being numerous, will be spread across an area of several square kilometers, and therefore give access to the highest energies from 1 TeV up to $\sim$300 TeV, since the chance of at least one SST being within the Cherenkov light pool produced by very rare multi-TeV γ-rays is greatly increased. The SSTs will be placed only on the southern site because one of their primary objectives will be the study of Galactic sources, which are best

seen from the Southern Hemisphere. The SST will have a mirror with a diameter of ∼4 meters and its camera will be equipped with 6 mm silicon photomultipliers (SiPM). The MSTs, featuring a compromise in terms of quantity / mirror area, will cover the energy range in the middle, 100 GeV – 10 TeV. The MST mirror will be around 12 meters in diameter and its camera will employ PMTs.

Unlike its predecessors, CTA will be operated as an open, proposal-driven observatory for the first time in the Cherenkov astronomy. After end of a proprietary period (typically 1 year), CTA data will become publicly available at the CTA data archive. Giving free, unrestricted access to scientists worldwide, beyond the consortium and the traditional Cherenkov astronomy community, is expected to maximize the scientific output of CTA.

7.1.2 Future science with CTA

Science questions that will be studied with CTA can be grouped into 3 broad topics:

- Probing extreme astrophysical environments
- Origin of high-energy cosmic particles
- Exploring new fundamental physics

– Probing extreme astrophysical environments

VHE γ-ray emission from distant sources is a tracer of various high-energy processes, often emerging due to extreme physical conditions, that provide mechanisms of particle acceleration and emission. Study of VHE γ-ray emitters gives us a powerful tool to reveal the characteristics of those violent environments and test fundamental laws of physics in the regimes that are unreachable in human-made laboratories. A few examples of such extreme environments include AGN jets, vicinity of neutron stars and black holes, supernova explosions, cosmic voids, etc.

Relativistic AGN jets: The high sensitivity of CTA should lead to an order of magnitude increase in the number of AGN detected in the VHE regime, providing a very important input for AGN unification schemes and studies on AGN evolution (for more information, see Sol et al. (2013)). Unprecedented characteristics of CTA will enable us to substantially advance in our understanding of the nature of AGN flaring phenomenon: higher-quality spectral measurements in a broader energy range and light curves with much better time resolution should more strongly constrain scenarios for

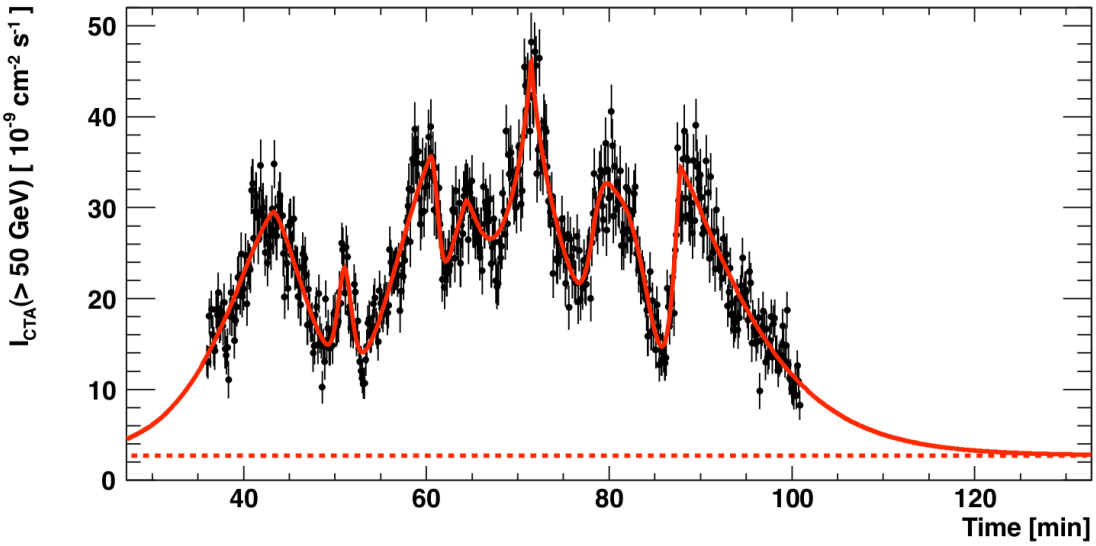


Figure 7.3: Simulated CTA light curve for the rapid flare of PKS 2155-304. (image source: Sol et al. (2013))

the flaring activity. Rapid AGN flares ($t_{\rm var} \lesssim 10$ min) are of special interest: due to enhanced flux sensitivity, CTA will be able to probe very fine details of the flux variation patterns during these exceptional events, providing valuable information on their origin (see Fig. 7.3). Wider spectral range and better spectral and timing resolution will allow CTA to impose more stringent constraints on the AGN emission mechanisms and possibly disentangle leptonic and hadronic emission models. Obviously, coordinated programs of CTA with instruments in other energy bands for a good MWL coverage during observations are crucial for the future physical modeling efforts. Also, CTA will study the physics of relativistic jets, including jet formation, dynamics and strength of the magnetic field in the AGN jet. More information on the expected scientific return of CTA related to AGN studies can be found in Zech et al. (2019).

Neutron stars and black holes: CTA will explore poorly understood physical processes that operate in the surroundings of neutron stars by studying binary systems with a neutron star, and performing mapping and spectral analysis of pulsar wind nebulae. One of the key studies of CTA devoted to black holes, is focused on a binary system Cygnus X-1 with a stellar mass black hole. This X-ray binary is a radio emitter and has jets (a microquasar), and also shows evidence for TeV emission, as reported by MAGIC (Albert et al. 2007c). Better sensitivity of CTA will help to get an insight into non-thermal physical processes happening in the vicinity of stellar mass black holes, in particular, in microquasars. Such studies should also enable to establish the missing

connection between the small-scale jets in microquasars, and large-scale jets in AGN, improving our understanding of jet formation and its general properties at different spatial scales. Also, combining γ-ray and gravitational waves data for neutron star and black hole mergers will provide invaluable information about extreme physical processes in play during these events.

Cosmic voids: These are regions of extremely low density between filaments, containing no or almost no galaxies. The VHE γ-ray beam provided by certain blazars, represents a unique tool to probe magnetic and radiation fields inside the voids which intervene between the Earth and a distant blazar. The **EBL** contains valuable information on the cosmological evolution of the Universe, in particular, evolution of stars and galaxies. Due to strong foreground comprising zodiacal light and light from the Milky Way, direct EBL measurements are very much complicated. One could infer spectral properties of the EBL indirectly, by studying the effect of EBL absorption in the VHE γ-ray blazar spectra. A large sample of blazars detected by CTA at different redshifts, with high-quality broad-band spectra stretching much further, than those measured with current IACTs, will allow CTA to reconstruct accurately the EBL spectrum from optical to far infrared wavelengths at redshift zero, as well as to measure the time evolution of the EBL for the first time, using distant blazars with redshifts up to $z \sim 1$. The effect of γ-γ pair production in the intergalactic medium provides also a possibility to probe **intergalactic magnetic fields (IGMF)**. Electrons and positrons, resulting from interactions of VHE γ-rays emitted by a blazar with low-energy EBL photons, can produce γ-rays of GeV energies via IC upscattering of CMB photons. If the magnetic field in the intergalactic medium is zero, the trajectories of the electrons and positrons are not deflected, and all GeV γ-rays will follow the path of the primary γ-rays. On the other hand, if the IGMF is non-zero, the cascade γ-ray emission will arrive from a direction which is not exactly aligned with that of the source, leading to a decrease of the cascade contribution to point-source flux. Depending on the strength of the IGMF, the deflections of the electrons and positrons could be very small (non-resolvable), so that the cascade emission will arrive within the PSF, but with a time delay due to deflection-induced path difference (*pair echo*), or large enough that the secondary γ-rays form an extended halo-like emission (*pair halo*). Insufficient sensitivity and angular resolution of currently operating γ-ray telescopes precludes the detection of these effects. Non-detection of the cascade emission from the direction of TeV emitting blazars by *Fermi*-LAT imposes a lower limit on the IGMF strength ($B \geq 3 \times 10^{-16}$ G) (Neronov & Vovk 2010). Thanks to its improved angular resolution and sensitivity, CTA is expected to detect either pair halos around point-like TeV emitting sources or detect delayed echo emission, and impose strong constraints on the strength of the IGMF. Some authors however argue that these effects do not exist: electrons and positrons produced in the interactions

of VHE γ-rays with EBL photons, will just heat the very low-density intergalactic plasma without inducing cascades, due to plasma beam instabilities, growing on time-scales much shorter than the one of the inverse Compton cooling (Broderick et al. (2012) ; Schlickeiser et al. (2013)). CTA will be able to verify this hypothesis, and if this is the case, to constrain the plasma heating rate.

– Origin of high-energy cosmic particles

Spectra of various astrophysical objects extend to the highest energies achievable with currently operating IACTs ($\sim$10-30 TeV), and are likely to stretch much further (e.g. blazars, pulsar wind nebulae, supernova remnants, etc). This implies that these sources host remarkably powerful particle accelerators, with acceleration mechanisms provided by extreme physical conditions in these objects. The efficiency of these accelerators can be very high, sometimes close to that achieved at the LHC in CERN, so that the maximum particle energy saturates the theoretical Hillas limit, as for example in the Crab Nebula, spectrum of which stretches as far as up to $\sim$ 100 TeV. Such great efficiency is perplexing and challenges existing theoretical models. Even more puzzling are the maximal energies attained in the astrophysical accelerators: e.g. the spectrum of cosmic rays extends up to $\sim 10^{20}$ eV, which is some seven orders of magnitude higher than the maximal energies of protons achieved at the LHC. It is not clear which acceleration mechanisms and cosmic sources can boost particles to such extreme energies. CTA will address the open questions related to origin of CRs and particle acceleration in the variety of astrophysical environments. In particular, thanks to its much wider energy range, spectral resolution and sensitivity, CTA will (*i*) explore and characterize sites of particle acceleration up to PeV energies (PeVatrons) in our Galaxy, (*ii*) verify whether the supernova remnants are the only major sources of Galactic cosmic rays, (*iii*) provide important clues on the sources of UHECRs. These objectives will be achieved via surveys targeting particle acceleration sites, and via high-quality measurements of spectra up to extreme end of the CTA energy range ($\sim$ 300 TeV) and of light curves with fine time resolution, both needed to get an insight into details of poorly understood physical processes powering the highly efficient particle accelerator machines in the Universe. The key targets of CTA for these goals will be pulsar wind nebulae, supernova remnants and AGN.

– Exploring new fundamental physics

CTA should be also able to resolve some of the mysteries in the fundamental physics, including some of the "exotic" effects. This includes searches for dark matter (DM), Lorentz Invariance Violation (LIV) and axion-like particles (ALPs). The nature of **dark matter** is unknown, one the most favored candidates are weakly interacting massive particles (WIMPs), appearing in a variety of Standard Model (SM) exten-

sions. WIMPs have masses in a range from a few of GeV to hundreds of TeV, and could have been produced thermally in the Early Universe. WIMPs can *annihilate*, and the annihilation cross-section needed to yield the observed DM abundance is remarkably close to that of the weak interaction. Another DM candidates are sterile neutrino, able to *decay* into a Standard Model neutrino and a photon, and axion-like particles, proposed as a solution to strong CP problem of quantum chromodynamics, which can experience an oscillation into a γ-ray in a magnetic field. CTA, with its extended energy range, enhanced sensitivity and better angular and spectral resolution, is expected to derive much stronger constraints on the DM models, than those obtained with present-day γ-ray instruments, via observations of the Galactic center, as well as various other targets, and searches for γ-ray excess features in the spectra. Both detection and non-detection of these features will result in profound implications for the current models of DM. CTA will also search for signatures of **Lorentz invariance violation**, which appears in different theories of quantum gravity. LIV induces time delays in propagation of γ-rays of different energies (speed of photons in vacuum depends on the energy), which CTA will attempt to detect by virtue of its improved time resolution. Both positive and negative result will allow to constrain existing theories predicting LIV. Finally, CTA will hunt for **axion-like particles**. The unusually hard γ-ray spectra of some blazars, as well as detection of VHE γ-ray emission from some FSRQs (e.g. 3C 279 by MAGIC Collaboration et al. (2008)), are quite challenging to interpret. One possible explanation is that the γ-rays experience photon-ALP oscillations in the source (similar to neutrino oscillations) in the presence of the ambient magnetic field. γ-rays that convert into ALPs avoid pair-production on internal photon fields in the jet, as well as on the EBL, and propagate freely towards the Earth. Closer to the observer, ALPs can oscillate back to γ-rays in the magnetic field of the Milky Way. As a result, an excess in a specific energy range in the spectrum will appear, with the increase depending on the ALP mass. The detection or non-detection of features due to reduced opacity in the γ-ray spectra of blazars will enable CTA to test the respective physical theories and constrain physical parameters of ALPs. The photon-axion mixing inside AGN jets is examined in detail by Harris & Chadwick (2014).

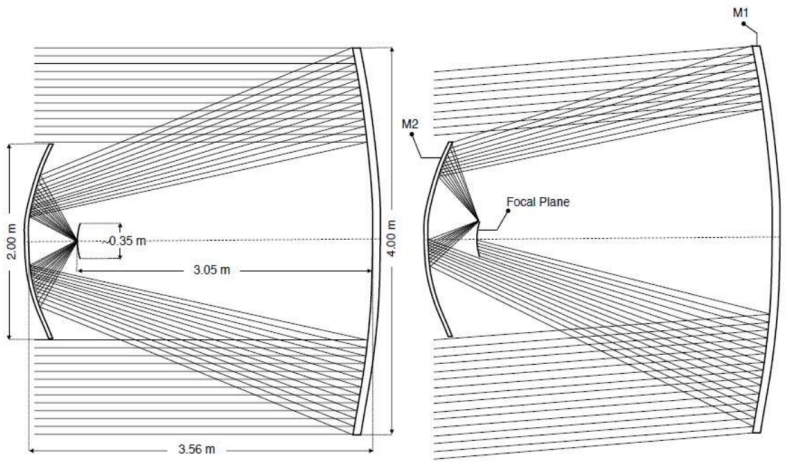


Figure 7.4: Optical design of the GCT for the on-axis observations (left) and at the edge of the FoV at 4.5° (right). Labels "M1" and "M2" denote primary and secondary mirror respectively (image source: Le Blanc et al. (2018))

7.2 Gamma-Ray Cherenkov Telescope: overview and prototyping

7.2.1 Overview

The Gamma-ray Cherenkov Telescope (GCT) is one of three designs proposed for the SST sub-array of the CTA. The optical system of GCT is based on a Schwarzschild-Couder (S-C) design, which features two mirrors that collect Cherenkov light and reflect in on the camera (see Fig. 7.4), contrary to the traditional single-dish Davies-Cotton configuration employed by the currently operating IACTs. The S-C design had never been implemented in Cherenkov astronomy before CTA, due to challenges in manufacturing of aspherical non-conic mirrors up to recent time. Modern technology surmounted those difficulties and allows relatively easy production process. In such optical systems aberrations of the primary mirror can be compensated by those of the secondary, yielding a better quality of air shower images. The design is developed in a way to correct spherical and coma aberrations and minimize astigmatism. It allows relatively wide fields of view combined with smaller plate scales, and therefore is well adapted to small-sized silicon photomultiplier (SiPM) pixel cameras. The high potential of dual-mirror telescopes with aspherical mirrors for wide FoV ground-based γ-ray observations was pointed out rather recently by Vassiliev et al. (2007).

The features of S-C design allow for a compact and light-weight mechanical

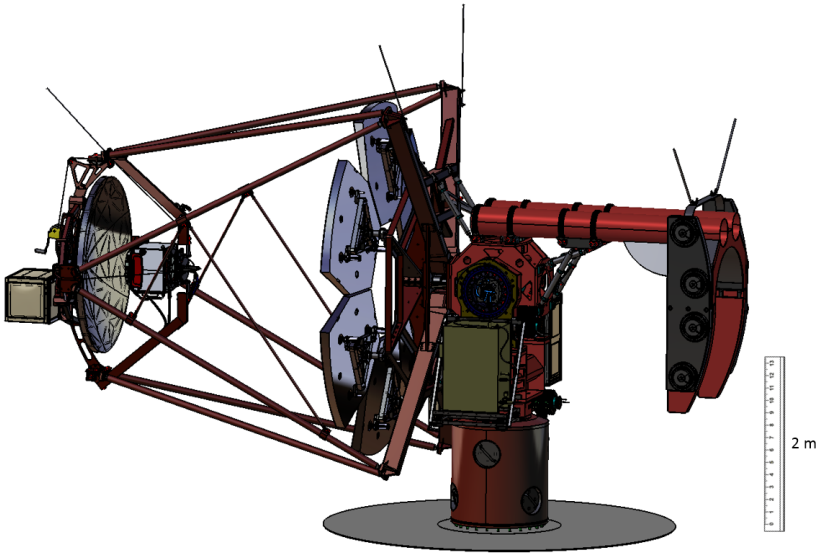


Figure 7.5: Final design of GCT (CAD model), comprising the mechanical structure, primary mirror consisting of six segments, secondary mirror, and the camera between the two mirrors. A 2-meter ruler is added for scale (image source: Dmytriiev et al. (2019b))

structure of the telescope. The Observatory of Paris and the National Institute for Earth Sciences and Astronomy (INSU/CNRS) have jointly developed the mechanical structure of GCT, as well as its control system and aluminum mirror segments. The mechanical structure was optimized in a way to ease its maintenance, shipping, assembly and to minimize the costs. An important advantage of the compact GCT design is that all the telescope sub-systems are directly accessible from the ground. The GCT uses an altitude-azimuth mount with a range in azimuth from $-90°$ to $450°$ and of up to $90°$ in altitude.

The primary mirror (M1) has a diameter of 4 meters, the secondary one (M2) of 2 meters, their separation is 3.56 meters. The primary mirror is composed from six hexagonal aluminum segments (petals). The main parameters of the telescope structure are listed in Tab. 7.1, and the final design is illustrated in Fig. 7.5.

Both CHEC-like (Zorn et al. 2018) and ASTRI-like (Catalano et al. 2018) cameras can be mounted on the GCT.

Optical Parameters		**Mechanical Parameters**	
Field of view	8.2° – 9.2°	M1 diameter	4 m
Focal length	2283 mm	M2 diameter	2 m
Plate scale	39.6 mm/°	Telescope size (*parking position*)	4.1 m × 5.7 m × 8.5 m
Throughput	> 60%	Telescope mass (*with CHEC camera*)	10.8 tons
Effective mirror area (*corrected for shadowing*)	7 m^2 (on-axis) 6.25 m^2 (4° off-axis)	Distance M1 to M2	3.56 m
PSF D80 on-axis	Aim: 3 mm (0.076°)	Distance M2 to camera	0.51 m

Table 7.1:: Main characteristics of the GCT design. PSF D80 is the diameter of a circle containing 80% of light energy.

7.2.2 Prototyping

In order to characterize the performances of the GCT design proposed for CTA, a prototype of GCT (pGCT) has been installed at the Meudon site of the Observatory of Paris. It recorded the first Cherenkov light in November 2015, making it the first CTA prototype to detect VHE Cherenkov events. For this very first observational campaign, the CHEC-M camera was installed on the prototype, and the telescope registered CR-induced air showers. A second observational campaign with the pGCT took place in March and April 2017, during which thousands of cosmic ray showers were recorded (see Fig. 7.6), with the distribution of shower parameters consistent with the one from MC simulations based on the expected instrument performance (Sol et al. 2017). Also, tracking of VHE γ-ray sources was done during this campaign, namely Mrk 421 and Mrk 501, in order to test tracking system and without an aim of detection of the sources. These two campaigns allowed to test the design, in particular, assess suitability of the S-C optical system for detection of air showers, ease of maintenance and operation, reliability of the control system. The results of both campaigns showed that the telescope is very well adapted to perform VHE γ-ray observations. Also, the experience of the prototype operation enabled to better understand the instrument and helped to further improve it.

Since the first campaign, the prototype was equipped with two circular aluminum M1 segments out of six, dummies were installed at the place of the other segments to simulate the correct weight distribution. Three actuators located behind each mirror segment allow to adjust finely its orientation. As a result of the compromise between quality and cost in 2014 when the two mirror segments were manufactured, a better quality of machining and high accuracy of the mirror surface shape was preferred over the quality of aluminum polishing to mirror finish. As a consequence, the mirror

Figure 7.6: Left: The GCT prototype at the Meudon site of the Observatory of Paris. The telescope is equipped with two circular M1 panels out of six. Right: examples of air shower images detected by the CHEC-M camera during the spring 2017 campaign. Credit: Observatoire de Paris. (image source: Dmytriiev et al. (2019b))

segments do not have state-of-the-art micro-roughness characteristic, whereas the mirror global shape is fully in agreement with the expectations. Various tests in 2018 were carried out on aluminum witness-samples to determine an optimal technology and manufacturing process in order to achieve a better mirror quality, especially in terms of the surface polishing. The most successful sequence of production steps was found to be: machining, lapping, nickel plating, polishing and optical coating. This manufacturing process was used to produce another two M1 segments with improved micro-roughness, which were recently (August 2020) installed on the prototype (see Fig. 7.7).

More details on the mechanical structure, different design characteristics, control-command system and mirrors of the prototype can be found in Dournaux et al. (2014) and Le Blanc et al. (2018).

7.3 Gamma-Ray Cherenkov Telescope: performance

In order to prove the suitability of the GCT telescope for the CTA project, specifically its compliance with the CTA requirements for the technical specifications, relative ease in production in 70 copies and durability (at least 30 years of operations), the

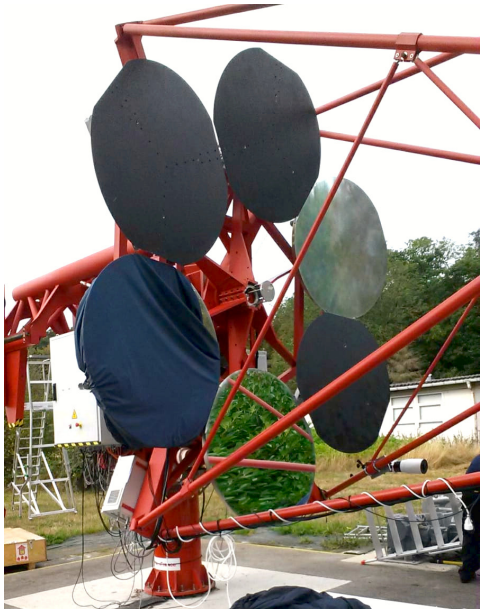


Figure 7.7: A photo showing new M1 segment installed in August 2020 with improved surface polishing (bottom) to be compared with the old M1 element produced in 2014 (top-right). Credit: H. Sol, 2020.

performance of the GCT has to be characterized in detail. This includes in particular performance of the mechanical structure, control system and of the optical system. The first two performances were assessed via a variety of tests conducted by the engineering team (Dournaux et al. (2014) ; Dournaux et al. (2016)), which proved that the mechanical structure is robust enough and is able to withstand emergency stops without suffering any damage, as well as showed that the pointing and tracking performance of GCT obeys very well the CTA requirements. The optical performance of the GCT was characterized with several sets of measurements on artificial light source, and stars, and via simulations and modeling efforts, done by the author of this manuscript in cooperation with engineers from GEPI laboratory of the Observatory of Paris, with the results presented further below[2]. Perspectives of the optical performance with the enhanced mirrors are discussed in sub-section 7.3.3.

[2]Partial results of the presented work have been published in (1) the proceedings of the SPIE in a contribution "Final characterisation and design of the Gamma-ray Cherenkov Telescope (GCT) for the Cherenkov Telescope Array" by O. Le Blanc et al., (including A. Dmytriiev) (July 2018) (Le Blanc et al. 2018), and (2) the proceedings of the ICRC in a contribution "Assessment of the GCT Prototype's Optical System Implementation and Other Key Performances for the Cherenkov Telescope Array" by A. Dmytriiev et al. (July 2019) (Dmytriiev et al. 2019b)

7.3.1 Ideal optical performance

The relevant attributes of the optical system performance include the *effective mirror area*, importance of the *shadowing* effect and the point-spread function (*PSF*). We determine the ideal characteristics by using ray-tracing simulations, performed with the **ROBAST** code (ROOT-based simulator for ray tracing) (Okumura et al. 2016). It is a non-sequential ray-tracing simulation library developed for use in optical simulations of γ-ray and cosmic-ray telescopes. To assess the performance of an optical system with this software one needs first to specify the 3D configuration of different parts of the optical system (their shapes, sizes, relative positions, etc.), and then perform simulation runs in which a beam (a large number of rays) from a source on a finite/infinite distance is cast on the instrument. The ROBAST code allows to follow the path of each ray in the optical system and study the response of the instrument to the light signal. We first characterize the optical performance of the prototype and then the one of the final GCT design. We consider in this sub-section *ideal* optical performance, i.e. without taking into account various imperfections of the optical system (e.g. mirror micro-roughness), with the non-ideal optical performance studied in the next sub-section.

Prototype

As the basis, we used the 3D ROBAST simulation of the mechanical structure of the prototype by Rulten et al. (2016), which included main elements of the mechanical structure, as well as the polynomial coefficients of the M1 mirror curved surface parametrization. We updated the model by taking into account a few design modifications introduced after 2016, and including two circular segments of M1 into our simulation. As a result, we reproduced the actual design of the prototype as of 2017 – 2020, the corresponding computational ROBAST model is illustrated in Fig. 7.8.

For this model, we calculate with ROBAST the associated *ideal* optical performance. The effective mirror collecting area is computed in the simulation via two complementary methods: (1) $A_{\text{eff}} = \frac{N_{\text{focused}}}{N_{\text{simulated}}} \times A$, where N is the number of photons/rays and A is the area of M1, (2) $A_{\text{eff}} = dA_{\text{single-photon}} \times N_{\text{simulated}}$, where $dA_{\text{single-photon}}$ is the effective area per one photon. The top panel of Fig. 7.9 shows the effective area depending on the off-axis angle calculated with the two approaches, which show a very good agreement. One can see that the effective collection mirror area of the prototype equipped with two circular petals of the M1, is of the order of ~ 2 m^2. We also estimate the effect of shadowing due to the elements of the mechanical structure of the prototype (e.g. masts, trusses, camera body, etc.), occurring on

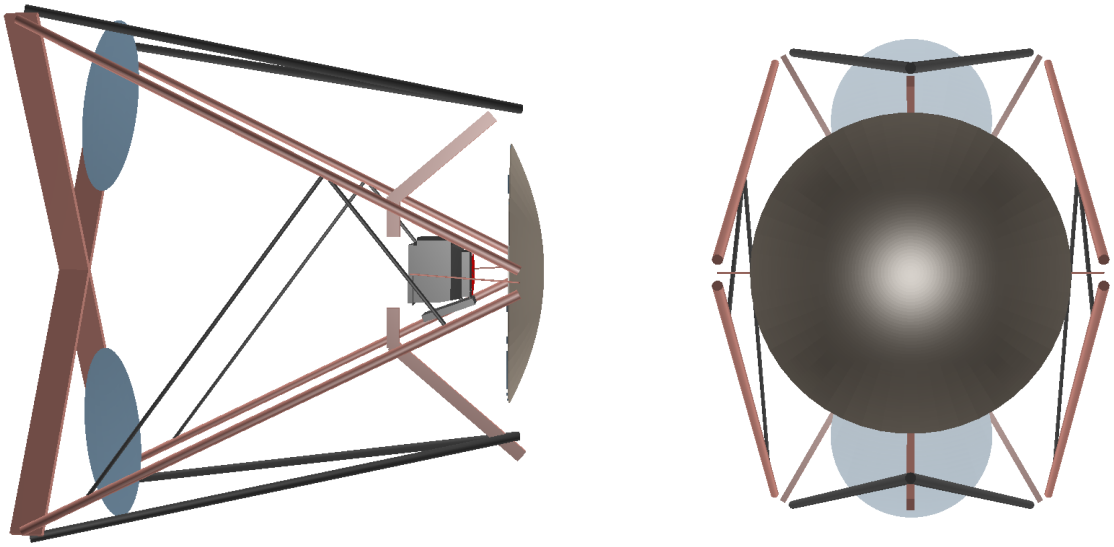


Figure 7.8: Computational ROBAST 3D model of the GCT prototype (*left*: side view, *right*: front view, i.e. from the side of M2), including two circular M1 segments, monolithic M2 mirror, the masts and trusses of the optical support structure, reinforcing bars, and the camera housing. The red surface indicates the ideal focal surface.

the way of light rays. The shadowing percentage is computed as $Y_{\text{sh}} = 1 - \frac{A_{\text{eff,ws}}}{A_{\text{eff,wos}}}$, where $A_{\text{eff,ws}}$ is the effective collection area of the system with the full mechanical structure included in the ray-tracing simulation (top panel of Fig. 7.9), and $A_{\text{eff,wos}}$ is the effective area for the case where all the obscuring elements of the structure are removed from the simulation, i.e. only M1, M2 and the focal plane are retained. The comparison between these two effective areas is depicted in bottom-left panel of Fig. 7.9, and the corresponding shadowing fraction is presented in bottom-right panel of Fig. 7.9. One sees, that the effective area with the structure included drops once we deviate off-axis, whereas an opposite effect is seen when the structure elements are excluded. This increase of the effective area occurs due to reduced obscuration of the two M1 petals by M2. Next, we calculate the PSF characteristics using ROBAST, namely the 80% containment radius (PSF R80) and the progression of the light spot in the ideal focal plane with an increasing off-axis angle. R80 is defined in two ways: (1) as the radius of the circle enclosing 80% of the light energy, and (2) as the half-length of the square enclosing 80% of the light energy. Bottom panel of Fig. 7.10 displays the 80% containment radius of the *ideal* PSF as the function of the off-axis angle, evaluated using these two approaches. One could notice that the *ideal* PSF of the prototype is of the order of 2′. The spot in the ideal focal plane of the prototype depending on the off-axis angle is illustrated in the top panel of Fig. 7.10, with the

zoom into the spot on-axis and at the edge of the FoV shown in the middle panel of Fig. 7.10.

GCT

Having characterized the optical performance of the prototype, we now proceed to calculation of the same characteristics for the GCT. For that, in the ROBAST computational model we replace the two circular segments of M1 with the full GCT primary mirror, comprising six hexagonal petals. Also we replace the CHEC-M camera body with the one of CHEC-S. We reproduce in the simulation the exact contour shape of the mirror segments based on the final design drawings, while the parametrization of the global curved surface shape of M1 remains the same as for prototype. The resulting computational ROBAST model is depicted in Fig. 7.11.

With this simulated design we calculate the effective area, shadowing effect and PSF characteristics. The effective mirror collecting area of the GCT with the full mechanical structure is displayed in the top panel of Fig. 7.12. One could see that the effective area of the GCT equipped with the full M1 mirror is ≈ 7.05 m^2 (on-axis), and ≈ 6.15 m^2 for observations at the edge of the FoV, well above the CTA requirement of ≥ 5 m^2 for this parameter. We also determine the effective area of GCT, when the obscuring mechanical structure elements are removed from the simulation. The comparison between the two collecting areas is presented in the bottom-left panel of Fig. 7.12. One could notice that, contrary to the case of the prototype, the effective area of the GCT without the mechanical structure elements is decreasing with increasing off-axis angle. This happens because the full six-segment M1 mirror is now simulated, so that during the off-axis observations, the M2 will inevitably obscure a larger part of M1 mirror, while unveiling only a small part. The related shadowing effect for GCT is illustrated in the bottom-right panel of Fig. 7.12. The shadowing percentage for GCT observations on-axis is $\approx 12\%$, and increases to $\approx 20\%$ when observing at the edge of the FoV. This result is consistent with estimates based on a previous design (11% for on-axis observation (Rulten et al. 2016)). Also, one sees that the shadowing fraction for the GCT is lower than for its prototype, due to the non-isotropic placing of the mechanical structure elements. As the next step, we deduce the properties of the *ideal* PSF of the GCT. The 80% containment radius depending on the off-axis angle is shown in the bottom panel of Fig. 7.13. One sees that the *ideal* PSF of the GCT has an extension of the order of $1.5'$ for observations close to the axis, and widens to $\sim 4'$ when observing close to the edge of the FoV. The top panel of Fig. 7.13 represents the spot progression in the ideal focal plane of the GCT with an increasing off-axis angle. A higher resolution image of the spot

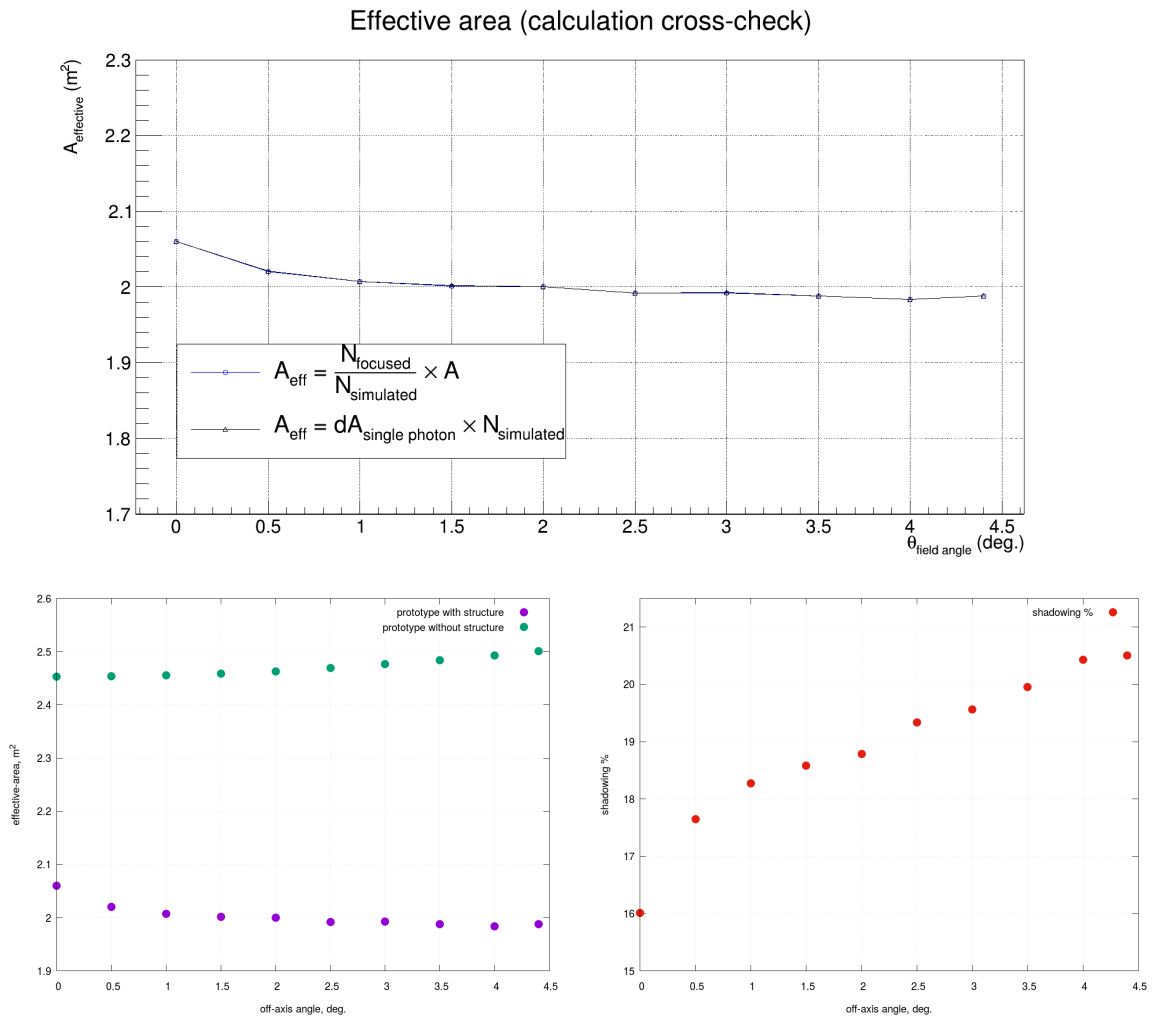


Figure 7.9: *Top*: effective collection mirror area of the GCT prototype with two circular M1 segments as a function of the off-axis angle, computed via two different approaches. *Bottom left*: comparison of the effective area with the full mechanical structure (violet points, same as top panel) and without the obscuring elements of the structure (green points). *Bottom right*: percentage of shadowing induced by the elements of the structure as a function of the off-axis angle. Ideal (100 %) photon detection efficiency and mirror reflectance is assumed in the simulation.

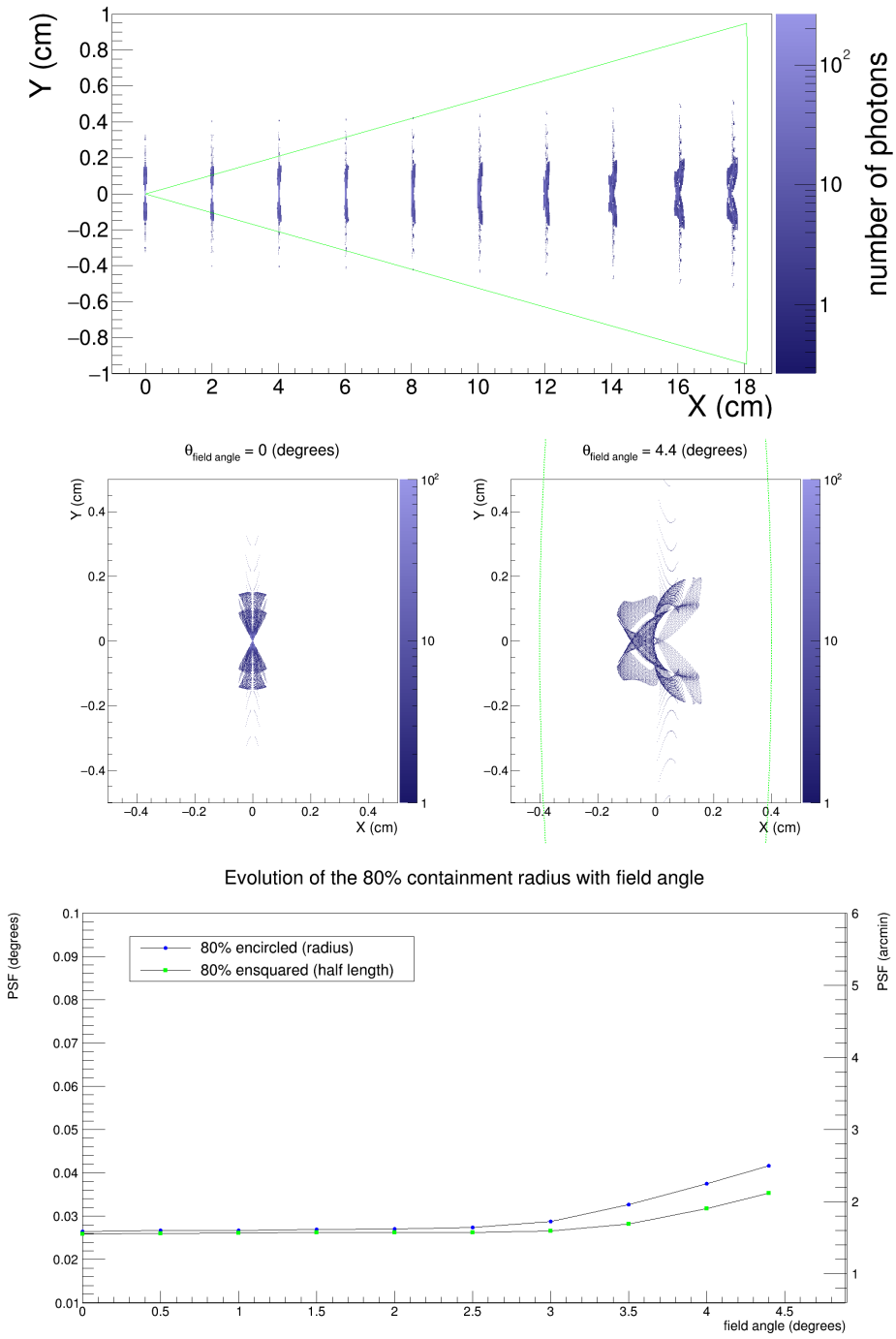


Figure 7.10: *Top*: progression of the light spot in the ideal focal plane of the prototype with increasing off-axis angle (from 0° (leftmost) to 4.5° (rightmost)). *Middle*: A zoom into the leftmost and the rightmost spots from the top panel. *Bottom*: 80% containment radius of the prototype ideal PSF depending on the off-axis angle, for the R80 definitions using encircled (blue points) and ensquared (green points) 80% of energy.

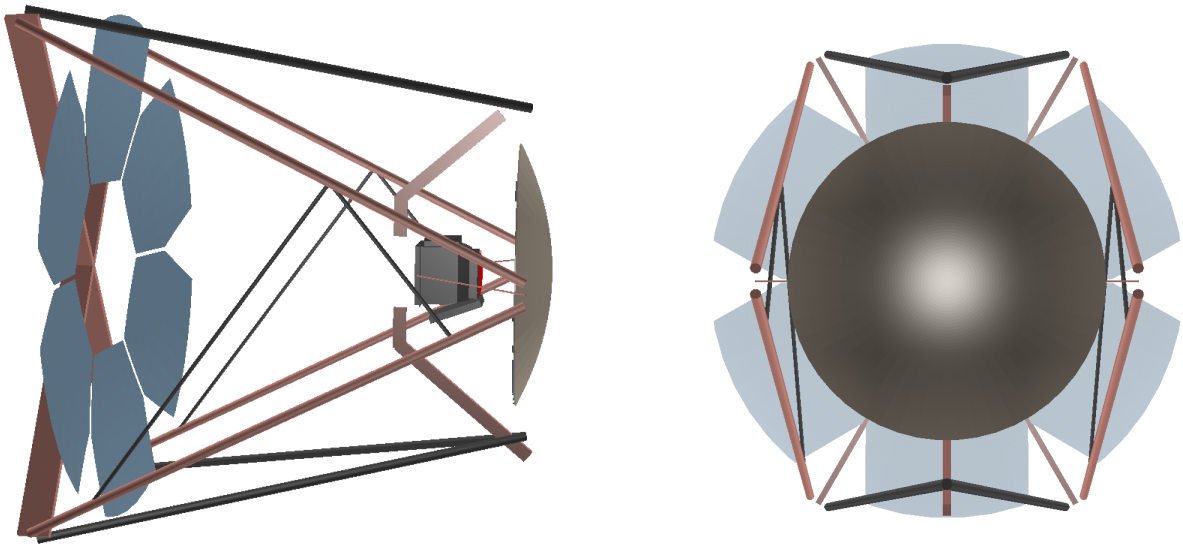


Figure 7.11: Computational ROBAST 3D model of the GCT (*left*: side view, *right*: front view, i.e. from the side of M2), including six hexagonal M1 segments, monolithic M2 mirror, the masts and trusses of the optical support structure, reinforcing bars, and the camera housing. The red surface shows the ideal focal surface.

observed on-axis and at the edge of the field of view (FoV) is provided in the middle panel of Fig. 7.13.

7.3.2 Non-ideal optical performance

Up to now we were considering the performance of the ideal optical design. However, in reality, different imperfections may be present in the optical system, most notably *tip*, *tilt* and *micro-roughness* of the M1 mirror segments. The first two are due to misalignment errors, i.e. slight deviation of the primary mirror segment orientation from the desired one, stemming e.g. from the limited accuracy of the mirror actuators. Tip and tilt of a mirror petal are produced by its rotation (for a small angle) around two axes perpendicular to the segment normal. Specifically, "tip" is a rotation around an axis normal to the sagittal plane, i.e. around an axis dividing the segment into an upper and lower parts, and "tilt" is a rotation around an axis normal to the tangential (transverse) plane, i.e. around an axis of symmetry dividing the segment into two equal parts. These two imperfections are visualized in Fig. 7.14. Mirror surface micro-roughness originates from imperfect polishing of aluminum to a mirror finish, introducing surface irregularities, and concerns not only the M1, but also the M2. The micro-roughness is characterized by a root mean square (RMS) value of the size of surface irregularity R_{q}. All three defects lead to deviation of the light

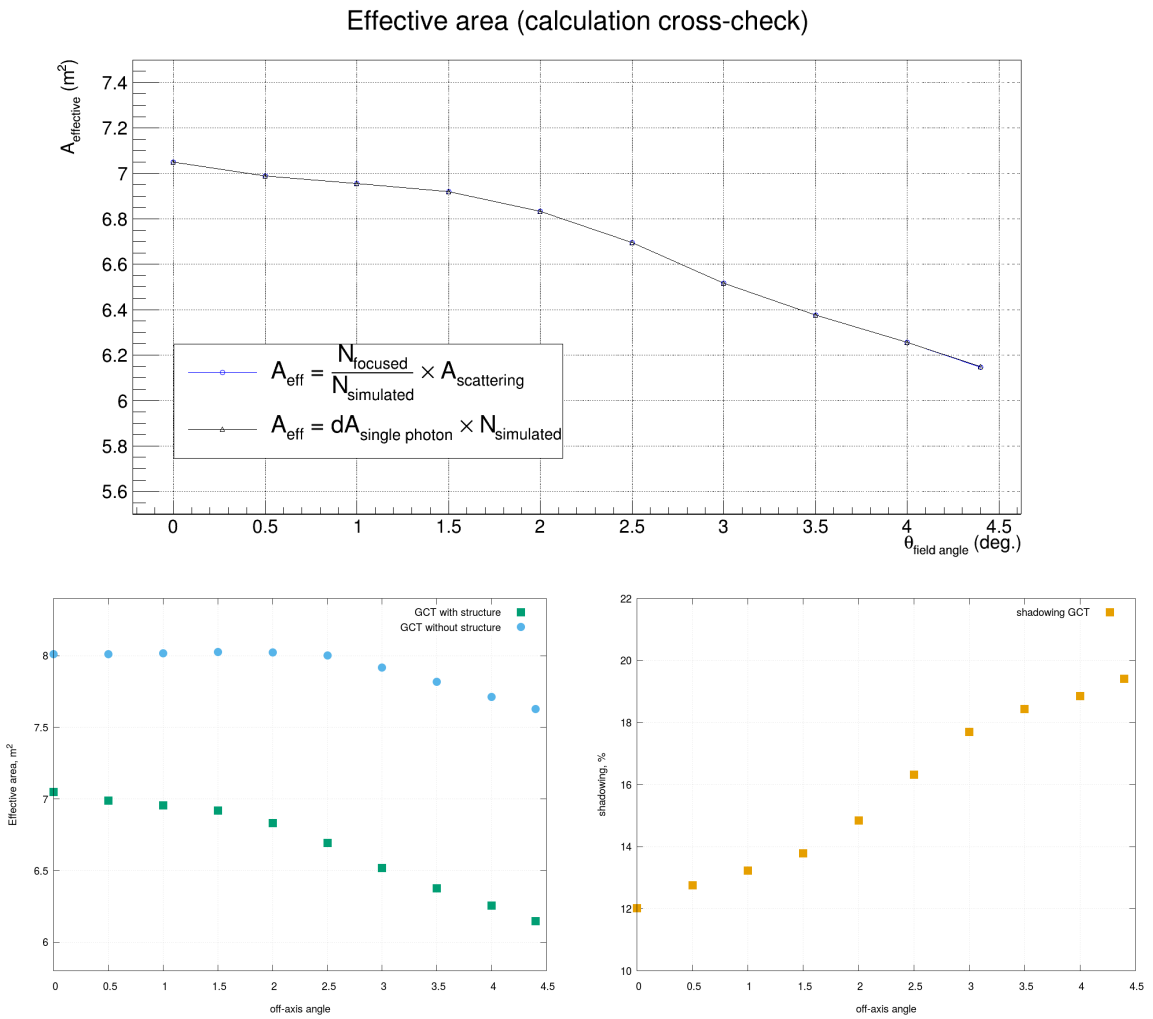


Figure 7.12: *Top*: effective mirror collection area of the GCT with the full M1 mirror comprising six segments depending on the off-axis angle, calculated via two different methods. *Bottom left*: effective area with the full mechanical structure (green points, same as top panel) compared to the one computed without the structure elements (blue points). *Bottom right*: shadowing percentage as a function of the off-axis angle. Ideal (100 %) efficiency of photon detection and mirror reflectivity is assumed in the simulation.

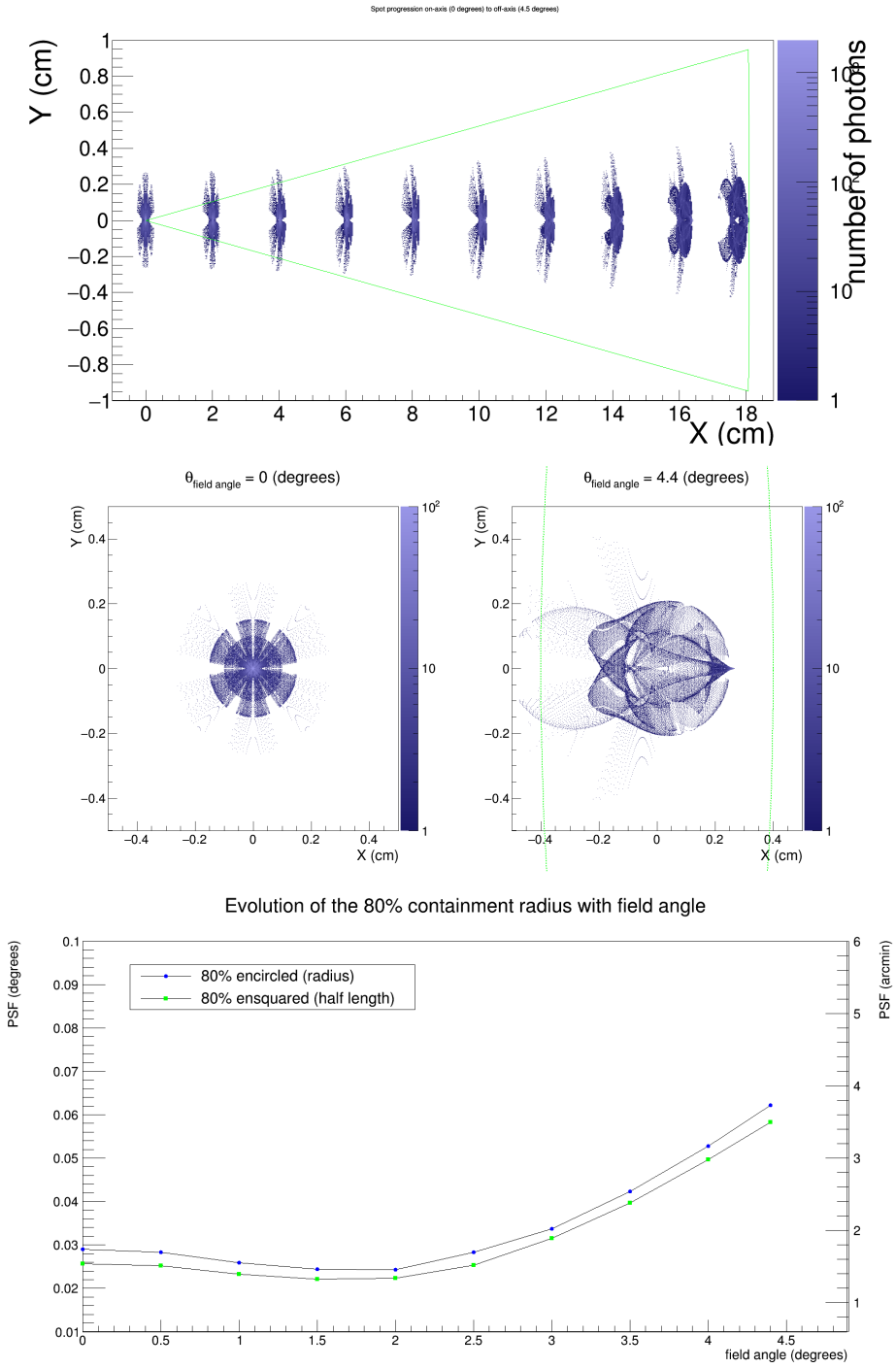


Figure 7.13: *Top*: the light spot appearing in the ideal focal plane of the GCT for different off-axis angles ranging from 0° (leftmost) to 4.5° (rightmost). *Middle*: A higher resolution image of the spot observed on-axis (left) and at the edge of the FoV (right). *Bottom*: 80% containment radius of the ideal PSF of GCT as a function of the off-axis angle, for the R80 defined as encircled (blue points) and ensquared (green points) 80% of energy. The plate scale is 39.6 mm/°.

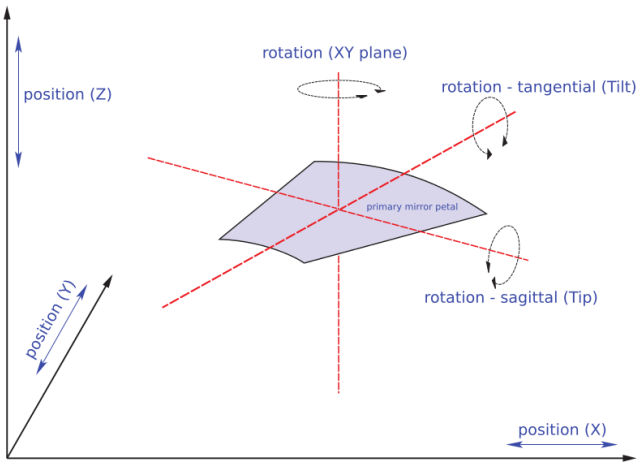


Figure 7.14: Illustration of the optical imperfections associated with mirror segment misalignment. "Tip" is a rotation around an axis normal to the sagittal plane, and "tilt" is a rotation around an axis normal to the tangential (transverse) plane. (image source: Rulten et al. (2016))

rays from their path in the ideal case, and cause the degradation of the PSF. Using ROBAST, we investigate the impact of these imperfections on the PSF of the GCT prototype, with a goal to assess their importance. Such a study will also allow to explore how the instrument PSF can be improved. For the moment, we only consider optical defects related to the M1 petals.

First, we consider the effect of tip and tilt, simulating these defects by intentionally misaligning the mirror segments in the computation model, achieved by specifying non-zero rotation errors for the mirrors petals (using the designated `SetRotationErrors` method of the ROBAST mirror class). We calculated the PSF with different values of tip and tilt, the influence of these imperfections on the PSF R80 is shown in Fig. 7.15. One could see that the PSF size widens as the tip or tilt increases.

Next, we examine the effect of the micro-roughness on the PSF. The rugosity of mirror surface causes the light not only to reflect from it, but also to diffuse. This effect can be simulated with ROBAST using a designated method `SetRoughness` of the ROBAST mirror class. When the method is activated, all the light rays arriving at the mirror surface are reflected with a Gaussian scatter around the direction of the ideal reflection, i.e. a random angle $\delta\theta$ is added to the direction of ideal reflection for each ray, and $\delta\theta$ is distributed normally with a dispersion $\sigma = \theta_{\text{diff}}$ (diffusion angle), specified as an argument of the `SetRoughness` method. For instance, in a case of a parallel beam of light rays falling on a mirror with a micro-roughness, as a

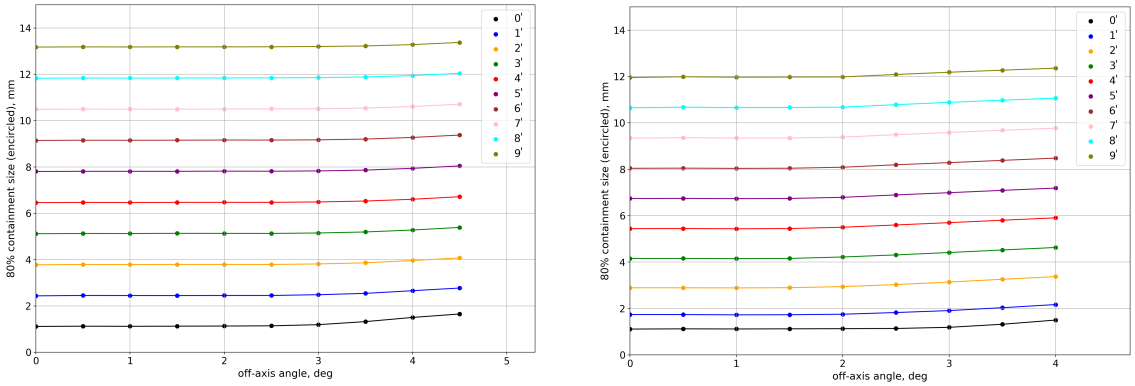


Figure 7.15: Impact of the tip and tilt of the GCT prototype M1 mirror segments on the PSF over the FoV, obtained using ROBAST simulations. *Left*: the effect of the tip. Lines with different colors indicate PSF R80 (encircled, in mm) as a function of the off-axis angle for different values of tip, ranging from $0'$ (no tip, black line) to $9'$ (olive line). *Right*: the same as left, but for the effect of the tilt. The plate scale is 39.6 mm/°.

consequence of the Gaussian distribution, 68% of the light rays will be scattered in a random direction within a cone around the mirror normal with an opening angle of $\theta_{\rm diff}$. The diffusion angle $\theta_{\rm diff}$ is defined by the level of the mirror roughness: the larger is the typical size of a surface irregularity $R_{\rm q}$, the wider diffusion cone one would expect. The Gaussian distribution for the scatter used in ROBAST is clearly an approximation, the real distribution may be more complex.

We simulated the PSF of the prototype, with M1 mirror segments having micro-roughness, for different values of $\theta_{\rm diff}$. The dependence of the PSF size on the diffusion angle and hence the micro-roughness level is illustrated in Fig. 7.16. One concludes that the **micro-roughness has a very important effect on the PSF** and degrades it significantly. This occurs due to the light diffusion, which causes the point source appear blurred, widening the PSF.

7.3.3 Modeling of the PSF of the pGCT

As was already discussed in sub-section 7.2.2, during the manufacturing of the prototype M1 segments, accurate shaping of the global curved mirror surface according to the design was preferred over the high quality of polishing to favor a high quality of the specular PSF, and as a result the prototype M1 petals do not possess the lowest achievable with the current technology micro-roughness characteristic. Based on the study conducted in the previous sub-section, one can expect that the real PSF

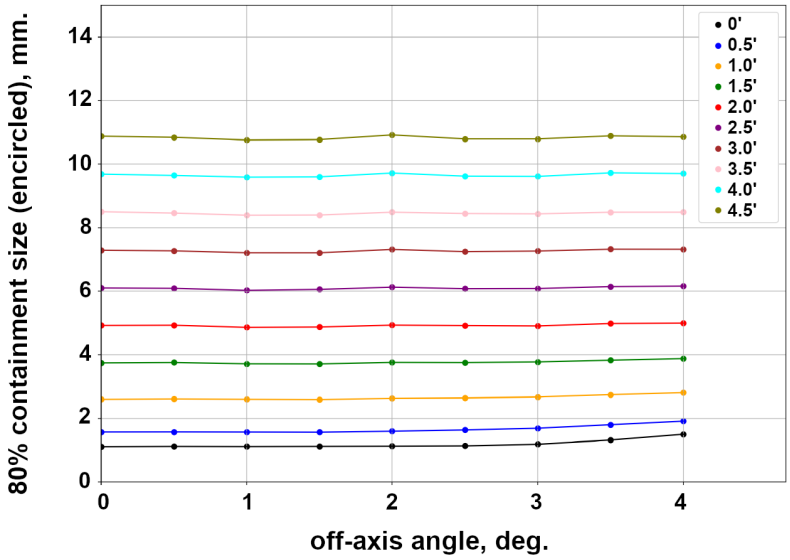


Figure 7.16: Effect of the micro-roughness of M1 mirror segments of the GCT prototype on the PSF over the FoV, derived using ROBAST simulations. Lines with different colors represent PSF R80 (encircled, in mm) depending on the off-axis angle for different values of diffusion angle θ_{diff}, ranging from $0'$ (no roughness, black line) to $4.5'$ (olive line). The plate scale is 39.6 mm/°.

observed with the pGCT will appear degraded due to the mirror rugosity. This can be verified by observations of point-like sources with the prototype, e.g. stars. If the measured PSF indeed could be explained by mirror micro-roughness, it is possible to predict the PSF that will be observed if one equips the instrument with mirror segments having a better polishing quality. Specifically, we will consider perspectives for the PSF with the new optimized mirrors, that were installed recently (August 2020).

PSF observed with pGCT in 2018

The response of the prototype telescope to point sources is most easily measured by observing stars, as they provide a signal with high photon statistic at all wavelengths of the optical spectral range, as well as high signal-to-noise ratios. The PSF measurements with the prototype were performed in summer and early autumn 2018, with an ATIK CCD camera positioned in the center of the focal plane. All measurements were carried out with the single circular best M1 element. A selection of bright stars were observed on-axis on several clear nights at different elevation and azimuth angles to determine the PSF. In addition, observations were carried out with a set

of broad-band filters to test the wavelength behavior of the PSF. Together with the images of the sources, auxiliary measurements, including background measurements, were performed. The reduced images were then analyzed in two different ways to extract the shape of the PSF.

1. Fit with a single 1D Gaussian: The image barycenter was determined and a first attempt was made at fitting the signal in the row of pixels that cross the barycenter with a single Gaussian function. A one-dimensional diameter $D80_{1D}$ containing 80% of the surface under the fitted Gaussian function was determined as $D80_{1D} = 2 \cdot 1.28\,\sigma$, where σ is the standard deviation of a single Gaussian. A typical result is shown in the left panel of Fig. 7.17. A single Gaussian profile was ruled out for all star images. Nevertheless, the resulting standard deviation provides a first rough estimate of the overall spread in the point-source response.

2. Fit with the sum of two 1D Gaussians: Next, the same data were fitted with the sum of two Gaussian functions, where amplitudes and widths were left free to vary, while both functions were required to have a common peak position. For each component, a one-dimensional diameter was determined in the same way as described above, averaged over the x- and y-direction ($D80_{1D,w}$ for the wide and $D80_{1D,n}$ for the narrow component). As can be seen in the right panel of Fig. 7.17, the double Gaussian function provides a good fit to the star images and shows that the PSF can be interpreted as a combination of a narrow and a broad component.

The resulting fit parameters (averaged over the sample of observed stars) are:

1. *Single 1D Gaussian*: $D80_{1D} = 8.1 \pm 0.5$ mm (the value shows only an indicative extent of the PSF)

2. *Sum of two 1D Gaussians*: wide component: $D80_{1D,w} = 14.6 \pm 1.5$ mm, narrow component: $D80_{1D,n} = 5.4 \pm 0.3$ mm.

The ratio of the area below the narrow Gaussian and the wide one, is $A_n/A_w = 0.9$.

It was observed that the width of the large component varies significantly between observations, while the width of the narrow component is very stable. Also, measurements done with broad-band filters in red and blue wavelength bands, show that the observed PSF size is smaller for the red filter, than for the blue one, with a significant difference of about 20%. Such a behavior is expected, when the PSF is significantly polluted by diffuse light due to mirror rugosity. For a given micro-roughness, the effect of diffusion is smaller for longer wavelengths. Finally, a cross-check was done by performing PSF measurements using the second M1 segment mounted on the prototype instead of the first one. The second M1 petal has a higher level of micro-roughness due to a different nickel treatment. It was found that when observing only

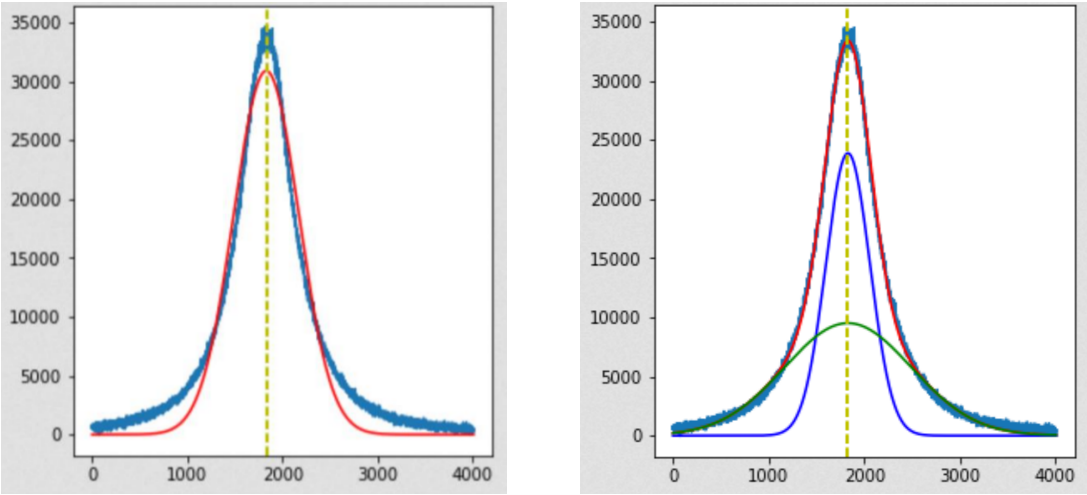


Figure 7.17: Single- and double Gaussian fits to the profile of the observed PSF of the pGCT averaged over several tens of stellar images. *Left*: profile of a single row crossing the star image barycentre (blue), fitted with a single Gaussian function (red). *Right*: the same data fitted with a sum of two Gaussian functions. The individual Gaussians are shown with green (wide component) and blue (narrow component) curves, and their sum is represented by a red curve. Units are ADC counts vs. pixel counts. (Credit: A. Zech)

with the second M1 element installed, the PSF becomes much wider, than with a first segment, having a lower rugosity. It was concluded that in the case of increased micro-roughness of the M1 element, diffusion on the M1 significantly enhances the wide extension of the broad Gaussian component, and also contributes non-negligibly to the narrow component.

Therefore, based on the PSF properties inferred from the observations, we presume that the observed PSF shape is caused by micro-roughness of the telescope mirrors, while the tip and tilt are considered to be negligible, mostly shifting the image but not greatly affecting the PSF. Further on, we will verify whether the PSF behavior can be really described with the light diffusion due to micro-roughness.

Modeling of the observed PSF with ROBAST

Using ROBAST, we try to reproduce the observed PSF shape by simulating the light diffusion on both M1 and M2. We use the computational ROBAST model of the pGCT, described in sub-section 7.3.1, in which we exclude the second M1 element, retaining only one single M1 segment (as the PSF measurements were done in this configuration). Before, in the sub-section 7.3.2, studying the impact of micro-

roughness on the PSF, we simulated the case where 100% of the light rays were diffused when arriving on a mirror with a micro-roughness. In reality, only a fraction of the light rays will undergo the diffusion, and therefore the interaction of the light rays with a mirror surface can be considered as the following: a fraction $Y_{\rm diff}$ of the light rays are diffused (in a random direction), and the remaining fraction $1 - Y_{\rm diff}$ is reflected. The ratio of scattered rays $N_{\rm diff}$ on a single surface over incoming rays $N_{\rm tot}$ is given by (e.g. Singh et al. (1996))

$$Y_{\rm diff} = \frac{N_{\rm diff}}{N_{\rm tot}} = 1 - \exp\left[-\left(\frac{4\pi\,\Delta\cos\theta_{\rm i}}{\lambda}\right)^2\right] \tag{7.1}$$

where Δ is the average micro-roughness parameter, approximately equal or at least related to the micro-roughness RMS $R_{\rm q}$, $\theta_{\rm i}$ is the incidence angle with respect to the surface normal, and λ is the wavelength, which we set to $\lambda = 500$ nm. We denote the fraction of light rays diffused by M1 as $Y_{\rm diff,M1}$, and by M2 as $Y_{\rm diff,M2}$.

We adopt a simplified approach to treat the diffusion of light on pGCT mirrors. After arriving at the M1/M2 mirror surface, the fraction $Y_{\rm diff,M1/M2}$ of light rays undergoes scattering in a random direction around the mirror normal distributed according to a Gaussian distribution with a standard deviation $\theta_{\rm diff}$. We assume that both M1 and M2 have the same level of the micro-roughness Δ, and scatter rays with the same standard deviation $\theta_{\rm diff}$. Based on the design geometry, we consider that the light rays are arriving at the M1 surface with an average incidence angle of $\theta_{\rm i,M1} = 10°$, and at the M2 with an average incidence angle of $\theta_{\rm i,M2} = 32°$. As a consequence, M1 diffuses a somewhat higher fraction of light rays than M2, $Y_{\rm diff,M1} > Y_{\rm diff,M2}$. As a result of the propagation of light rays in the optical system of the telescope, and their interaction with the M1 segment and the M2 in a manner described above (reflection plus diffusion), **four distinct components arise**:

1. Specular-specular: reflection of light rays from M1 and then M2 (pure reflection and no diffusion). The fraction of rays undergoing this process is $Y_{\rm ss} = (1 - Y_{\rm diff,M1})\,(1 - Y_{\rm diff,M2})$.

2. Diffuse-specular: diffusion of light rays on M1 and then their reflection from M2 (cross-term). The fraction is $Y_{\rm ds} = Y_{\rm diff,M1}\,(1 - Y_{\rm diff,M2})$.

3. Specular-diffuse: reflection of light rays from M1 and then their diffusion on M2 (cross-term). The fraction is $Y_{\rm sd} = (1 - Y_{\rm diff,M1})\,Y_{\rm diff,M2}$.

4. Diffuse-diffuse: diffusion of light rays on M1 and then M2 (pure diffusion, no

reflection). The fraction is $Y_{\rm dd} = Y_{\rm diff,M1}\, Y_{\rm diff,M2}$.

The sum of the weights of each component, $Y_{\rm ss} + Y_{\rm ds} + Y_{\rm sd} + Y_{\rm dd} = 1$, as it should be.

We simulate the PSF of the pGCT degraded by micro-roughness, by simulating each of four processes described above separately, and stacking the resulting four light spots appearing in the focal plane, with the corresponding weights $Y_{\rm ss/ds/sd/dd}$. We vary two related parameters, $\theta_{\rm diff}$ and Δ to reproduce the measured PSF shape, by requiring that the cut through the barycenter of the simulated stacked spot image (simulated PSF profile) is consistent with the sum of two Gaussian functions with the parameters deduced from fitting the observational PSF data (widths $\mathrm{D80}_{\rm 1D,w}$ and $\mathrm{D80}_{\rm 1D,n}$, and the areas ratio $A_{\rm n}/A_{\rm w}$). In this way, we ensure that we not only match the global observed double Gaussian profile of the PSF, but also accurately describe the contributing narrow and wide components. The value of $\theta_{\rm diff}$ is constrained by the width of the observed PSF, while the ratio between the narrow and wide PSF components constrains the value of Δ. As there are two direct and independent relations between the two sought parameters and the two key quantities measured from the double Gaussian fit of the PSF (width and areas ratio), a combination of $\theta_{\rm diff}$ and Δ can be deduced in a unique way. As a result, **we reproduce the observed PSF** with $\theta_{\rm diff} = 4.3'$, and $\Delta = 40$ nm. The standard deviation of the wide component of the double Gaussian fit appears to be $\sigma_{\rm fit,w} = 5.7$ mm, which translates into $\mathrm{D80}_{\rm fit,w} = 14.6$ mm, consistent with the observed value $\mathrm{D80}_{\rm 1D,w} = 14.6 \pm 1.5$ mm. For the narrow component, $\sigma_{\rm fit,n} = 1.0$ mm, which translates into $\mathrm{D80}_{\rm fit,n} = 2.6$ mm, somewhat lower than the measured value of $\mathrm{D80}_{\rm 1D,n} = 5.4 \pm 0.3$ mm. The ratio of narrow component to the wide one is $A_{\rm fit,n}/A_{\rm fit,w} = 0.9$, exactly as observed. Therefore, the wide PSF component and the ratio between the components is very well reproduced, while the width of the narrow PSF component appears to be slightly underestimated. This may be due to possible effects from misalignment or from intermediate spatial frequencies on the specular component, which were not taken into account in this simulation. The resulting spot 1D profile describing the observed PSF is illustrated in the left panel of Fig. 7.19, and the simulated stacked spot itself is displayed in Fig. 7.18. The inferred value of micro-roughness parameter Δ appears to be quite close to the experimentally measured value of the M1 RMS $R_{\rm q} = 55$ nm. The resulting fractions of diffused light (Eq. 7.1) are $Y_{\rm diff,M1} \approx 63\%$ and $Y_{\rm diff,M2} \approx 52\%$, so the PSF is dominated by scattered light for the segments of mirrors produced in 2014.

The simulations confirm that, for parameters that are consistent with the characteristics of the prototype mirrors, the narrow PSF component comprises specular reflection (specular-specular component) and single diffusion on the M2 mirror

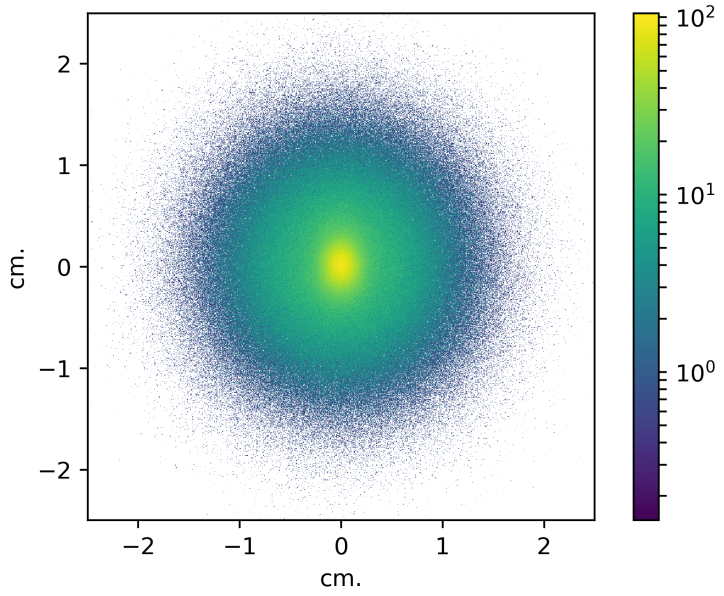


Figure 7.18: Stacked four-component light spot in the focal plane of the pGCT simulated with ROBAST for the standard deviation $\theta_{\text{diff}} = 4.3'$, and micro-roughness parameter $\Delta = 40$ nm, with which the measured PSF is reproduced. The spot represents the simulated response of the pGCT to point-like source, i.e. the PSF. The color bar on the right shows surface density of photons.

(specular-diffused component), while diffusion on the M1 (diffuse-specular component) and double diffusion (first M1 then M2, diffuse-diffuse component) are responsible for the wide component of the observed PSF. Contribution of each of the four components is shown in the right panel of Fig. 7.19).

Perspectives with the mirrors with higher-quality polishing

Having reproduced with the simulations the PSF observed with the pGCT equipped with the M1 segments from 2014, we are now going to predict the PSF that will be observed with the new M1 petals having better quality of polishing, by extrapolating the simulation results obtained previously.

Based on experimental measurements carried out for small mirror samples, which yielded an RMS of micro-roughness of $R_{\text{q}} = 21$ nm and an approximate value of $\theta_{\text{diff,new}} \sim 1.2'$ derived from their bidirectional reflectance distribution function (BRDF), measured at a wavelength of 325 nm, a second simulation was carried out with these values, corresponding to an improved micro-roughness compared to the mirrors from 2014. The results of this simulation are shown in the left panel of

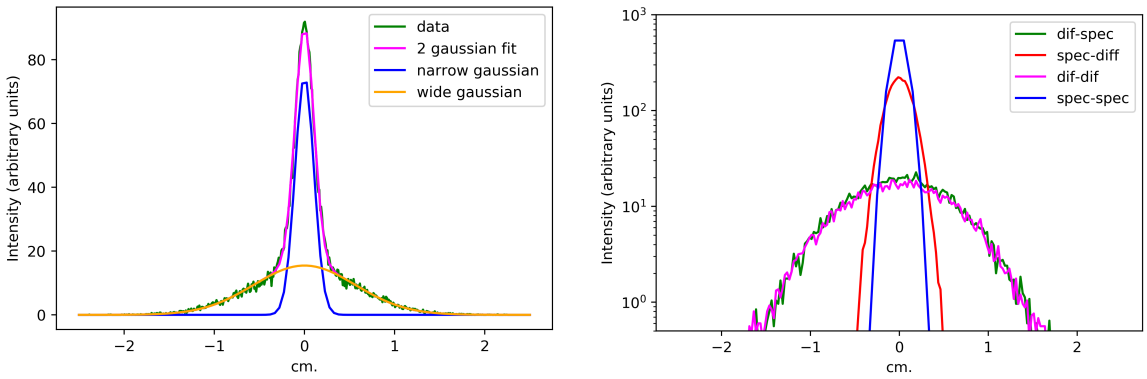


Figure 7.19: Illustration of the components contributing to the observed PSF. *Left*: 1D profile of the simulated PSF with a double Gaussian fit which reproduces the present observations of the prototype PSF. Green curve (labeled "data") represents the cut through barycenter of the simulated composite PSF (Fig. 7.18). The magenta line displays a double Gaussian fit to the simulated PSF profile, with yellow and blue curves indicating the wide and narrow components of the double Gaussian fit respectively. *Right*: Different components that are responsible for the overall shape of the simulated PSF (green curve in the left panel) shown in log scale. Green line represents a component due to diffusion only on M1 ("dif-spec"), red line – diffusion only on M2 ("spec-dif"), magenta line – diffusion on M1 and M2 ("dif-dif"), and blue line – absence of diffusion ("spec-spec").

Fig. 7.20. It can be seen that the impact of diffusion is very significantly reduced with these parameters. The specular component dominates the overall PSF and the diffuse components are much narrower than with the current mirrors. The width of the narrow Gaussian fit component is now $\mathrm{D80_{n,new}} = 2.25$ mm, and the width of the wide component $\mathrm{D80_{w,new}} = 4.15$ mm. The corresponding 80% containment diameter (for encircled energy) is $\mathrm{D80_{new}} = 4.7$ mm for the overall PSF, which is well within the CTA requirements. The four components of the predicted PSF are depicted in the right panel of Fig. 7.20.

It should be stressed that this simulation assumes an improvement of the mirror roughness quality that has been proven to be achievable. An even better performance with respect to micro-roughness is expected with the mirrors which are already installed on the prototype in August 2020 (see Fig. 7.7). Laboratory tests show that an R_{q} of $\sim$7 nm is achieved for these new M1 petals. Using ROBAST, we find that with such level of micro-roughness, D80 decreases to $\sim$2.5 mm on-axis ($\sim 0.063^{\circ}$), neglecting alignment errors. This result has to be validated by PSF measurements on sky (planned to be performed in autumn 2020) with the newly installed M1 segments. Such small PSF would be very well adapted to a new generation of Cherenkov

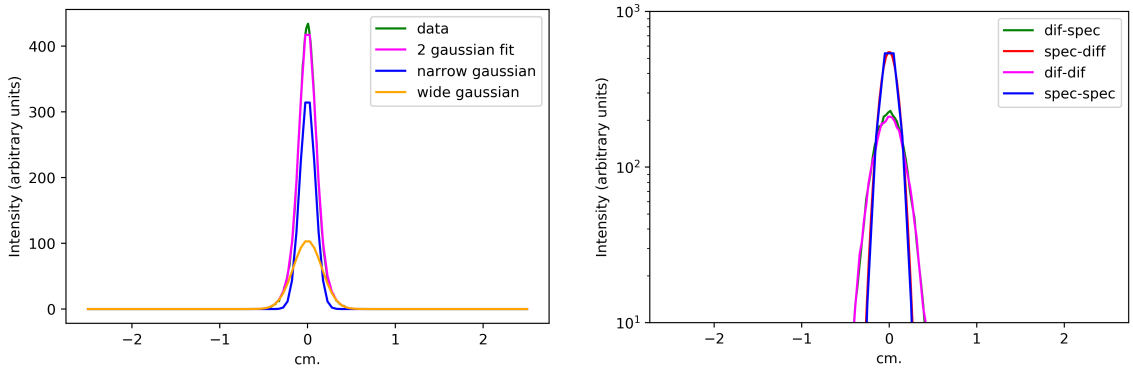


Figure 7.20: Illustration of the components contributing to the PSF predicted for the mirrors having micro-roughness of $R_{\rm q} = 21$ nm and a standard deviation of $\theta_{\rm diff,new} \sim 1.2'$. *Left*: double Gaussian fit of simulated 1D stellar profiles shown in log scale in the case of mirrors with improved roughness of $R_{\rm q} = 21$ nm, at the wavelength of 325 nm. *Right*: Different components that are responsible for the overall shape of the simulated PSF in the left panel (analogous to the right panel of Fig. 7.19, but for the mirrors with improved micro-roughness).

cameras equipped with smaller SiPM ensuring a 3 mm pixel size.

7.4 Discussion and perspective

Unfortunately, relatively recently, another design (ASTRI) was selected by the CTA management for the SST section of CTA. Nevertheless, if the excellent PSF of GCT of ≈ 3 mm will be confirmed by on-sky measurements, this would prove the interest and potential of the GCT optical system for a new generation of finer pixel cameras.

Chapter 8

Conclusion and perspective

In this thesis, three closely intertwined studies were conducted. The core study is devoted to the physical modeling of the brightest detected up to now VHE γ-ray flare of the archetypal blazar Mrk 421, a representative of BL Lac class of type HBL. To reveal the physical processes in the source causing such extreme activity, I developed a time-dependent numerical code "EMBLEM" designed to simulate broad-band varying emission of blazars. The code computes the evolution of the electron distribution in the VHE γ-ray emission zone and the associated SSC emission. We apply the EMBLEM code to the MWL data set of the outburst, and propose a novel global physical scenario to explain the observed flaring behavior, in which the flare is initiated by the second-order Fermi re-acceleration of high-energy electrons in the turbulent region spontaneously arising around the VHE γ-ray stationary emitting zone. Turbulent re-acceleration of the escaping particle flux from the VHE γ-ray emission region boosts electrons much further in energy and significantly hardens the spectrum of the particle population in the transient turbulent zone, causing a dramatic flare in the VHE regime. Our scenario reasonably describes the MWL data set of the flare. We are able not only to fit satisfactorily the available spectral measurements in different energy bands, but also rather well reproduce the temporal behavior of the flux during the flare in different energy ranges, i.e. the detailed shape of multi-band light curves. This is, in fact, one of the first efforts in the literature to fit the profiles of MWL light curves of blazar flares using physical modeling.

In our scenario we make a strong general assumption: the flaring state is achieved by a moderate non-destructive perturbation of the steady state, so that the flux continuously evolves from a quiescent to an elevated level, and then smoothly goes down to the initial low-state level. This assumption allows to drastically reduce the number of free parameters used to describe the flaring behavior. Although such

assumption seems quite intuitive, it is not immediately obvious that a flare should represent a sudden transition from a long-term steady state to a high state triggered by a local perturbation. Various authors for example consider that the blazar emission follows a completely stochastic pattern (e.g. Marscher (2014)) rather than a long-term baseline quiescent state with superimposed isolated flares on top of it. In this case, the nature of a flare is not ascribed to an isolated perturbation of a steady state, but to a stochastic effect. More observational data with a much better time coverage and resolution is needed to clarify the origin of the blazar variability. A substantial progress in this direction is expected with the advent of CTA.

An interesting extension of our flare scenario would be to explore in more detail various scenarios of injection of the turbulent energy in the jet environment around the VHE γ-ray emission region, including the turbulent energy injection time-scales and energetics. Replacing the rather *ad-hoc* turbulent injection term with its more realistic physical counterpart would allow to better delineate the domain of the scenario applicability, as well as to have a more realistic description of the turbulent acceleration process. Other natural further development of the scenario is modeling of the secondary, more modest flare during the February 2010 outburst of Mrk 421.

The code I developed is a versatile and powerful tool allowing to explore many more flares of Mrk 421, as well as of other BL Lac objects. An obvious next application of the EMBLEM code would be modeling of a number of other intermediate time-scale ($t_{\rm var} \sim 1$ d) flares of Mrk 421 (e.g. June 2008 flare), to verify whether the underlying nature of violent γ-ray flux variations on these time-scales in this object can be attributed to the intermittent turbulence in the vicinity of the VHE γ-ray production site, providing a test of the universality of our physical scenario. Another property of our scenario which can be explored for different BL Lac objects and not only Mrk 421, is that a strong VHE γ-ray flare is accompanied by rather weak variability in the optical band. Application of the code to a selection of flares of a variety of BL Lac objects (e.g. Mrk 501, BL Lacertae, PKS 2155-304, etc.) showing this peculiar MWL behavior would allow to test whether flares of this specific kind can be explained with our turbulent re-acceleration scenario, independently of the object.

Furthermore, the effort of modeling blazar flares with my code is not limited to consideration of only one physical scenario. It is quite tempting to apply the code to some unusual behavior observed in blazars that is impossible to describe with simple models (e.g. peculiar features in the light curve profile, extremely hard spectra, orphan flares, etc.) and construct new physical scenarios to explain the observed peculiarities. This includes both the available archival data sets and those from future MWL campaigns. Extending the EMBLEM code with treatment of

magnetic reconnection process would broaden the range of possible scenarios. Apart from flares happening on time-scales of ∼1 day, it is also particularly interesting to study blazar high-states occurring on time-scales of months or even years. For this goal, the currently neglected *adiabatic expansion* of the VHE blob has to be implemented.

Another highly attractive opportunity is to extend the EMBLEM code for the purpose of *FSRQ modeling*. The external Compton process is already treated in the code: in our modeling of the extreme Mrk 421 flare we take into account the cross-scattering effects between the steady-state VHE emitting zone and the radiative turbulent region, which induce an additional inverse Compton component. To be able to model FSRQs, we have to implement in the numerical code the external radiation fields of the BLR, dusty torus and the accretion disk. Application of the code to FSRQ flares can be then done fairly quickly: the flexibility of my code enables a relatively easy and fast adaptation of the external Compton routine to the case of the interaction of soft photons of BLR, dusty torus and the accretion disk with high-energy particles in VHE γ-ray emission region. The first immediate application of the extended EMBLEM code would be physical modeling of 3C 279 flares studied in this thesis. Besides 3C 279, there are many more FSRQs challenging existing scenarios of flaring activity, and so the extension of the code opens a vast diversity of possible applications to a rich palette of unusual behaviors observed in different FSRQs.

Apart from the EMBLEM code, another tool elaborated in this thesis can be used to study BL Lac flares. The *analytical method* I developed to test the viability of the one-zone shock acceleration scenario can be applied to a large collection of flares of different BL Lac objects with a goal to establish the defining common characteristics of flares that can or cannot be explained with the simple one-zone scenario.

A slightly more distant but a very enticing prospect is the application of the extended EMBLEM code and the analytical method to the *future data of CTA* on blazar flares. A significantly improved performance of CTA compared to currently operating IACT systems will allow to uncover many more interesting behaviors in blazars and investigate them in detail. Of particular interest is fast variability, studies of different types of observed noises, exploring the general properties of the VHE blob and whether it can be described as a kind of a "mini hot-spot".

At the moment of writing of this chapter, measurements of the PSF of GCT are performed with the recently installed new M1 segment having a drastically enhanced mirror polishing quality. In case the measurements agree well with the predictions made in Chapter 7, this would fully validate on sky the high level properties expected from the initial GCT optical design. A proof of better PSF quality of GCT compared

to the other designs opens a door to possible future developments in Cherenkov astronomy.

Finally, I would like to conclude this thesis with an insightful quote by Peter De Vries: "*The universe is like a safe to which there is a combination. But the combination is locked up in the safe*".

Appendices

Appendix A

Analytical solution of the kinetic equation for the case of shock perturbing a steady-state electron spectrum

Here we solve the Eq. 6.13 describing perturbation of particle population in the emitting zone by a passing shock. This process is described by two key parameters of the shock: $t_{\rm FI}$ which is the Fermi-I acceleration time-scale characterizing the efficiency of the shock in acceleration of particles, and $t_{\rm cs}$ which a duration of the acceleration phase by the shock, equal to the time of transit of the shock through the emitting zone. The latter parameter is related to the rising time of the flux n the flare light curve.

A.1 Assumptions and boundary conditions

We assume that the physical parameters of the emitting zone do not change during the passage of the shock. The shock enters the blob at $t = 0$.

The initial condition is that the electron spectrum at $t = 0$ is the steady state solution $N_{\rm e,0}(\gamma)$:

$$N_{\rm e,FI}(\gamma, t = 0) = N_{\rm e,0}(\gamma) \tag{A.1}$$

APPENDIX A. ANALYTICAL SOLUTION OF THE KINETIC EQUATION FOR THE CASE OF SHOCK PERTURBING A STEADY-STATE ELECTRON SPECTRUM

The $N_{\rm e,0}(\gamma)$ is the asymptotic stationary solution of the kinetic equation with only injection, escape and cooling terms, and no acceleration process, deduced from the Eq. 6.13 (setting shock acceleration term to zero):

$$\frac{\partial}{\partial\gamma}\left(b_{\rm c}\gamma^2\cdot N_{\rm e,0}(\gamma)\right)-\frac{N_{\rm e,0}(\gamma)}{t_{\rm esc}}+Q_{\rm inj}(\gamma)=0 \tag{A.2}$$

We neglect the inverse Compton cooling, so the $b_{\rm c}$ in constant in time. We also assume $t_{\rm esc}$ is constant in time and not energy-dependent: $t_{\rm esc}\sim R_{\rm b}/c$, where $R_{\rm b}$ is size of the emitting region. We require a boundary condition such that the electron spectrum tends to zero at the maximal Lorentz factor $\gamma_{\rm max}$: $N_{\rm e,0}(\gamma=\gamma_{\rm max})=0$. With this condition, this equation has the following solution:

$$N_{\rm e,0}(\gamma)=\frac{1}{b_{\rm c}\gamma^2}\int_{\gamma}^{\gamma_{\rm max}}Q_{\rm inj}(\gamma')\cdot\exp\left(\frac{1/\gamma'-1/\gamma}{b_{\rm c}t_{\rm esc}}\right)d\gamma' \tag{A.3}$$

The multiplicative term in this expression and the negative exponent in the exponential describe how the injection effect is respectively damped by the cooling and by the escape. Now Let us consider how this spectrum is modified with time when the shock acceleration is acting on this electron population. Let us decompose the electron spectrum in the emitting zone during the passage of the shock into the initial and perturbed parts:

$$N_{\rm e,FI}(\gamma,t)=N_{\rm e,0}(\gamma)+N_{\rm e,p}(\gamma,t) \tag{A.4}$$

The initial electron spectrum is the steady state solution, and the time-dependent perturbed part is the one causing the flux increase. We plug this expression into the kinetic equation Eq. 6.13, which yields:

$$\begin{aligned}\frac{\partial N_{\rm e,p}(\gamma,t)}{\partial t}=&\frac{\partial}{\partial\gamma}\left(b_{\rm c}\gamma^2\cdot N_{\rm e,0}(\gamma)\right)-\frac{N_{\rm e,0}(\gamma)}{t_{\rm esc}}+Q_{\rm inj}(\gamma)-\\&-\frac{\partial}{\partial\gamma}\left(\frac{\gamma}{t_{\rm FI}}\cdot N_{\rm e,0}(\gamma)\right)+\frac{\partial}{\partial\gamma}\left(\left[b_{\rm c}\gamma^2-\frac{\gamma}{t_{\rm FI}}\right]N_{\rm e,p}(\gamma,t)\right)-\frac{N_{\rm e,p}(\gamma,t)}{t_{\rm esc}}\end{aligned} \tag{A.5}$$

The first three terms in the right hand side comprise the right hand side of the kinetic equation for the steady state (Eq. A.2) which is equal to zero. So, these terms disappear from the equation and we obtain:

$$\frac{\partial N_{\mathrm{e,p}}(\gamma,t)}{\partial t} = \frac{\partial}{\partial \gamma}\left[W(\gamma)\, N_{\mathrm{e,p}}(\gamma,t)\right] - \frac{N_{\mathrm{e,p}}(\gamma,t)}{t_{\mathrm{esc}}} - \frac{\partial}{\partial \gamma}\left(\frac{\gamma}{t_{\mathrm{FI}}} \cdot N_{\mathrm{e},0}(\gamma)\right) \quad \text{(A.6)}$$

where $W(\gamma) = b_{\mathrm{c}}\gamma^2 - \dfrac{\gamma}{t_{\mathrm{FI}}}$.

This equation describes the time evolution of the perturbed time-dependent addition $N_{\mathrm{e,p}}(\gamma,t)$ to the steady state solution. Here t_{FI} is assumed constant in time and energy-independent. Plugging the expression for $N_{\mathrm{e},0}(\gamma)$ (Eq. A.3) to the equation, and evaluating the last term (free term depending only on γ), we find:

$$\frac{\partial N_{\mathrm{e,p}}(\gamma,t)}{\partial t} = \frac{\partial}{\partial \gamma}\left[W(\gamma)\, N_{\mathrm{e,p}}(\gamma,t)\right] - \frac{N_{\mathrm{e,p}}(\gamma,t)}{t_{\mathrm{esc}}} + \mathcal{F}(\gamma) \quad \text{(A.7)}$$

with $\mathcal{F}(\gamma) = \left[1 - \frac{1}{b_{\mathrm{c}}\gamma\, t_{\mathrm{esc}}}\right] \cdot \dfrac{N_{\mathrm{e},0}(\gamma)}{t_{\mathrm{FI}}} + \dfrac{Q_{\mathrm{inj}}(\gamma)}{b_{\mathrm{c}}\gamma\, t_{\mathrm{FI}}}$.

We obtained the final form of the equation governing how $N_{\mathrm{e,p}}(\gamma,t)$ is evolving with time. The function $\mathcal{F}(\gamma)$ can be considered as a complex injection function composed of two terms: a scaled steady-state electron spectrum and a scaled injection spectrum. From Eq. A.1 and A.4 we deduce that the initial condition for $N_{\mathrm{e,p}}(\gamma,t)$:

$$N_{\mathrm{e,p}}(\gamma, t=0) = 0 \quad \text{(A.8)}$$

A.2 Solving by characteristics

We use the method of characteristics to solve the equation Eq. A.7. We first search for characteristic curves in the γ-t space along which the equation for $N_{\mathrm{e,p}}(\gamma,t)$ becomes an ordinary differential equation. Then we solve this equation along a characteristic curve. Let us rewrite the Eq. A.7 in the following form (expanding the partial derivative over γ):

$$\frac{\partial N_{\mathrm{e,p}}(\gamma,t)}{\partial t} + (-1)\cdot W(\gamma)\frac{\partial N_{\mathrm{e,p}}(\gamma,t)}{\partial \gamma} = \mathcal{F}(\gamma) - \frac{N_{\mathrm{e,p}}(\gamma,t)}{\tau(\gamma)} \quad \text{(A.9)}$$

where $\frac{1}{\tau(\gamma)} = \frac{1}{t_{\rm esc}} + \frac{1}{t_{\rm FI}} - 2b_{\rm c}\gamma$.

Let us consider a characteristic curve $(\gamma(t), t)$. The left hand side of the equation can be now represented as a full derivative of $N_{\rm e,p}(\gamma(t), t)$ with respect to time, and also as a directional derivative of $N_{\rm e,p}(\gamma(t), t)$ in the direction of $(-W(\gamma)\,,\,1)$ in the γ-t plane. By the chain rule, we have:

$$\frac{dN_{\rm e,p}(\gamma(t),t)}{dt} = \frac{\partial N_{\rm e,p}(\gamma(t),t)}{\partial t} + \frac{d\gamma(t)}{dt}\cdot\frac{\partial N_{\rm e,p}(\gamma(t),t)}{\partial \gamma} \tag{A.10}$$

We see that along the characteristic curve $(\gamma(t), t)$ our equation in partial derivatives transforms into an ordinary differential equation:

$$\frac{dN_{\rm e,p}(\gamma(t),t)}{dt} = \mathcal{F}(\gamma) - \frac{N_{\rm e,p}(\gamma(t)\,,t)}{\tau(\gamma)} \tag{A.11}$$

$$\frac{d\gamma(t)}{dt} = -W(\gamma) \tag{A.12}$$

Let us solve the Eq. A.12 for the characteristic curve in the γ-t space. We choose an initial point on our characteristic as (ξ,0), so the equation has to satisfy the boundary condition $\gamma(t=0) = \xi$. The solution of the Eq. A.12 with this boundary condition is:

$$\gamma_{(\xi)}(t) = \frac{1}{b_{\rm c}t_{\rm FI}(1-e^{-t/t_{\rm FI}}) + \frac{1}{\xi}e^{-t/t_{\rm FI}}} \tag{A.13}$$

This formula defines a characteristic curve in the γ-t space. For given γ and t, Let us find the starting Lorentz factor ξ of the characteristic that passes through point (γ,t):

$$\xi = \xi(\gamma, t) = \frac{\gamma e^{-t/t_{\rm FI}}}{1 - b_{\rm c}t_{\rm FI}\gamma\,(1-e^{-t/t_{\rm FI}})} \tag{A.14}$$

Now let us solve the initial value problem (Eq. A.11 and A.8). We restrict the $N_{\rm e,p}(\gamma, t)$ to the characteristic (Eq. A.13), noting $N_{\rm e,p}(\gamma_{(\xi)}(t), t) \Rightarrow u(t)$ at a given ξ and solve the differential equation Eq. A.11 along the characteristic curve. We have:

$$\frac{du(t)}{dt} + \frac{u(t)}{\tau(\gamma_{(\xi)}(t))} = \mathcal{F}(\gamma_{(\xi)}(t)) \tag{A.15}$$

This is a simple linear non-homogeneous first order differential equation, which can be solved with the help of an integrating factor. The equation has the following general solution:

$$u(t) = \frac{1}{\mu(t)} \left[\int_0^t \mu(t')\, \mathcal{F}(\gamma_{(\xi)}(t'))\, dt' \, + \, C \right] \tag{A.16}$$

with the integrating factor:

$$\mu(t) = e^{\int \frac{1}{\tau(\gamma_{(\xi)}(t))}\, dt} \tag{A.17}$$

From the initial condition Eq. A.8 which is $u(t=0) = 0$, we get the constant of integration $C = 0$.

Now let us calculate the $\mu(t)$ function. First we evaluate the exponent in A.17:

$$\begin{aligned} \int \frac{1}{\tau(\gamma_{(\xi)}(t))}\, dt \, &= \int \left(\frac{1}{t_{\rm esc}} + \frac{1}{t_{\rm FI}} \right) dt \, - \, 2b_{\rm c} \int \frac{1}{b_{\rm c} t_{\rm FI} + (1/\xi - b_{\rm c} t_{\rm FI})\, e^{-t/t_{\rm FI}}}\, dt \, = \\ &= \left[\frac{1}{t_{\rm esc}} + \frac{1}{t_{\rm FI}} \right] t \, - \, 2 \ln \left[1/\xi \, + \, b_{\rm c}\, t_{\rm FI}\, (e^{t/t_{\rm FI}} - 1) \right] \end{aligned} \tag{A.18}$$

The integrating factor $\mu(t)$ is then:

$$\mu(t) = \frac{e^{(1/t_{\rm esc} + 1/t_{\rm FI}) \cdot t}}{\left[1/\xi \, + \, b_{\rm c} t_{\rm FI} (e^{t/t_{\rm FI}} - 1) \right]^2} \tag{A.19}$$

A.3 Final solution

The transition from $u(t)$ back to $N_{\rm e,p}(\gamma, t)$ is achieved by substitution of $\xi = \xi(\gamma, t)$ to the expression for $u(t)$ (Eq. A.16): $N_{\rm e,p}(\gamma, t) = u(t)|_{\xi = \xi(\gamma,t)}$.

$$N_{\rm e,p}(\gamma, t) = u(t)|_{\xi=\xi(\gamma,t)} = \int_0^t \left[\frac{\mu(t')}{\mu(t)}\, \mathcal{F}(\gamma_{(\xi)}(t')) \right]_{|_{\xi=\xi(\gamma,t)}} dt' \tag{A.20}$$

APPENDIX A. ANALYTICAL SOLUTION OF THE KINETIC EQUATION FOR THE CASE OF SHOCK PERTURBING A STEADY-STATE ELECTRON SPECTRUM

Let us proceed with the substitution $\xi = \xi(\gamma, t)$ to the components of the integrand.

First we evaluate $\mu(t)|_{\xi=\xi(\gamma,t)}$. Substituting the expression for the initial Lorentz factor $\xi = \xi(\gamma, t)$ from Eq. A.14, we get:

$$\mu(t)|_{\xi=\xi(\gamma,t)} = \mu(t,\gamma) = \gamma^2\, e^{(1/t_{\rm esc} - 1/t_{\rm FI})\cdot t} \tag{A.21}$$

Next, we compute the form of the $\mu(t')|_{\xi=\xi(\gamma,t)}$, again substituting the expression for ξ from Eq. A.14:

$$\mu(t')|_{\xi=\xi(\gamma,t)} \;=\; \mu(t',\gamma,t) \;=\; \frac{e^{t'/t_{\rm esc} + (t'-2t)/t_{\rm FI}}}{\left[1/\gamma \;+\; b_{\rm c}\, t_{\rm FI}\, (e^{(t'-t)/t_{\rm FI}} - 1)\right]^2} \tag{A.22}$$

Then we have to calculate the Lorentz factor $\gamma_{(\xi)}(t')|_{\xi=\xi(\gamma,t)}$ that appears in the function $\mathcal{F}(\gamma_{(\xi)}(t'))$. We use Eq. A.13, A.14 and after simple and obvious transformations we obtain:

$$\gamma_{(\xi)}(t')|_{\xi=\xi(\gamma,t)} = \Gamma(\gamma,t,t') = \frac{\gamma\cdot e^{(t'-t)/t_{\rm FI}}}{1 \;+\; \gamma b_{\rm c}\, t_{\rm FI}\, (e^{(t'-t)/t_{\rm FI}} \;-\; 1)} \tag{A.23}$$

We note that the denominator of $\mu(t',\gamma,t)$ in Eq. A.22 multiplied by γ^2 is exactly the square of denominator of the $\Gamma(\gamma,t,t')$, so for the sake of simplicity we express $\mu(t',\gamma,t)$ via $\Gamma(\gamma,t,t')$:

$$\mu(t',\gamma,t) = \Gamma(\gamma,t,t')^2\cdot e^{(1/t_{\rm esc} - 1/t_{\rm FI})\cdot t'} \tag{A.24}$$

Now we evaluate the expression under the integral in Eq. A.20, using previously derived components, where the substitution was done (Eq. A.21, A.24 and A.23):

$$\begin{aligned}\left[\frac{\mu(t')}{\mu(t)}\,\mathcal{F}(\gamma_{(\xi)}(t'))\right]_{|_{\xi=\xi(\gamma,t)}} &= \frac{\Gamma(\gamma,t,t')^2}{\gamma^2}\cdot e^{(1/t_{\rm esc} - 1/t_{\rm FI})\cdot(t'-t)}\times\\ \times\left[\frac{N_{\rm e,0}(\Gamma(\gamma,t,t'))}{t_{\rm FI}}\cdot\left(1-\frac{1}{b_{\rm c}t_{\rm esc}\Gamma(\gamma,t,t')}\right) + \frac{Q_{\rm inj}(\Gamma(\gamma,t,t'))}{b_{\rm c}t_{\rm FI}\Gamma(\gamma,t,t')}\right] &=\\ = \frac{\Gamma(\gamma,t,t')\cdot e^{(1/t_{\rm esc} - 1/t_{\rm FI})\cdot(t'-t)}}{b_{\rm c}\, t_{\rm FI}\,\gamma^2}\times&\\ \times\left[Q_{\rm inj}(\Gamma(\gamma,t,t')) \;+\; \left(b_{\rm c}\Gamma(\gamma,t,t') \;-\; \frac{1}{t_{\rm esc}}\right)\cdot N_{\rm e,0}(\Gamma(\gamma,t,t'))\right]&\end{aligned} \tag{A.25}$$

APPENDIX A. ANALYTICAL SOLUTION OF THE KINETIC EQUATION FOR THE CASE OF SHOCK PERTURBING A STEADY-STATE ELECTRON SPECTRUM

We can now write down the final solution for the total electron spectrum $N_{\rm e,FI}(\gamma, t)$ using Eq. A.4, A.20 and A.25:

$$N_{\rm e,FI}(\gamma,t) = N_{\rm e,0}(\gamma) \; + \; \int_0^t \frac{\Gamma(\gamma,t,t') \cdot e^{(1/t_{\rm esc} \, - \, 1/t_{\rm FI}) \cdot (t' \, - \, t)}}{b_{\rm c} \, t_{\rm FI} \, \gamma^2} \times$$
$$\times \; \left[Q_{\rm inj}(\Gamma(\gamma,t,t')) \; + \; \left(b_{\rm c} \Gamma(\gamma,t,t') \; - \; \frac{1}{t_{\rm esc}} \right) \cdot N_{\rm e,0}(\Gamma(\gamma,t,t')) \right] \, dt' \quad \text{(A.26)}$$

Let us explore the final solution. At the moment when the shock just enters the blob ($t = 0$), the electron spectrum is, as expected, the steady-state solution. Also, when the shock acceleration is extremely weak ($t_{\rm FI} \to \infty$), we see that the electron spectrum will remain the steady-state one and not evolve in time, which is in agreement with the expectations (very weak shock will not perturb the electron spectrum).

APPENDIX A. ANALYTICAL SOLUTION OF THE KINETIC EQUATION FOR THE CASE OF SHOCK PERTURBING A STEADY-STATE ELECTRON SPECTRUM

References

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010a, Nature, 463, 919

Abdo, A. A., Ackermann, M., Agudo, I., et al. 2010b, ApJ, 716, 30

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2011, ApJ, 736, 131

Abeysekara, A. U., Archambault, S., Archer, A., et al. 2015, ApJ, 815, L22

Abeysekara, A. U., Benbow, W., Bird, R., et al. 2020, ApJ, 890, 97

Abraham, Z., & Romero, G. E. 1999, A&A, 344, 61

Acciari, V. A., Aliu, E., Arlen, T., et al. 2011a, ApJ, 730, L20

—. 2011b, ApJ, 738, 25

Acciari, V. A., Arlen, T., Aune, T., et al. 2014, Astroparticle Physics, 54, 1

Achterberg, A. 1979, A&A, 76, 276

Aharonian, F., Akhperjanian, A., Beilicke, M., et al. 2002, A&A, 393, 89

Aharonian, F., Akhperjanian, A. G., Aye, K. M., et al. 2004, A&A, 425, L13

Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2007, ApJ, 664, L71

Aharonian, F. A. 2000, , 5, 377

Aharonian, F. A., Khangulyan, D., & Costamante, L. 2008, MNRAS, 387, 1206

Ahnen, M. L., Ansoldi, S., Antonelli, L. A., et al. 2016, A&A, 593, A91

Albert, J., Aliu, E., Anderhub, H., et al. 2007a, ApJ, 663, 125

—. 2007b, ApJ, 669, 862

—. 2007c, ApJ, 665, L51

Aleksić, J., Ansoldi, S., Antonelli, L. A., et al. 2015, A&A, 578, A22

Anderhub, H., Backes, M., Biland, A., et al. 2013, Journal of Instrumentation, 8, P06008

Andreopoulos, Y., Agui, J. H., & Briassulis, G. 2000, Annual Review of Fluid Mechanics, 32, 309

Antonucci, R. R. J., & Miller, J. S. 1985, ApJ, 297, 621

Armstrong, P., Chadwick, P. M., Cottle, P. J., et al. 1999, Experimental Astronomy, 9, 51

Asano, K., & Hayashida, M. 2018, The Astrophysical Journal, 861, 31

Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, ApJ, 697, 1071

Baixeras, C., Bastieri, D., Bigongiari, C., et al. 2004, Nuclear Instruments and Methods in Physics Research A, 518, 188

Barkov, M. V., Aharonian, F. A., Bogovalov, S. V., Kelner, S. R., & Khangulyan, D. 2012, ApJ, 749, 119

Barrau, A., Bazer-Bachi, R., Beyer, E., et al. 1998, Nuclear Instruments and Methods in Physics Research A, 416, 278

Becherini, Y., Djannati-Ataï, A., Marandon, V., Punch, M., & Pita, S. 2011, Astroparticle Physics, 34, 858

Becker, P. A., Le, T., & Dermer, C. D. 2006, ApJ, 647, 539

Bell, A. R. 1978, MNRAS, 182, 147

Blandford, R., Meier, D., & Readhead, A. 2019, ARA&A, 57, 467

Blandford, R. D., & Payne, D. G. 1982, MNRAS, 199, 883

Blandford, R. D., & Znajek, R. L. 1977, MNRAS, 179, 433

Błażejowski, M., Sikora, M., Moderski, R., & Madejski, G. M. 2000, ApJ, 545, 107

Błażejowski, M., Blaylock, G., Bond, I. H., et al. 2005, ApJ, 630, 130

Bloom, S. D., & Marscher, A. P. 1996, ApJ, 461, 657

Boettcher, M. 2012, arXiv e-prints, arXiv:1205.0539

Böttcher, M., & Dermer, C. D. 2002, ApJ, 564, 86

Böttcher, M., Reimer, A., & Marscher, A. P. 2009, ApJ, 703, 1168

Böttcher, M., Reimer, A., Sweeney, K., & Prakash, A. 2013, ApJ, 768, 54

Broderick, A. E., Chang, P., & Pfrommer, C. 2012, ApJ, 752, 22

Bromberg, O., & Tchekhovskoy, A. 2016, MNRAS, 456, 1739

Burn, B. J. 1975, A&A, 45, 435

Cao, G., & Wang, J. 2013, PASJ, 65, 109

Casadio, C., Gómez, J. L., Jorstad, S. G., et al. 2015, ApJ, 813, 51

Catalano, O., Capalbi, M., Gargano, C., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII, 1070237

Catanese, M., Bradbury, S. M., Breslin, A. C., et al. 1997, ApJ, 487, L143

Cawley, M. F., Fegan, D. J., Harris, K., et al. 1990, Experimental Astronomy, 1, 173

Celotti, A., & Ghisellini, G. 2008, MNRAS, 385, 283

Cerruti, M., Zech, A., Boisson, C., & Inoue, S. 2015, MNRAS, 448, 910

Chadwick, P. M., Dickinson, M. R., Dipper, N. A., et al. 1998, Astroparticle Physics, 9, 131

Chadwick, P. M., Lyons, K., McComb, T. J. L., et al. 1999, ApJ, 513, 161

Chang, J., & Cooper, G. 1970, Journal of Computational Physics, 6, 1

Cherenkov Telescope Array Consortium, Acharya, B. S., Agudo, I., et al. 2019, Science with the Cherenkov Telescope Array, doi:10.1142/10986

Chiaberge, M., & Ghisellini, G. 1999, MNRAS, 306, 551

Chitnis, V. R., Pendharkar, J. K., Bose, D., et al. 2009, The Astrophysical Journal, 698, 1207–1220

Coppi, P. S., & Blandford, R. D. 1990, MNRAS, 245, 453

Coroniti, F. V., & Eviatar, A. 1977, ApJS, 33, 189

Courvoisier, T. J. L. 2013, High Energy Astrophysics, doi:10.1007/978-3-642-30970-0

Crusius, A., & Schlickeiser, R. 1986, A&A, 164, L16

Curtis, H. D. 1918, Publications of Lick Observatory, 13, 9

Cutini, S. 2015, The Astronomer's Telegram, 7633, 1

Czerny, B. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 360, AGN Variability from X-Rays to Radio Waves, ed. C. M. Gaskell, I. M. McHardy, B. M. Peterson, & S. G. Sergeev, 265

Daum, A., Hermann, G., Heß, M., et al. 1997, Astroparticle Physics, 8, 1

de Gouveia Dal Pino, E. M., Piovezan, P. P., & Kadowaki, L. H. S. 2010, A&A, 518, A5

De Naurois, M. 2012, Habilitation à diriger des recherches, Université Pierre et Marie Curie - Paris VI

de Naurois, M., & Rolland, L. 2009, Astroparticle Physics, 32, 231

Dermer, C. D., & Giebels, B. 2016, Comptes Rendus Physique, 17, 594

Dermer, C. D., Miller, J. A., & Li, H. 1996, ApJ, 456, 106

Dermer, C. D., & Schlickeiser, R. 1993, ApJ, 416, 458

—. 2002, ApJ, 575, 667

Dermer, C. D., Schlickeiser, R., & Mastichiadis, A. 1992, A&A, 256, L27

Dmytriiev, A., Sol, H., & Zech, A. 2019a, in International Cosmic Ray Conference, Vol. 36, 36th International Cosmic Ray Conference (ICRC2019), 662

Dmytriiev, A., Dangeon, L., Fasola, G., et al. 2019b, in International Cosmic Ray Conference, Vol. 36, 36th International Cosmic Ray Conference (ICRC2019), 663

Domínguez, A., Primack, J. R., Rosario, D. J., et al. 2011, MNRAS, 410, 2556

Dournaux, J.-L., Huet, J.-M., Amans, J.-P., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9151, Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation, 915102

Dournaux, J. L., Amans, J. P., Dangeon, L., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9912, Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation II, 99124Z

Emery, G., Cerruti, M., Dmytriiev, A., et al. 2019, in International Cosmic Ray Conference, Vol. 36, 36th International Cosmic Ray Conference (ICRC2019), 668

Fath, E. A. 1909, Lick Observatory Bulletin, 149, 71

Fermi, E. 1949, Physical Review, 75, 1169

Fortson, L., VERITAS Collaboration, & Fermi-LAT Collaborators. 2012, in American Institute of Physics Conference Series, Vol. 1505, High Energy Gamma-Ray Astronomy: 5th International Meeting on High Energy Gamma-Ray Astronomy, ed. F. A. Aharonian, W. Hofmann, & F. M. Rieger, 514–517

Fossati, G., Maraschi, L., Celotti, A., Comastri, A., & Ghisellini, G. 1998, MNRAS, 299, 433

Fossati, G., Buckley, J. H., Bond, I. H., et al. 2008, ApJ, 677, 906

Fraija, N., Cabrera, J. I., Benítez, E., & Hiriart, D. 2015, arXiv e-prints, arXiv:1508.01438

Franceschini, A., Rodighiero, G., & Vaccari, M. 2008, A&A, 487, 837

Furth, H. P., Killeen, J., & Rosenbluth, M. N. 1963, Physics of Fluids, 6, 459

Gaidos, J. A., Akerlof, C. W., Biller, S., et al. 1996, Nature, 383, 319

Galtier, S. 2016, Introduction to Modern Magnetohydrodynamics

Garrigoux, T. 2015, Theses, Université Pierre et Marie Curie - Paris VI

Gedalin, M. 1993, Phys. Rev. E, 47, 4354

Ghisellini, G., Righi, C., Costamante, L., & Tavecchio, F. 2017, MNRAS, 469, 255

Ghisellini, G., & Svensson, R. 1991, MNRAS, 252, 313

Ghisellini, G., Tavecchio, F., & Chiaberge, M. 2005, A&A, 432, 401

Ghisellini, G., Tavecchio, F., Foschini, L., et al. 2010, MNRAS, 402, 497

Giannios, D., Uzdensky, D. A., & Begelman, M. C. 2009, MNRAS, 395, L29

Ginzburg, V. L., & Syrovatskii, S. I. 1964, The Origin of Cosmic Rays

Giovannini, G., Taylor, G. B., Arbizzani, E., et al. 1999, ApJ, 522, 101

Gould, R. J. 1979, A&A, 76, 306

Goyal, A. 2020, MNRAS, 494, 3432

Grimm, H. J., Gilfanov, M., & Sunyaev, R. 2002, A&A, 391, 923

Gu, M., Cao, X., & Jiang, D. R. 2001, MNRAS, 327, 1111

H. E. S. S. Collaboration, Abramowski, A., Acero, F., et al. 2010, A&A, 520, A83

—. 2013, A&A, 554, A107

H. E. S. S. Collaboration, Abramowski, A., Aharonian, F., et al. 2015, Science, 347, 406

H. E. S. S. Collaboration, Abdalla, H., Abramowski, A., et al. 2017, A&A, 598, A39

H. E. S. S. Collaboration, Abdalla, H., Adam, R., et al. 2020a, Nature, 582, 356

—. 2020b, Nature, 582, 356

Harris, J., & Chadwick, P. M. 2014, , 2014, 018

Hartman, R. C., Bertsch, D. L., Fichtel, C. E., et al. 1992, ApJ, 385, L1

Haugbølle, T. 2011, ApJ, 739, L42

Hayashida, M., Madejski, G. M., Nalewajko, K., et al. 2012, ApJ, 754, 114

Hayashida, M., Nalewajko, K., Madejski, G. M., et al. 2015, ApJ, 807, 79

Henri, G., & Pelletier, G. 1991, ApJ, 383, L7

Hervet, O., Boisson, C., & Sol, H. 2015, A&A, 578, A69

HESS Collaboration, Abramowski, A., Aharonian, F., et al. 2016, Nature, 531, 476

Hillas, A. M. 1985, in International Cosmic Ray Conference, Vol. 3, 19th International Cosmic Ray Conference (ICRC19), Volume 3, 445

Holder, J., Atkins, R. W., Badran, H. M., et al. 2006, Astroparticle Physics, 25, 391

Isobe, N., Koyama, S., Kino, M., et al. 2017, ApJ, 850, 193

Jaekel, U., & Schlickeiser, R. 1992, Journal of Physics G: Nuclear and Particle Physics, 18, 1089

Jaffe, W., Ford, H. C., Ferrarese, L., van den Bosch, F., & O'Connell, R. W. 1993, Nature, 364, 213

Jones, F. C. 1968, Physical Review, 167, 1159

Jorstad, S. G., Marscher, A. P., Lister, M. L., et al. 2004, AJ, 127, 3115

Kadowaki, L. H. S., de Gouveia Dal Pino, E. M., & Singh, C. B. 2015, ApJ, 802, 113

Kardashev, N. S. 1962, Soviet Ast., 6, 317

Karimabadi, H., Roytershteyn, V., Vu, H. X., et al. 2014, Physics of Plasmas, 21, 062308

Kataoka, J., Mattox, J. R., Quinn, J., et al. 1999, ApJ, 514, 138

Katarzyński, K., Sol, H., & Kus, A. 2001, Astronomy and Astrophysics, 367, 809

—. 2003, A&A, 410, 101

Kazanas, D., & Mastichiadis, A. 1999, ApJ, 518, L17

Kerr, R. P. 1963, Phys. Rev. Lett., 11, 237

Khiali, B., de Gouveia Dal Pino, E. M., & Sol, H. 2015, arXiv e-prints, arXiv:1504.07592

Kim, J. Y., Krichbaum, T. P., Lu, R. S., et al. 2018, A&A, 616, A188

Kirk, J. G., Guthmann, A. W., Gallant, Y. A., & Achterberg, A. 2000, ApJ, 542, 235

Kirk, J. G., Rieger, F. M., & Mastichiadis, A. 1998, A&A, 333, 452

Krennrich, F., Bond, I. H., Bradbury, S. M., et al. 2002, ApJ, 575, L9

Kruells, W. M. 1992, A&A, 260, 49

Larionov, V. M., Jorstad, S. G., Marscher, A. P., et al. 2008, A&A, 492, 389

Larionov, V. M., Villata, M., Raiteri, C. M., et al. 2016, MNRAS, 461, 3047

Lazarian, A., & Vishniac, E. T. 1999, ApJ, 517, 700

Le Blanc, O., Fasola, G., Huet, J. M., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10700, Ground-based and Airborne Telescopes VII, 1070010

Le Bohec, S., Degrange, B., Punch, M., et al. 1998, Nuclear Instruments and Methods in Physics Research A, 416, 425

Lense, J., & Thirring, H. 1918, Physikalische Zeitschrift, 19, 156

Levinson, A., & Rieger, F. 2011, ApJ, 730, 123

Li, Y., Wiita, P. J., Schuh, T., Elghossain, G., & Hu, S. 2018, ApJ, 869, 32

Liska, M. 2019, Phd thesis, University of Amsterdam

Liu, F. K., Liu, B. F., & Xie, G. Z. 1997, A&AS, 123, 569

Longair, M. S. 2011, High Energy Astrophysics

Loureiro, N. F., Schekochihin, A. A., & Cowley, S. C. 2007, Physics of Plasmas, 14, 100703

Madsen, K. K., Reynolds, S., Harrison, F., et al. 2015, ApJ, 801, 66

MAGIC Collaboration, Albert, J., Aliu, E., et al. 2008, Science, 320, 1752

MAGIC Collaboration, Ahnen, M. L., Ansoldi, S., et al. 2016, , 2016, 039

MAGIC Collaboration, Acciari, V. A., Ansoldi, S., et al. 2019, Nature, 575, 459

MAGIC collaboration, Acciari, V. A., Ansoldi, S., et al. 2020, arXiv e-prints, arXiv:2001.08678

Malyshkin, L. M., Linde, T., & Kulsrud, R. M. 2005, Physics of Plasmas, 12, 102902

Mannheim, K. 1993, A&A, 269, 67

Maraschi, L., Ghisellini, G., & Celotti, A. 1992, ApJ, 397, L5

Marscher, A. P. 2014, ApJ, 780, 87

Marscher, A. P., & Gear, W. K. 1985, ApJ, 298, 114

Marscher, A. P., Jorstad, S. G., D'Arcangelo, F. D., et al. 2008, Nature, 452, 966

Marscher, A. P., Jorstad, S. G., Larionov, V. M., et al. 2010, ApJ, 710, L126

Marshall, H. L., Miller, B. P., Davis, D. S., et al. 2002, ApJ, 564, 683

Marziani, P., Sulentic, J. W., Dultzin-Hacyan, D., Calvani, M., & Moles, M. 1996, ApJS, 104, 37

Mastichiadis, A., & Kirk, J. G. 1997, A&A, 320, 19

McEnery, J., Bond, I. H., Boyle, P. J., et al. 1997, in International Cosmic Ray Conference, Vol. 3, International Cosmic Ray Conference, 257

McKinney, J. C., & Narayan, R. 2007, MNRAS, 375, 531

Mead, A. R. G., Ballard, K. R., Brand, P. W. J. L., et al. 1990, A&AS, 83, 183

Meliani, Z., & Keppens, R. 2009, ApJ, 705, 1594

Mertens, F., Lobanov, A. P., Walker, R. C., & Hardee, P. E. 2016, A&A, 595, A54

Migliori, G., Orienti, M., Coccato, L., et al. 2020, MNRAS, 495, 1593

Moderski, R., Sikora, M., Coppi, P. S., & Aharonian, F. 2005, MNRAS, 363, 954

Morris, P. J., Potter, W. J., & Cotter, G. 2019, MNRAS, 486, 1548

Mouri, H., Hori, A., & Takaoka, M. 2009, Physics of Fluids, 21, 065107

Mücke, A., & Protheroe, R. J. 2001, Astroparticle Physics, 15, 121

Mücke, A., Protheroe, R. J., Engel, R., Rachen, J. P., & Stanev, T. 2003, Astroparticle Physics, 18, 593

Neronov, A., Semikoz, D., Aharonian, F., & Kalashev, O. 2002, Phys. Rev. Lett., 89, 051101

Neronov, A., & Vovk, I. 2010, Science, 328, 73

Okumura, A., Noda, K., & Rulten, C. 2016, Astroparticle Physics, 76, 38

Orienti, M., Brunetti, G., Nagai, H., et al. 2017, MNRAS, 469, L123

Orienti, M., Prieto, M. A., Brunetti, G., et al. 2012, MNRAS, 419, 2338

Ostrowski, M., & Bednarz, J. 2002, A&A, 394, 1141

O'Sullivan, S., Reville, B., & Taylor, A. M. 2009, MNRAS, 400, 248

Paliya, V. S., Böttcher, M., Diltz, C., et al. 2015a, ApJ, 811, 143

Paliya, V. S., Sahayanathan, S., & Stalin, C. S. 2015b, ApJ, 803, 15

Park, B. T., & Petrosian, V. 1996, ApJs, 103, 255

Parker, E. N. 1957, J. Geophys. Res., 62, 509

Penna, R. F., Narayan, R., & Sadowski, A. e. 2013, MNRAS, 436, 3741

Petschek, H. E. 1964, Magnetic Field Annihilation, Vol. 50, 425

Pittori, C., Cutini, S., Gasparrini, D., et al. 2008, The Astronomer's Telegram, 1583, 1

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1989, Numerical recipes in C. The art of scientific computing

Ptitsyna, K., & Neronov, A. 2016, A&A, 593, A8

Punch, M., Akerlof, C. W., Cawley, M. F., et al. 1992, Nature, 358, 477

Raiteri, C. M., Villata, M., Acosta-Pulido, J. A., et al. 2017, Nature, 552, 374

Rulten, C., Zech, A., Okumura, A., Laporte, P., & Schmoll, J. 2016, Astroparticle Physics, 82, 36

Rybicki, G. B., & Lightman, A. P. 1979, Radiative processes in astrophysics

Sahu, S., Miranda, L. S., & Rajpoot, S. 2016, European Physical Journal C, 76, 127

Salpeter, E. E. 1964, ApJ, 140, 796

Sambruna, R. M., Aharonian, F. A., Krawczynski, H., et al. 2000, ApJ, 538, 127

Schlickeiser, R. 1989, ApJ, 336, 243

Schlickeiser, R., Krakau, S., & Supsar, M. 2013, ApJ, 777, 49

Schmidt, M. 1963, Nature, 197, 1040

Shukla, A., Chitnis, V. R., Vishwanath, P. R., et al. 2012, Astronomy and Astrophysics, 541, A140

Sikora, M., Begelman, M. C., Madejski, G. M., & Lasota, J.-P. 2005, ApJ, 625, 72

Sikora, M., Begelman, M. C., & Rees, M. J. 1994, ApJ, 421, 153

Sikora, M., Błażejowski, M., Begelman, M. C., & Moderski, R. 2001, ApJ, 554, 1

Singh, K. K., Yadav, K. K., Chandra, P., et al. 2015, Astroparticle Physics, 61, 32

Singh, S., Solak, H., & Cerrina, F. 1996, Review of Scientific Instruments, 67, 3355

Sinha, A., Shukla, A., Misra, R., et al. 2015, A&A, 580, A100

Sironi, L., Keshet, U., & Lemoine, M. 2015, Space Sci. Rev., 191, 519

Sironi, L., & Spitkovsky, A. 2010, The Astrophysical Journal, 726, 75

Sironi, L., & Spitkovsky, A. 2014, ApJ, 783, L21

Skilling, J. 1975, MNRAS, 172, 557

Sol, H., Greenshaw, T., Le Blanc, O., White, R., & GCT project, C. 2017, in International Cosmic Ray Conference, Vol. 301, 35th International Cosmic Ray Conference (ICRC2017), 822

Sol, H., Pelletier, G., & Asseo, E. 1989, MNRAS, 237, 411

Sol, H., Zech, A., Boisson, C., et al. 2013, Astroparticle Physics, 43, 215

Spada, M., Ghisellini, G., Lazzati, D., & Celotti, A. 2001, MNRAS, 325, 1559

Spitkovsky, A. 2007, The Astrophysical Journal, 673, L39

Sweet, P. A. 1958, in IAU Symposium, Vol. 6, Electromagnetic Phenomena in Cosmical Physics, ed. B. Lehnert, 123

Tammi, J., & Duffy, P. 2009, MNRAS, 393, 1063

Tluczykont, M. 2011, arXiv e-prints, arXiv:1106.1035

Tramacere, A., Massaro, E., & Taylor, A. M. 2011, ApJ, 739, 66

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803

Vassiliev, V., Fegan, S., & Brousseau, P. 2007, Astroparticle Physics, 28, 10

VERITAS Collaboration, Aliu, E., Arlen, T., et al. 2011, Science, 334, 69

Villata, M., & Raiteri, C. M. 1999, A&A, 347, 30

Villata, M., Raiteri, C. M., Sillanpaa, A., & Takalo, L. O. 1998, MNRAS, 293, L13

Virtanen, J. J. P., & Vainio, R. 2005, ApJ, 621, 313

Völk, H. J., & Bernlöhr, K. 2009, Experimental Astronomy, 25, 173

Wagner, R. M. 2006, PhD thesis, Max-Planck-Institut für Physik, Föhringer Ring 6, 80805 München, Germany

Weekes, T. C., Cawley, M. F., Fegan, D. J., et al. 1989, ApJ, 342, 379

Wentzel, D. G. 1974, ARA&A, 12, 71

Woo, J.-H., & Urry, C. M. 2002, ApJ, 579, 530

Yan, D., Zhang, L., Yuan, Q., Fan, Z., & Zeng, H. 2013, ApJ, 765, 122

Zacharias, M., Böttcher, M., Jankowsky, F., et al. 2017, ApJ, 851, 72

Zech, A., Mazin, D., Biteau, J., et al. 2019, KSP: Active Galactic Nuclei, ed. B. S. Acharya, I. Agudo, I. Al Samarai, R. Alfaro, J. Alfaro, C. Alispach, R. Alves Batista, J. P. Amans, E. Amato, G. Ambrosi, E. Antolini, L. A. Antonelli, C. Aramo, M. Araya, T. Armstrong, F. Arqueros, L. Arrabito, K. Asano, M. Ashley, M. Backes, C. Balazs, M. Balbo, O. Ballester, J. Ballet, A. Bamba, M. Barkov, U. Barres de Almeida, J. A. Barrio, D. Bastieri, Y. Becherini, A. Belfiore, W. Benbow, D. Berge, E. Bernardini, M. G. Bernardini, M. Bernardos, K. Bernlöhr, B. Bertucci, B. Biasuzzi, C. Bigongiari, A. Biland, E. Bissaldi, J. Biteau, O. Blanch, J. Blazek, C. Boisson, J. Bolmont, G. Bonanno, A. Bonardi, C. Bonavolontà, G. Bonnoli, Z. Bosnjak, M. Böttcher, C. Braiding, J. Bregeon, A. Brill, A. M. Brown, P. Brun, G. Brunetti, T. Buanes, J. Buckley, V. Bugaev, R. Bühler, A. Bulgarelli, T. Bulik, M. Burton, A. Burtovoi, G. Busetto, R. Canestrari, M. Capalbi, F. Capitanio, A. Caproni, P. Caraveo, V. Cárdenas, C. Carlile, R. Carosi, E. Carquín, J. Carr, S. Casanova, E. Cascone, F. Catalani, O. Catalano, D. Cauz, M. Cerruti, P. Chadwick, S. Chaty, R. C. G. Chaves, A. Chen, X. Chen, M. Chernyakova, M. Chikawa, A. Christov, J. Chudoba, M. Cieślar, V. Coco, S. Colafrancesco, P. Colin, V. Conforti, V. Connaughton, J. Conrad, J. L. Contreras, J. Cortina, A. Costa, H. Costantini, G. Cotter, S. Covino, R. Crocker, J. Cuadra, O. Cuevas, P. Cumani, A. D'Aì, F. D'Ammando, P. D'Avanzo, D. D'Urso, M. Daniel, I. Davids, B. Dawson, F. Dazzi, A. De Angelis, R. de Cássia dos Anjos, G. De Cesare, A. De Franco, E. M. de Gouveia Dal Pino, I. de la Calle, R. de los Reyes Lopez, B. De Lotto, A. De Luca, M. De Lucia, M. de Naurois, E. de Oña Wilhelmi, F. De Palma, F. De Persio, V. de Souza, C. Deil, M. Del Santo, C. Delgado, D. della Volpe, T. Di Girolamo, F. Di Pierro, L. Di Venere, C. Díaz, C. Dib, S. Diebold, A. Djannati-Ataï, A. Domínguez, D. Dominis Prester, D. Dorner, M. Doro, H. Drass, D. Dravins, G. Dubus, V. V. Dwarkadas, J. Ebr, C. Eckner, K. Egberts, S. Einecke, T. R. N. Ekoume, D. Elsässer, J. P.

Ernenwein, C. Espinoza, C. Evoli, M. Fairbairn, D. Falceta-Goncalves, A. Falcone, C. Farnier, G. Fasola, E. Fedorova, S. Fegan, M. Fernand ez-Alonso, A. Fernández-Barral, G. Ferrand, M. Fesquet, M. Filipovic, V. Fioretti, G. Fontaine, M. Fornasa, L. Fortson, L. Freixas Coromina, C. Fruck, Y. Fujita, Y. Fukazawa, S. Funk, M. Füßling, S. Gabici, A. Gadola, Y. Gallant, B. Garcia, R. Garcia López, M. Garczarczyk, J. Gaskins, T. Gasparetto, M. Gaug, L. Gerard, G. Giavitto, N. Giglietto, P. Giommi, F. Giordano, E. Giro, M. Giroletti, A. Giuliani, J. F. Glicenstein, R. Gnatyk, N. Godinovic, P. Goldoni, G. Gómez-Vargas, M. M. González, J. M. González, D. Götz, J. Graham, P. Grand i, J. Granot, A. J. Green, T. Greenshaw, S. Griffiths, S. Gunji, D. Hadasch, S. Hara, M. J. Hardcastle, T. Hassan, K. Hayashi, M. Hayashida, M. Heller, J. C. Helo, G. Hermann, J. Hinton, B. Hnatyk, W. Hofmann, J. Holder, D. Horan, J. Hörandel, D. Horns, P. Horvath, T. Hovatta, M. Hrabovsky, D. Hrupec, T. B. Humensky, M. Hütten, M. Iarlori, T. Inada, Y. Inome, S. Inoue, T. Inoue, Y. Inoue, F. Iocco, K. Ioka, M. Iori, K. Ishio, Y. Iwamura, M. Jamrozy, P. Janecek, D. Jankowsky, P. Jean, I. Jung-Richardt, J. Jurysek, P. Kaaret, S. Karkar, H. Katagiri, U. Katz, N. Kawanaka, D. Kazanas, B. Khélifi, D. B. Kieda, S. Kimeswenger, S. Kimura, S. Kisaka, J. Knapp, J. Knödlseder, B. Koch, K. Kohri, N. Komin, K. Kosack, M. Kraus, M. Krause, F. Krauß, H. Kubo, G. Kukec Mezek, H. Kuroda, J. Kushida, N. La Palombara, G. Lamanna, R. G. Lang, J. Lapington, O. Le Blanc, S. Leach, J. P. Lees, J. Lefaucheur, M. A. Leigui de Oliveira, J. P. Lenain, R. Lico, M. Limon, E. Lindfors, T. Lohse, S. Lombardi, F. Longo, M. López, R. López-Coto, C. C. Lu, F. Lucarelli, P. L. Luque-Escamilla, E. Lyard, M. C. Maccarone, G. Maier, P. Majumdar, G. Malaguti, D. Mandat, G. Maneva, M. Manganaro, S. Mangano, A. Marcowith, J. Marín, S. Markoff, J. Martí, P. Martin, M. Martínez, G. Martínez, N. Masetti, S. Masuda, G. Maurin, N. Maxted, D. Mazin, C. Medina, A. Melandri, S. Mereghetti, M. Meyer, I. A. Minaya, N. Mirabal, R. Mirzoyan, A. Mitchell, T. Mizuno, R. Moderski, M. Mohammed, L. Mohrmann, T. Montaruli, A. Moralejo, D. Morcuende-Parrilla, K. Mori, G. Morlino, P. Morris, A. Morselli, E. Moulin, R. Mukherjee, C. Mundell, T. Murach, H. Muraishi, K. Murase, A. Nagai, S. Nagataki, T. Nagayoshi, T. Naito, T. Nakamori, Y. Nakamura, J. Niemiec, D. Nieto, M. Nikołajuk, K. Nishijima, K. Noda, D. Nosek, B. Novosyadlyj, S. Nozaki, P. O'Brien, L. Oakes, Y. Ohira, M. Ohishi, S. Ohm, N. Okazaki, A. Okumura, R. A. Ong, M. Orienti, R. Orito, J. P. Osborne, M. Ostrowski, N. Otte, I. Oya, M. Padovani, A. Paizis, M. Palatiello, M. Palatka, R. Paoletti, J. M. Paredes, G. Pareschi, R. D. Parsons, A. Pe'er, M. Pech, G. Pedaletti, M. Perri, M. Persic, A. Petrashyk, P. Petrucci, O. Petruk, B. Peyaud, M. Pfeifer, G. Piano, A. Pisarski, S. Pita, M. Pohl, M. Polo, D. Pozo, E. Prandini, J. Prast, G. Principe, D. Prokhorov, H. Prokoph, M. Prouza, G. Pühlhofer, M. Punch,

S. Pürckhauer, F. Queiroz, A. Quirrenbach, S. Rainò, S. Razzaque, O. Reimer, A. Reimer, A. Reisenegger, M. Renaud, A. H. Rezaeian, W. Rhode, D. Ribeiro, M. Ribó, T. Richtler, J. Rico, F. Rieger, M. Riquelme, S. Rivoire, V. Rizi, J. Rodriguez, G. Rodriguez Fernandez, J. J. Rodríguez Vázquez, G. Rojas, P. Romano, G. Romeo, J. Rosado, A. C. Rovero, G. Rowell, B. Rudak, A. Rugliancich, C. Rulten, I. Sadeh, S. Safi-Harb, T. Saito, N. Sakaki, S. Sakurai, G. Salina, M. Sánchez-Conde, H. Sandaker, A. Sandoval, P. Sangiorgi, M. Sanguillon, H. Sano, M. Santander, S. Sarkar, K. Satalecka, F. G. Saturni, E. J. Schioppa, S. Schlenstedt, M. Schneider, H. Schoorlemmer, P. Schovanek, A. Schulz, F. Schussler, U. Schwanke, E. Sciacca, S. Scuderi, I. Seitenzahl, D. Semikoz, O. Sergijenko, M. Servillat, A. Shalchi, R. C. Shellard, L. Sidoli, H. Siejkowski, A. Sillanpää, G. Sironi, J. Sitarek, V. Sliusar, A. Slowikowska, H. Sol, A. Stamerra, S. Stanič, R. Starling, Ł. Stawarz, S. Stefanik, M. Stephan, T. Stolarczyk, G. Stratta, U. Straumann, T. Suomijarvi, A. D. Supanitsky, G. Tagliaferri, H. Tajima, M. Tavani, F. Tavecchio, J. P. Tavernet, K. Tayabaly, L. A. Tejedor, P. Temnikov, Y. Terada, R. Terrier, T. Terzic, M. Teshima, V. Testa, S. Thoudam, W. Tian, L. Tibaldo, M. Tluczykont, C. J. Todero Peixoto, F. Tokanai, J. Tomastik, D. Tonev, M. Tornikoski, D. F. Torres, E. Torresi, G. Tosti, N. Tothill, G. Tovmassian, P. Travnicek, C. Trichard, M. Trifoglio, I. Troyano Pujadas, S. Tsujimoto, G. Umana, V. Vagelli, F. Vagnetti, M. Valentino, P. Vallania, L. Valore, C. van Eldik, J. Vand enbroucke, G. S. Varner, G. Vasileiadis, V. Vassiliev, M. Vázquez Acosta, M. Vecchi, A. Vega, S. Vercellone, P. Veres, S. Vergani, V. Verzi, G. P. Vettolani, A. Viana, C. Vigorito, J. Villanueva, H. Voelk, A. Vollhardt, S. Vorobiov, M. Vrastil, T. Vuillaume, S. J. Wagner, R. Wagner, R. Walter, J. E. Ward, D. Warren, J. J. Watson, F. Werner, M. White, R. White, A. Wierzcholska, P. Wilcox, M. Will, D. A. Williams, R. Wischnewski, M. Wood, T. Yamamoto, R. Yamazaki, S. Yanagita, L. Yang, T. Yoshida, S. Yoshiike, T. Yoshikoshi, M. Zacharias, G. Zaharijas, L. Zampieri, F. Zand anel, R. Zanin, M. Zavrtanik, D. Zavrtanik, A. A. Zdziarski, A. Zech, H. Zechlin, V. I. Zhdanov, A. Ziegler, & J. Zorn, 231–272

Zel'dovich, Y. B. 1964, Soviet Physics Doklady, 9, 195

Zhdankin, V., Uzdensky, D. A., Werner, G. R., & Begelman, M. C. 2018, MNRAS, 474, 2514

Zheng, Y. G., Kang, S. J., & Li, J. 2014, MNRAS, 442, 3166

Zhou, Y., & Matthaeus, W. H. 1990, J. Geophys. Res., 95, 14881

Zorn, J., White, R., Watson, J. J., et al. 2018, Nuclear Instruments and Methods in Physics Research A, 904, 44

Zweibel, E. G., & Yamada, M. 2009, ARA&A, 47, 291