Felix Aharonian
Lars Bergström
Charles Dermer

Astrophysics at Very High Energies

Saas-Fee Advanced Course 40

Felix Aharonian · Lars Bergström
Charles Dermer

Astrophysics at Very High Energies

Saas-Fee Advanced Course 40

Swiss Society for Astrophysics and Astronomy
Edited by R. Walter and M. Türler

Springer

Felix Aharonian
DIAS Dublin and MPIK Heidelberg
Dublin
Ireland

and

Max Planck Institute für Kernphysik
Heidelberg
Germany

Lars Bergström
Department of Physics
Stockholm University
Stockholm
Sweden

Charles Dermer
Space Science Division
U.S. Naval Research Laboratory
Washington
USA

Volume Editors
Roland Walter
Marc Türler
ISDC, Geneva Observatory
University of Geneva, Versoix
Switzerland

Cover Illustration: A composite image of the Crab Nebula showing the X-ray (blue), and optical (red) images superimposed. The X-ray emitting region is smaller because the higher energy X-ray emitting electrons radiate away their energy more quickly than the lower energy optically emitting electrons. The inner ring is about one light year across (Credit: X-ray: NASA/CXC/ASU/J. Hester et al.; Optical: NASA/HST/ASU/J. Hester et al.). The inset features the gamma-ray variability of the central part of the pulsar wind nebula, as observed by the Fermi/LAT, indicating that electrons are accelerated up to PeV energies, the highest energy electrons ever measured (Credit: ISDC/M. Balbo).

Preface

The 40th "Saas-Fee Advanced Course" of the Swiss Society for Astrophysics and Astronomy devoted to Astrophysics at Very-High Energies was held from March 14 to 20, 2010 in Les Diablerets, in the Swiss Alps. It gathered 105 participants and included a Fermi hands-on tutorial and an INTEGRAL data analysis tutorial.

The course was organized as 28 lectures reviewing the state of knowledge, open questions, and forecasts in the field of high and very-high energy gamma-ray astrophysics, a field that has encountered a revolution in the last years with the success of Cherenkov astronomy and the launch of the Fermi mission. Impact of gamma-ray observations on our knowledge of particle acceleration in galactic and extragalactic sources were reviewed as well as the prospects for dark matter detection and advances from the multi-messenger approaches.

The lectures were given by three world experts in the field:

Prof. Felix A. Aharonian is Professor of Astrophysics at the Dublin Institute for Advanced Studies (Ireland) and the Head of the High Energy Astrophysics Theory Group at the Max-Planck-Institut für Kernphysik in Heidelberg (Germany). His main scientific interests are the processes in thermal and non-thermal relativistic plasmas, physics and astrophysics of relativistic winds and jets, non-thermal processes in large-scale AGN Jets and in clusters of galaxies, the origin of galactic and extra-galactic cosmic rays, the diffuse extragalactic background radiation, observational gamma-ray cosmology, and the imaging atmospheric Cherenkov array technique. Prof. Aharonian is involved in a number of major high energy experiments, in particular as a member of the Science Working Group of ASTRO-H, of the Collaboration Board of H.E.S.S. and of the Consortium Board of KM3NeT.

Prof. Lars Bergström is Professor at the University of Stockolm (Sweden) and the Head of the Cosmology, Particle Astrophysics and String Theory Group. He is also Director of the Oskar Klein Centre for Cosmoparticle Physics. One of his main lines of research is the investigation of the nature of dark matter, in particular supersymmetric and Kaluza-Klein particles and the prediction of indirect detection rates of various dark matter particle candidates. His group is active in many aspects of observational and theoretical supernova cosmology, gravitational lensing, determination of cosmological parameters, models for dark matter, and

string cosmology. He is collaborating in various experiments to search for evidences for dark matter, in particular Fermi, IceCube and PAMELA.

Dr. Charles D. Dermer is the Head of the Space Radiations Section in the Space Science Division of the Naval Research Laboratory in Washington, DC. His interests cover many areas of astrophysics, including cosmic rays, the multi-wavelength astronomy of blazars, the physics of neutron stars and black holes, gamma-ray bursts, merging clusters of galaxies, and solar flares. He uses theoretical modeling, supported by numerical simulations of the basic physical processes involving high-energy interactions between particles and photons in magnetized plasma, to identify the nature of high-energy astronomical sources and the physical mechanisms responsible for the observed gamma ray and particle emissions. Dr. Dermer was one of four GLAST Interdisciplinary scientists, and is currently a full Fermi Collaboration Member. He has served on numerous review and study panels, including the Advanced Compton Telescope Working Group and the VERITAS External Oversight Committee.

This volume of the Saas-Fee lecture notes provides a broad overview of astrophysics at high and very-energy energies, as well as an introduction to multi-messenger astronomy and the possible nature of dark matter. Prof. Felix Aharonian presents the breakthrough in very-high energy gamma-rays achieved by the current generation of Cherenkov telescopes. He describes the main results and their implications for theoretical models of the TeV gamma-ray emission with a focus on Galactic sources. Dr. Charles Dermer follows a similar approach for the other breakthrough in high-energy observations achieved by the Fermi gamma-ray space telescope in the spectral window of GeV gamma-rays. In this second part, emphasis is given on the physics at play in blazars—the most extremely luminous and variable active galactic nuclei—as constrained by the unprecedented gamma-ray observations by Fermi. The third contribution to this book is of a quite different nature. Prof. Lars Bergström gives us a broad overview of multi-messenger astronomy and the quest of identifying the nature of dark matter both theoretically and experimentally. He perfectly succeeded in making the challenges of current astroparticle physics and theoretical cosmology understandable to astronomers.

We are very grateful to the lecturers for their enthusiasm in communicating their deep knowledge, their brilliant lectures, as well as for writing the rich manuscripts composing this book. We extend our warmest thanks to the course secretary, Martine Logossou, for her effective administration of registrations, of the budget, and her organizational help during the course. We acknowledge the design of the course poster by Jean-Christophe Leyder. We also would like to thank all speakers of the INTEGRAL tutorial session held on Thursday afternoon. In particular, Peter Kretschmar from ESA and our colleagues from the ISDC Data Centre for Astrophysics: Enrico Bozzo, Carlo Ferrigno, Lucia Pavan, Nicolas Produit, Claudio Ricci, and Reiner Rohlfs. Last but not least, we thank Elizabeth Hays and Elizabeth Ferrara from the NASA Fermi Science Support Center and

Andrea Tramacere from the ISDC to offer the course participants the opportunity of a hands-on session on Fermi data analysis.

One of the highlights of the course was the concert "Il Viaggio d'Amore", a love journey from the Renaissance to nowadays, offered by Arianna Savall Figueras and Petter Udland Johansen. Many participants made a memorable walk with torches from Les Diablerets to the little church of Vers l'Eglise where the concert took place. It was a magical evening and we would like to thank again the two outstanding performers for their delighting music.

The Eurotel-Victoria provided—as so often in the past—a pleasant environment for the Saas-Fee Course and a generous banquet dinner. The organization of this course would not have been possible without the financial support of the Swiss Society for Astrophysics and Astronomy (SSAA), the Swiss Institute of Particle Physics (CHIPP), and the Swiss Academy of Sciences (SCNAT). We are very grateful to these organizations for their contribution, which allowed the participants to attend a very diverse, interesting, and successful 40th Saas-Fee Course.

Versoix, October 2012   Roland Walter
                                                                                               Marc Türler

Contents

Sources of GeV Photons and the Fermi Results
Charles D. Dermer

Gamma Rays at Very High Energies

Felix Aharonian

1 Introduction

Cosmic gamma rays carry key information about high energy phenomena in a large variety of astrophysical environments. Being a part of *modern astrophysics* and *astroparticle physics*, gamma-ray astronomy is a discipline in its own right [52]. It addresses an impressively broad range of topics related to the non thermal processes in the Universe, including acceleration, propagation, and radiation of relativistic particles on all astronomical scales: from compact objects like pulsars (neutron-stars) and microquasars (accreting stellar mass black holes) to giant jets and lobes of radio-galaxies and galaxy clusters.

The gamma-ray phenomena generally proceed under extreme physical conditions in environments characterized with huge gravitational, magnetic and electric fields, relativistic bulk motions and shock waves, highly excited (turbulent) plasma, etc. Consequently, any coherent description and interpretation of phenomena related to gamma-rays requires deep knowledge of many disciplines of physics like nuclear and particle physics, quantum and classical electrodynamics, special and general relativity, plasma physics, magnetohydrodynamics, etc.

The energy range covered by gamma ray astronomy spans from 0.1 MeV to 100 EeV (throughout these lectures I will use the energy units which are common in high energy physics and astrophysics: $1\,\mathrm{keV} = 10^3\,\mathrm{eV}$, $1\,\mathrm{MeV} = 10^6\,\mathrm{eV}$, $1\,\mathrm{GeV} = 10^9\,\mathrm{eV}$, $1\,\mathrm{TeV} = 10^{12}\,\mathrm{eV}$, $1\,\mathrm{PeV} = 10^{15}\,\mathrm{eV}$, $1\,\mathrm{EeV} = 10^{18}\,\mathrm{eV}$). While the lower bound associates with the region of nuclear gamma-ray lines, the upper bound is determined by the highest energy particles observed in cosmic rays. Because of the

F. Aharonian (✉)
Dublin Institute for Advanced Studies, 31 Fitzwilliam Place, Dublin 2, Ireland
e-mail: felix.aharonian@dias.ie

F. Aharonian
Max Planck Institute für Kernphysik, Saupfercheckweg 31, 69117 Heidelberg, Germany
e-mail: felix.aharonian@mpi-hd.mpg.de

essentially different detection methods and approaches applicable to different energy bands, currently this enormous energy domain of cosmic electromagnetic radiation is covered inhomogeneously. In particular, so far cosmic gamma-rays are detected in 'low' (LE or MeV), 'high' (HE or GeV) and very high (VHE or TeV) energy bands.

1.1 Status of Observational Gamma Ray Astronomy

(i) *Low Energy Band: 0.1–100 MeV*

This energy interval is uniquely linked to several astrophysical phenomena, in particular to the nucleosynthesis of heavy elements related to the type Ia supernovae (SNIa), Gamma Ray Bursts, Solar flares, interactions of sub-relativistic cosmic rays with the interstellar gas and dust, production and annihilation of positrons, etc. While many aspects of these phenomena can be best probed with low energy gamma-rays, the MeV gamma-ray sky remains an almost unexplored frontier. The main challenge of low energy gamma-ray astronomy is the design and construction of detectors with a sensitivity compatible to the conservative flux predictions. Unfortunately, the combination of several principal factors—the low detection efficiency, the modest angular resolution and the high level of backgrounds of different origin—severely limit the potential of detectors operating in this energy region.

The minimum detectable energy fluxes, even after significant improvements as foreseen for the next generation of low-energy gamma-ray detectors, will still remain modest, hardly better than 10^{-12} erg/cm^2 s. Even so, low-energy gamma-rays are messengers of crucial astronomical information that cannot be obtained by other means. This concerns, for example the probes of the flux of sub-relativistic ($E \leq 100$ MeV) cosmic rays in the Interstellar Medium (ISM) through the prompt *de-excitation* gamma-ray lines (see e.g. Ref. [178]). On the other hand, gamma-ray continuum at MeV energies produced via *bremsstrahlung* of electrons (positrons), as well as at the *positron annihilation in flight* (in sources with high positrons-to-electron ratio) contains unique information about relativistic electrons below 100 MeV [58]. In the environments with low magnetic field, in particular in the ISM ($B \leq 10\,\mu$G), such an information is not accessible via synchrotron radio emission because it appears at non-visible frequencies below 1 MHz. The information about low-energy electrons and protons (nuclei) is important for the understanding of the energy balance between different forms of matter, magnetic fields and cosmic rays. In the galactic disk, such measurements have some other astrophysical implications, for example they provide direct estimates of the ionization and heating rates of the interstellar gas by low-energy cosmic-ray protons and electrons.

Another important implication of MeV gamma-ray emission is related to studies of mildly relativistic *thermal* plasmas formed in the vicinity of compact relativistic objects like neutron stars and black holes. The detection of characteristic MeV radiation dominated by the Comptonized free-free (bremsstrahlung) emission and electron–positron annihilation, gives direct information on electrons in such extreme thermal plasmas (see e.g. [223]). However, in most cases the fast radiative

cooling of electrons prevents their heating to temperatures beyond 10^9 K, thus the electron cooling proceeds through radiation in hard X-rays. Consequently, the proton temperature can significantly exceed the electron temperature. The formation of hot two-temperature plasmas, $T_i \gg T_e$, in strong shock waves or in accretion flows close to black holes, can be studied by detection of characteristic gamma-ray line emission produced through the chain of spallation and excitation reactions. A clear signature of radiation at the final stage of such hot plasmas when all nuclei are destroyed and the nucleonic component of plasma basically consists of protons and neutrons (with a small fraction of deuterium in equilibrium), is the continuum due to the proton-neutron bremsstrahlung and the broadened and "blue-shifted" deuterium line [69]. In accreting solar-mass black holes, this radiation is released, depending on T_i, typically between 1 and 30 MeV.

Finally, one of the major objectives of MeV gamma-ray astronomy remains the exploration of 0.511 MeV line emission due to annihilation of the positrons which are copiously produced in various astrophysical environments (see e.g. the recent review [210] on the annihilation line from the Galactic Center).

(ii) *High Energy Band: 0.1–100 GeV*

Before the launch of the *Fermi Gamma-ray Space Telescope* (formerly GLAST) in May 2008, the high energy space-based gamma-ray astronomy has been dominated by the results obtained with the EGRET telescope aboard Compton Gamma Ray Observatory. Because of the rather modest angular resolution of EGRET (of order of a few degrees), only two source populations—the active galactic nuclei and pulsars— have been clearly identified as high energy gamma-ray emitters. With the *Fermi* LAT (Large Area Telescope) the HE gamma-ray astronomy entered a new era. This instrument with significantly improved (compared to EGRET) angular resolution (0.6° at 1 GeV and better than 0.15° at energies above 10 GeV) and flux sensitivity (better than 10^{-12} erg/cm^2 s) [86], is a perfectly designed tool for deep gamma-ray surveys with an effective field of view of order of 2 steradian. Over the last three years, *Fermi* LAT has been releasing vast amount of important astronomical information. The high energy gamma-ray sky revealed by *Fermi* (see Fig. 1) is really very impressive! These results confirm, to a large extent, the optimistic pre-launch expectations concerning, in particular, the dramatic increase of the number of gamma-ray emitting pulsars and AGN, discovery of new classes of *compact/variable* and *extended* galactic and extragalactic gamma-ray sources, the detection of multi-GeV components of GRBs, etc. The second *Fermi* LAT gamma-ray source catalogue [200], based on the first two years of observations, consists of almost 2000 galactic and extragalactic gamma-ray emitters. While more than half of these objects are associated with counterparts representing known source populations (more than one hundred sources being firmly identified), the origin of approximately 1/3 of these objects remains an open issue. This concerns, first of all, the extended sources located in the galactic plane, e.g. SNRs and PWNe, for which the chance of confusion with the diffuse emission of the galactic disk is especially high. Because of the limited angular resolution, the most reliable approach for identification of GeV gamma-ray sources is the analysis based on temporal studies. In this regard it is quite natural that the best "astronomical
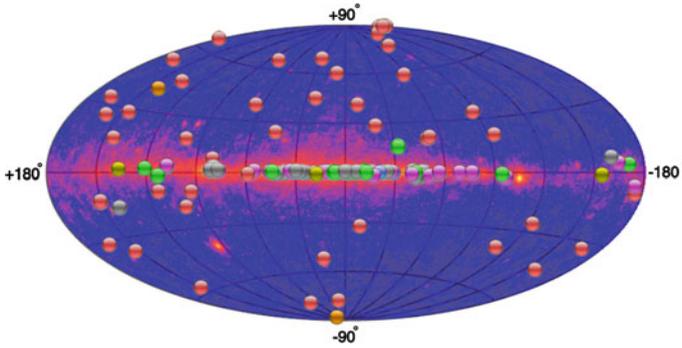
Fig. 1 The overall MeV/GeV gamma-ray sky (the *blue* to *red color* background) as seen by *Fermi* [200] and the positions of discrete TeV gamma-ray sources detected with ground-based instruments; the regularly updated version of this figure can be found on the "TeVCat" webpage: http://tevcat. uchicago.edu. The SNRs (in association with GMCs or without) are shown with a *green* symbol, the pulsar wind nebulae are shown in *violet*, the binary systems—in *yellow*, starburst galaxies—in *brown*, AGN (all types)—in *red*. The so-called "dark" sources without any reliable association with the well known objects are shown with a *grey* symbol

clocks"—the Pulsars—constitute the largest population of identified galactic GeV gamma-ray sources. The periodic character of gamma-ray emission of the galactic binary systems or the sporadic flares of AGN provide another tool for identification of variable gamma-ray sources based on simultaneous observations in different energy bands. In general, the multi-wavelength observations is a key component for identification of gamma-ray emitters, as well as for deeper understanding of the nature of these objects.

The *Fermi* observations significantly enhance our knowledge about the diffuse gamma-ray backgrounds of different origin. In particular, *Fermi* has extended the range of observations of the diffuse emission of the Galactic Disk and the isotropic (extragalactic) gamma-ray background to several hundreds of GeV and helped to clarify some controversial issues related to the contributions of different source populations and the relevant radiation mechanism.

A number of important results, especially at energies below 1 GeV, have been recently reported also by the Italian gamma-ray satellite AGILE [207]. In general, the observations by *Fermi* and *AGILE* support many phenomenological concepts and theoretical models in different areas of astrophysics. At the same time, these observations resulted in a number of "unplanned" discoveries and revealed some puzzling phenomena like flares of the Crab Nebula or existence of multi-kpc scale non-thermal structures—giant reservoirs of relativistic particles centered on the core of the Galaxy ("Fermi bubbles").

Concerning the next generation space-based gamma-ray detectors, it is likely that in the foreseeable future one cannot expect significant developments beyond the level achieved by *Fermi* LAT, except perhaps for the energy band below 1 GeV down to several tens of MeV. The design and construction of a space-based instrument in

this energy interval with an affordable effective detection area of order of $1\,m^2$ and angular resolution of about $1°$, would increase significantly the detection rate (photon statistics) and improve the flux sensitivity of *Fermi* LAT below 1 GeV potentially by an order of magnitude. The optimization of the pair-conversion tracking detection technique with a focus on energies around 100 MeV would be an attractive and promising strategy given the number of outstanding astrophysical questions (relevant to nearly all source populations) not fully addressed by *Fermi LAT*.

(iii) *Very High Energy Band: 0.1–100 TeV*

One of the most remarkable achievements of recent years in astrophysics was the sudden emergence of very-high-energy gamma-ray astronomy as a truly astronomical discipline. The observations conducted by HESS, MAGIC, VERITAS (see Figs. 4, 5, 6) and MILAGRO groups resulted in the discovery of many sources with a number in excess of 130 (see Fig. 1). These sources represent almost all major *non-thermal* astrophysical source populations, including shell type Supernova Remnants, Pulsar Wind Nebulae, Star Forming Regions, Giant Molecular Clouds, X-ray Binary Systems, Blazars, Radio-galaxies, Starburst Galaxies (for a review see e.g. [27, 154]). In general, this success was a big surprise, especially given the rather difficult past and the controversial history of the field over the last four decades (see e.g. [52, 248]). In this regard, a question naturally arises concerning the reasons which made possible this success. A likely answer to this question perhaps can be formulated as a fortunate combination of two independent factors:

(a) the practical realization of the great potential of stereoscopic arrays of Imaging Atmospheric Cherenkov Telescopes (IACTs) as effective multifunctional tools for spectral, temporal, and morphological studies of VHE gamma-ray sources;
(b) the existence of a large variety of perfectly designed machines—*TeVatrons*, *PeVatrons*, and *EeVatrons*—"factories" of relativistic matter where the effective particle acceleration is accompanied by effective radiation processes.

The discussion of the nature of VHE gamma-ray source populations constitutes a major purpose of this chapter.

1.2 Links to Other Disciplines

Gamma-ray astronomy has deep intrinsic links to other astronomical disciplines and cosmology. For example, relativistic electrons responsible for radio emission of the Galactic Disk, produce, at interactions with ambient gas and radiation fields, diffuse gamma-ray emission in the MeV and GeV energy bands. At higher (TeV) energies of electrons, the synchrotron radiation appears in the X-ray energy band, while the inverse Compton (IC) scattering of same electrons results in high energy gamma-rays in the TeV energy domain. The link between synchrotron X-rays and very high energy IC gamma-rays is common for source population of quite different origin, in particular for Supernova Remnants, Pulsar Wind Nebulae, X-ray binaries, Blazars,

etc. Another example is the interactions of cosmic ray (CR) protons with matter and radiation that produce π^0 mesons, the decay of which results in the so-called hadronic component of gamma-rays. Neutrinos from decays of $\pi^{\pm}$ mesons produced in same interactions, constitute the basis of high energy neutrino astronomy. The common origin of high energy neutrinos and hadronic gamma-rays as secondary products of interactions of cosmic rays is an indication of the important role of gamma-ray astronomy for realization of the so-called *multi-messenger* approach in the solution of the problem of origin of galactic and extragalactic cosmic rays—one of the key objectives of the new interdisciplinary research area called Astroparticle Physics.

The topics related to gamma-ray astronomy cannot be reduced to merely non-thermal phenomena. For example diffuse galactic and extragalactic infrared/optical radiation components have thermal origin and formally do not associate with high energy processes. On the other hand, they play an important role in production and absorption of high energy gamma-rays through the inverse Compton scattering and photon–photon pair production processes. This determines the deep links of gamma-ray astronomy to infrared and optical astronomies, as well as to cosmology. For example, the characteristic absorption features in the spectra of high energy gamma-rays arriving from distant extragalactic objects caused by interactions with the Extragalactic Background Light (EBL) contain unique cosmological information about the epochs of formation of galaxies and their evolution in the past. The indirect search for non-baryonic Dark Matter through high energy gamma-ray, as well as detection of gamma-radiation related to large scale cosmological structures (clusters of galaxies) are two other important (not yet realized) issues at the interface of gamma-ray astronomy and cosmology.

The above examples demonstrate the distinct feature of high energy gamma-ray astronomy as a *multi-disciplinary* research area and a key component of astroparticle physics in the context of *multi-wavelength* and *multi-messenger* approach in the studies of the most energetic processes in the Universe.

The links of the TeV domain to its closest neighbour—the GeV domain—are expected to be especially tight. It is generally believed that the TeV and GeV gamma-ray fluxes should strongly correlate, therefore the TeV gamma-ray sources should be also prominent GeV sources. However the GeV–TeV links are not so straightforward as it may look at first glance.

On the GeV–TeV Links

Over the last several years the number of VHE sources have been dramatically increased; presently (first half of 2012) it exceeds 130. At first glance, it looks a modest achievement compared to the almost 2000 sources detected by *Fermi* LAT. However, when one takes into account the limited observation time (the duty factor of observations with Cherenkov telescopes does not exceed 10 %) and the small field of view (less than 0.01 steradian) of IACT arrays, one may conclude that there should be a plenty of VHE gamma-ray emitters to be discovered by next-generation instruments.

It is expected that the next generation major ground based detector, the Cherenkov Telescope Array(CTA), with an order-of-magnitude improved sensitivity (see Sect. 2), will increase dramatically, by one or perhaps even two orders of magnitude, the number of VHE gamma-ray sources. However, the predictions on the number of VHE sources based on extrapolations from observations at other wavelengths, should be taken with a caution. Indeed, as long as we deal with poorly understood phenomena in a new energy band, often seemingly reasonable extrapolations could appear wrong. Moreover, since in sources of different power the acceleration, radiation and absorption processes can proceed with different efficiencies, the predictions based on the so-called 'LogN–LogS' relations, could lead to misleading conclusions even in the case of well established source populations. This rather general statement concerns all wavelengths of non thermal electromagnetic radiation, including the two closest neighbours—the GeV and TeV bands. For example, the increase of the overall luminosity of a compact object could reduce the maximum achievable energy of electrons (because of the enhanced inverse Compton cooling), and, at the same time, increase the probability for absorption of VHE gamma-rays inside the source. This kind of non-linear effects should have a strong impact on gamma-ray fluxes, but act in different energy bands quite differently. Consequently one should expect essentially different 'LogN–LogS' relations applied to GeV and TeV energies.

For copious gamma-ray production two conditions are required—an effective particle *accelerator* and a surrounding dense *target* in the form of gas, radiation or magnetic field. In the case of energy-dependent escape of particles (protons or electrons) from the accelerator, the resulting spectrum of particles inside the 'old' accelerator can be significantly *softer* than the spectrum of escaped particles outside the accelerator. Correspondingly, the gamma-ray spectrum from the accelerator will be softer compared to the energy spectrum of gamma-ray produced outside the accelerator. Since the diffusion of low energy protons is significantly slower, the impact of the escape is less critical for GeV gamma-rays. Thus an observer detecting GeV, but not TeV gamma-ray, from an accelerator at late epochs of its evolution, might conclude that we see an *active* source which however does not accelerate particles beyond TeV energies. But in reality, the accelerator could be a *dead* TeVatron or PeVatron. The ambiguity can be resolved through comparison of gamma-ray spectra detected from inside and outside the accelerator.

Because of propagation effects, the hadronic gamma-ray sources are expected to be more extended at TeV than at GeV energies. For example, the gamma-ray sources can be due to protons which left the accelerator and interact with nearby dense molecular clouds. In such cases gamma-rays can be produced predominantly from the clouds but not from the accelerator itself; the latter could be not active anymore, or the gas density inside the source cannot be sufficient for production of detectable gamma-ray fluxes. This might be a quite natural explanation of the so-called 'dark accelerators'—VHE gamma-ray sources from regions without counterparts observed at other wavelengths. For electrons, which suffer significant radiative (synchrotron and inverse Compton) losses on time-scales shorter than the escape time, just an opposite picture is expected—strong VHE gamma-ray emission of IC origin from a

compact region which coincides with the accelerator, and more extended emission at lower energies from run-away electrons.

Finally, the energy dependent absorption of gamma-rays due to photon–photon collisions inside the compact objects can significantly change the original (production) spectrum of gamma-rays. The impact of absorption can be different for different energy intervals of gamma-rays depending on the spectral energy distribution (SED) of the surrounding target photon gas. For example, in compact X-ray regions, e.g. in accretion flows close to black holes, the absorption effect is strongest at GeV energies, while in binary systems containing luminous optical stars, TeV gamma-rays suffer most severe absorption.

In summary, depending on specifics of acceleration and propagation of charged particles (protons and electrons), as well as depending on the combination of gamma-ray production/absorption mechanisms, one can expect quite different relations between the GeV and TeV fluxes. They can correlate, anti-correlate, or behave in a random fashion. Thus, the energy spectra of GeV gamma-ray sources should not necessarily extend to TeV energies, while the TeV gamma-ray sources could appear without a GeV counterpart. Therefore the GeV and TeV gamma-ray sources can be represented quite differently in the given source population. For example, while the *Fermi* observations have proved that a large fraction of pulsars are prolific MeV/GeV gamma-ray emitters, the ground-based observations show that Pulsar Wind Nebulae radiate most effectively in the TeV gamma-ray band. On the other hand it is remarkable that both GeV and TeV gamma-rays are detected from sources representing almost all 'suspected' non thermal source populations of galactic and extragalactic origin, in particular from Supernova Remnants (SNRs), Pulsar Wind Nebulae (PWNe), Giant MolecularClouds (GMCs), Compact Binaries (CBs), Starburst Galaxies, Radio Galaxies and Blazars. High energy gamma-rays are detected also from the Sun and Moon, as well as from Gamma Ray Bursts.

Figure 2 shows the populations of sources established as GeV and TeV gamma-ray emitters, and demonstrates the complex links related to several major research areas of astrophysics and astroparticle physics: Origin of Galactic and Extragalactic Cosmic Rays, Physics and Astrophysics of Compact Objects (Black Holes and Neutron stars), Relativistic Outflows (AGN jets and Pulsar Winds), Cosmology (Dark Matter, Extragalactic radiation and magnetic fields), etc.

2 Astrophysical Potential of Ground-Based Detectors

Earth's atmosphere is not transparent to gamma-rays of any energy. Therefore, their registration requires detectors installed on space platforms. However, the satellites cannot offer, at least in the foreseeable future, detection areas significantly exceeding $1\,\mathrm{m}^2$; this constrains the *effective* studies of tiny fluxes of cosmic gamma-rays to energies $\leq 100\,\mathrm{GeV}$. Fortunately, at higher energies an alternative method can be used for detection of gamma-rays. The method is based on the registration of atmospheric
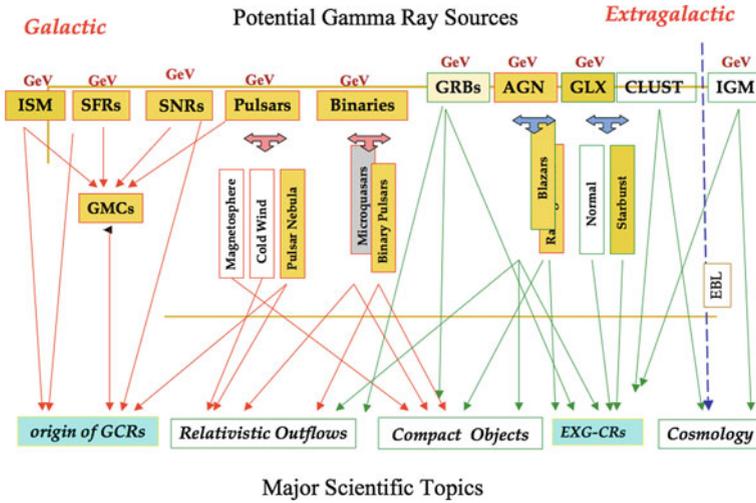
Fig. 2 GeV and TeV source populations and the links between the major scientific topics

showers (initiated by interactions of gamma-ray) either directly or through their Cherenkov radiation.

The faint and brief Cherenkov signal which lasts only several nanoseconds can be detected by large optical reflectors equipped with fast multi pixel cameras. With a telescope consisting of an optical reflector of diameter $D \approx 10$ m, and a multichannel camera with pixel size $0.1°$–$0.2°$ and field-of-view $\Theta \geq 3°$, primary gamma-rays of energy ≥ 100 GeV can be collected from distances as large as 100 m. This provides huge detection areas, $A \geq 3 \times 10^4 \, \text{m}^2$, which largely compensate the weak gamma-ray fluxes at these energies. The total number of photons in the registered Cherenkov light image is a measure of energy, the orientation of the image correlates with the arrival direction of the gamma-ray, and the shape of the image contains information about the origin of the primary particle (a proton or photon). The basic principles of operation of the IACT technique is illustrated in Fig. 3.

The stereoscopic observations of air showers with two or more 10 m diameter telescopes located at distances of about 100 m from each other provide a quite low energy threshold around 100 GeV, effective (by a factor of 100) rejection of hadronic showers, and good angular ($\approx 0.1°$) and energy ($\approx 15\,\%$) resolutions (see e.g. [53]). At energies around 1 TeV, this results in a minimum detectable energy flux of 10^{-13} erg/cm^2 (see Fig. 7), a quite impressive sensitivity even in the standards of advanced branches of astrophysics. In particular, it is much better than in any other gamma-ray domain, including the GeV energy band, where the sensitivity of *Fermi* LAT, even after dramatic improvement compared to the performance of the previous gamma-ray space-borne instruments, still cannot compete with the performance already achieved in the TeV energy band. Thanks to very large collection area, the IACT technique provides large gamma-ray photon statistics even from
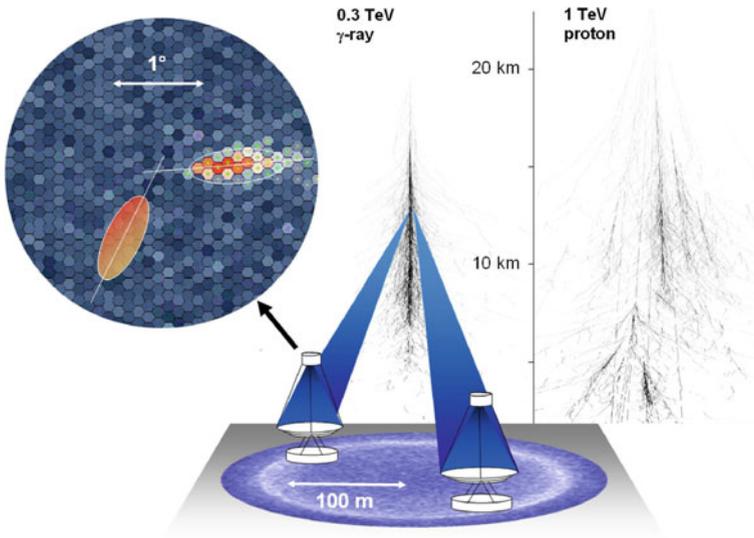
Fig. 3 The operation of the imaging atmospheric Cherenkov telescope technique (from Ref. [154])

Fig. 4 The HESS system of four 13 m diameter Cherenkov telescopes. While this system has been operating in Namibia since 2003, the new 28 m diameter telescope is under construction (the picture in the center of the original HESS array is a photo montage)

relatively modest TeV gamma-ray emitters. Coupled with good energy and angular resolutions, the rich photon statistics allows deep morphological, spectral, and temporal studies. This makes the IACT arrays perfect multifunctional and multipurpose astronomical tools for exploration of a broad range of non-thermal objects and phenomena, both of galactic and extragalactic origin. Currently, three major IACT arrays (see Figs. 4, 5, 6) HESS (High Energy Stereoscopic System), MAGIC (Major Atmospheric Imaging Cherenkov) and VERITAS (Very Energetic Radiation Imaging Telescope Array System)—located both in the northern (MAGIC, VERITAS) and southern (HESS) hemispheres, determine the status of VHE gamma-ray

Fig. 5 The system of two 17 m diameter MAGIC gamma-ray telescopes at the Roque de los Muchachos observatory (La Palma, Canary Islands)

Fig. 6 The VERITAS system of four 12 m diameter Cherenkov telescopes located in southern Arizona

astronomy. What concerns the previous generation instruments, one should mention, amongst others, the 10 m diameter single dish of the Whipple Observatory (south Arizona) and the HEGRA array of five relatively modest (4 m diameter) Cherenkov telescopes (La Palma, Canary Islands). These instruments, which can be considered as prototypes of the current IACT arrays, played a crucial role in the development of ground-based gamma-ray astronomy. While the Whipple collaboration pioneered the implementation and successful realization of the *imaging* atmospheric Cherenkov technique, the HEGRA collaboration convincingly demonstrated the power of the *stereoscope* approach. In this regard, it is not a big surprise that the performance of the current (and the next) generation IACT arrays are not far from the early predictions based on the extrapolation of the performance of the single Whipple dish and the HEGRA telescope array (see e.g. [52]).
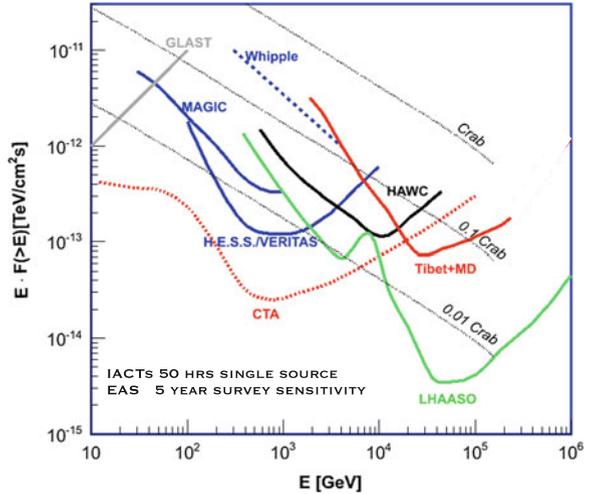
2.1 IACT Arrays

Figures 7, 8, 9, and 10 represent three characteristic examples demonstrating the
great performance of the stereoscopic IACT technique for morphological, spectral
and temporal studies, respectively.

(i) Figure 8 shows the VHE gamma-ray map of an extraordinary site—the central
several hundred parsec region of our Galaxy which harbours a variety of poten-
tial gamma-ray emitters. This region has been predicted as a possible gamma-ray
source also because of predicted sharp concentration of Dark Matter. Deep obser-
vations of HESS did reveal that this compact region (angular size less than $2°$) is
packed with several gamma-ray sources, including a point like source in the very
center of the Galactic Center, diffuse emission contributed by giant molecular
clouds, a composite supernova remnant, as well as an interesting but not yet
identified extended source.

(ii) Figure 9 demonstrates the power of the IACT technique for spectroscopic stud-
ies. The energy spectra of two active galactic nuclei, Mkn 421 and Mkn 501,
have been measured by the HEGRA IACT array. The spectra based on very
large gamma-ray statistics (60,000 and 40,000 from Mkn 421 and Mkn 501,
respectively), detected during high states of these objects, can be fitted with the
canonical "power-law with exponential cut-off" function. Note that the mea-
sured spectral points extend beyond $3E_0$; this is a remarkable result even for the
standards of laboratory experiments.

(iii) High detection rates of gamma-rays by IACT arrays are possible even for
relatively modest energy gamma-ray flux of about 10^{-11} erg/cm^2 s. This makes
these instruments powerful tools for temporal studies of highly variable VHE
sources. This is demonstrated in Fig. 10 for a major flare of the blazer
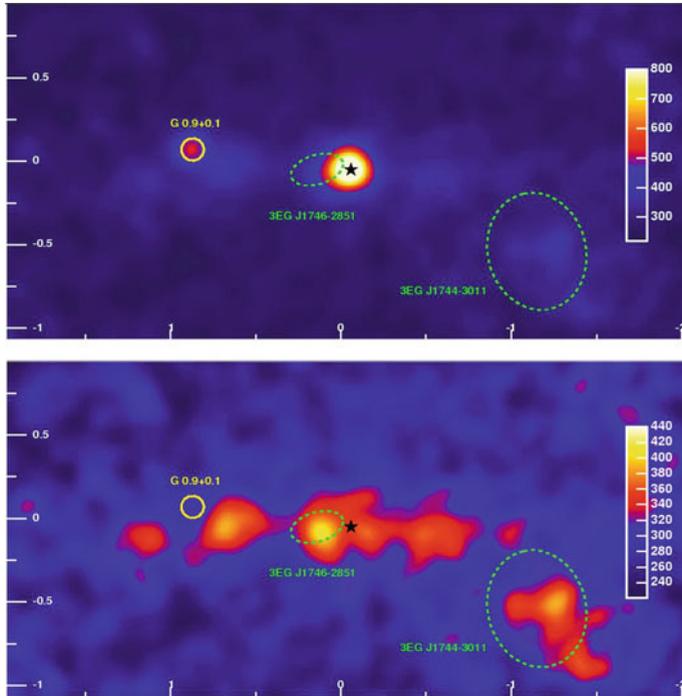
Fig. 8 The image of the several-hundred parsec region of the Galactic Center in TeV gamma-rays. It contains a point like source (angular radius less than a few arcminutes) the gravity center of which coincides with an accuracy of 13 arcseconds with the compact radio source Sgr A*—a super massive *black hole* at the dynamical center of the Milky Way [21, 32]. The second point like source located about one degree away positionally coincides with the composite supernova remnants G09+0.1 [37]. A prominent feature of this region is the ridge of diffuse emission tracing several well identified giant molecular clouds [39]. This complex region contains some other, not yet firmly identified, "hot spots"

PKS 2155-304 on the night of July 29–30, 2006. The outburst was so powerful that the detection rate of VHE gamma-rays by the HESS telescopes "jumped" a level of several (background-free) events per second leading to the discovery of variability of the source on an exceptionally short time-scale of about 2–3 arcmin.

2.2 Potential of EAS Arrays

The IACT arrays are designed for observations of point-like or moderately extended (with angular size 1° or less) objects with known celestial coordinates. However, the high sensitivity and relatively large (≥4°) field of view of IACT arrays allow effective all-sky surveys as demonstrated by the HESS collaboration. On the other hand, the potential of IACT arrays is limited for the search of very extended structures (like
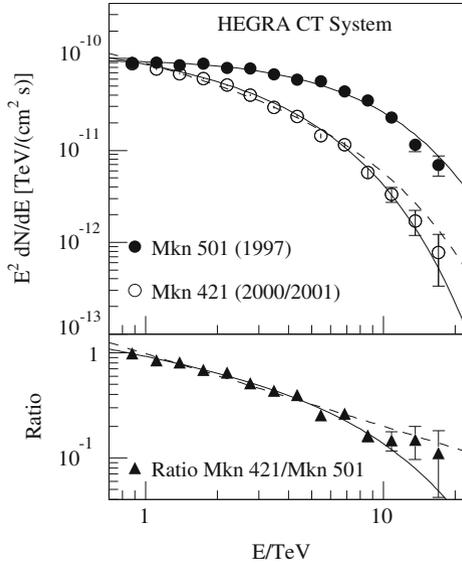
Fig. 9 The energy spectra of active galactic nuclei Mkn 501 and Mkn 421 measured by the HEGRA stereoscopic system of Cherenkov telescopes in the high states (from Ref. [30]). The spectra are well described by "power-law with exponential cut-off", $E^{-\Gamma} \exp(-E/E_0)$, with $\Gamma = 1.92$ and $E_0 = 6.2$ TeV for Mkn 501, and $\Gamma = 2.19$ and $E_0 = 3.6$ TeV for Mkn 421. To demonstrate the difference in two energy spectra and to reduce the impact of possible systematic effects, the ratio of the Mkn 421 and Mkn 501 spectra is shown in the *lower panel*
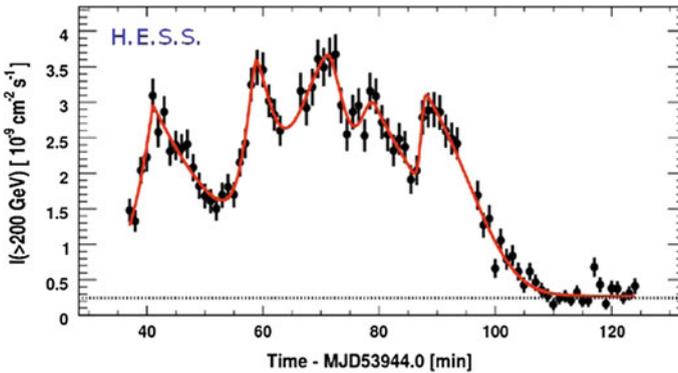
Fig. 10 The *light curve* of the exceptional outburst of the blazer PKS 2155-304 on the night of July 29–30 2006. More than 10,000 gamma-rays have been detected during 90 min leading to extraction of sharp flares on minute time-scales

diffuse emission of the galactic disk), as well as for the solitary events like GRBs. In this regard, the detection technique based on direct registration of particles that comprise the extensive air showers (EAS), is a complementary approach to the IACT technique.

The traditional EAS technique, based on scintillators or water Cherenkov detectors spread over large areas, works quite effectively for detection of cosmic rays at ultra-high energies, $E \geq 100$ TeV. In order to make this technique more adequate to purposes of gamma-ray astronomy, the detection energy threshold should be reduced by two orders of magnitude. This can be achieved using dense particle arrays located on very high altitudes. The feasibility of both approaches recently have been successfully demonstrated by the ARGO and Milagro groups. The significance map of the galactic plane region $l \in [30°, 220°]$ and $b \in [-10°, 10°]$ obtained with the Milagro detector [13] is shown in Fig. 11. Eight candidate sources at a median energy of $\sim$20 TeV have been found with pre-trial significance $\geq 4.5\sigma$. After accounting for the trials over ≈ 400 square degree region, four of these candidates survived as reliable detections with an after-trial statistical significance exceeding 4σ.

These results, as well as the prospects of continuous monitoring of a significant part of the sky, which might lead to exciting discoveries of yet unknown VHE transient phenomena in the Universe, justifies the new proposals of high altitude EAS detectors (see for a review [27]) like HAWK, a High Altitude Water Cherenkov Experiment under construction on a site close to Sierra Negra, Mexico. The 5 year survey sensitivity of HAWK at energies between 1 and 10 TeV is expected to be comparable to the sensitivity of *Fermi* around 1 GeV. In this regard *HAWK* will be complementary to *Fermi* for continuous monitoring of more than 1 steradian fraction of the sky at TeV energies. At higher energies, one should mention the ambitious LHAASO (Large High Altitude Air Shower Observatory) detector facility at Yangbajing, Tibet. This array consisting of several types of detectors of electromagnetic and muon components of air showers will cover huge area and achieve an impressive sensitivity at energies of several tens of TeV.

2.3 Future IACT Arrays

Planning of the next generation of Imaging Atmospheric Cherenkov Telescope (IACT) arrays has two major objectives: (i) an order of magnitude improvement of the flux sensitivity in the standard 0.1–10 TeV energy interval (TeV regime), and (ii) an aggressive expansion of the energy domain of IACT arrays in both directions—down to 10 GeV (multi-GeV regime) and well beyond 10 TeV (sub-PeV regime).

TeV Regime

The best performance the IACT technique is achieved in this energy regime, and still the potential is not saturated. The combination of three basic factors, (i) high
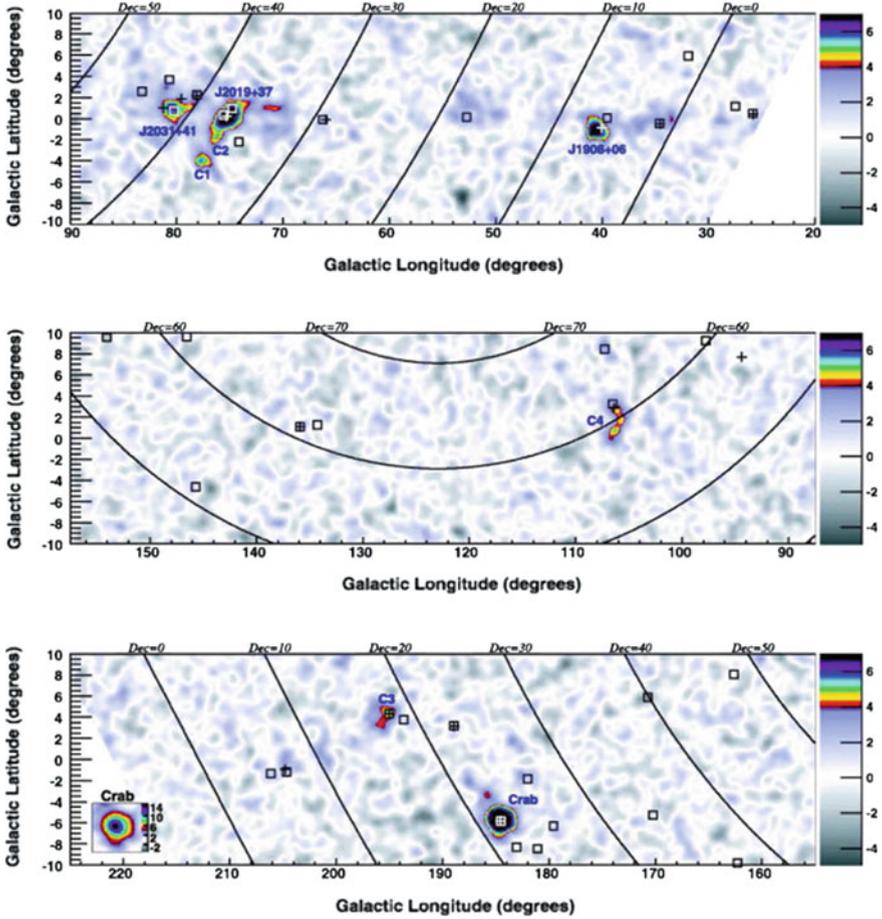
Fig. 11 Significance map of the Galactic plane at energies above 20 TeV produced on the basis of MILAGRO data [13]. The *color code* shows the pre-trials significance in this PSF-smoothed map. The maximum positive value of the color code saturates at 7σ, although three of the gamma-ray sources are detected with higher statistical significance

efficiency of detection/identification of electromagnetic showers, (ii) good accuracy of reconstruction of the direction and energy of primary gamma-ray, and (iii) large gamma-ray photon statistics, allows reduction of the minimum detectable energy flux to the level of 10^{-14} erg/cm^2 s, and improvement of the angular resolution to $\delta\theta \approx 2$–3 arcmin.

Such an impressive performance can be achieved by stereoscopic arrays consisting of tens of 10 m diameter class (HESS-type) telescopes. The flux sensitivity 10^{-14} erg/cm^2 s at TeV energies would be a great achievement even in the standards of the most advanced branches of observational astronomy. This should allow
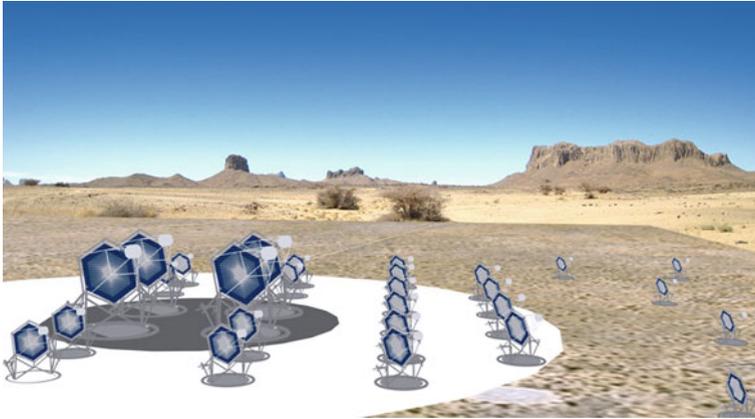
Fig. 12 Layout of the future Cherenkov telescope array

us to probe the gamma-ray luminosities of potential TeV emitters at the levels of $10^{32}(d/10\,\mathrm{kpc})^2$ erg/s for galactic sources and $10^{40}(d/100\,\mathrm{Mpc})^2$ erg/s for extra-galactic objects. Although for moderately extended sources, e.g. of angular size $\Psi \sim 1°$, the minimum detectable energy flux will be by a factor of $\Psi/\delta\theta \sim 10$–30 higher, yet it would be better than the energy flux sensitivities of the best current X-ray satellites, *Chandra*, *XMM*-Newton, *INTEGRAL* and *Suzaku*, i.e. should allow the deepest probes of non thermal high energy phenomena in extended sources, in particular in shell type Supernova Remnants (SNRs), Giant Molecular Clouds (GMCs), Pulsar Driven Nebulae (Plerions), Clusters of Galaxies, hypothetical Giant Pair Halos around AGN, etc. Such a system of 10–12 m diameter class IACTs with a field of view (FoV) of 6–$8°$, most likely will constitute the core of the Cherenkov Telescope Array (CTA)—an initiative towards the major ground-based gamma-ray detector (see Fig. 12). The huge area covered by tens of telescope should provide dramatic increase of the gamma-ray photon statistics. On the other hand, the detection of the cascades in many projections will improve the angular resolution, and the efficiency of suppression of hadronic showers. All these factors should allow an improvement of the sensitivity of CTA in the TeV region, compared to the current instruments, by a factor of five (see Fig. 7). Correspondingly, the required observation time of the objects would be reduced by a factor of 25.

Sub-PeV Regime

The general tendency of decreasing gamma-ray fluxes with energy becomes especially dramatic above 30 TeV. The reasons could be different, e.g. external and internal absorption of gamma-ray, limited efficiency of particle acceleration processes, escape of highest energy particles from the production region, etc. Any meaningful study of cosmic gamma-rays beyond 30 TeV requires detection areas

well exceeding $1 \, \text{km}^2$. An effective and straightforward approach would be the use of IACT arrays optimized for detection of gamma-rays in the region up to 100 TeV and beyond. This can be realized by a modest, approximately $10\text{–}30 \, \text{m}^2$ area reflectors separated from each other, depending on the scientific objectives and the configuration of the imagers, between 300 and 500 m. The requirement to the pixel size of imagers is also rather modest, $0.25°$ or so, however they should have large, up to $10°$ FoV for simultaneous detection of showers from distances of several 100 m [209]. A sub-array consisting of several tens of such telescopes is included in the concept of CTA with a primary goal of study of energy spectra of gamma-ray sources well beyond 10 TeV. It will serve as a powerful too for searches of galactic cosmic ray 'PeVatrons', albeit nearby ($R \ll 10$ Mpc radio galaxies and starburst galaxies can be considered high priority targets as well.

Sub-100 GeV Regime

The energy threshold ε_{th} of IACTs is generally defined as a characteristic energy at which the gamma-ray detection rate for a primary power-law spectrum with a photon index $\Gamma = 2\text{–}3$ achieves its maximum. It is well known, from Monte Carlo simulations and from the operation of previous generation IACTs, that in practice the best performance, in particular the minimum detectable energy flux, is achieved at energies exceeding several times ε_{th}. In this regard, for optimization of gamma-ray detection around 100 GeV, one should reduce the energy threshold of telescopes to $\varepsilon_{\text{th}} \leq 30 \, \text{GeV}$. This can be done by using very large, 20 m-diameter class reflectors. On the other hand, the reduction of the detection threshold to 30 GeV is an important scientific issue in its own right; the intermediate interval between 30 and 300 GeV is a crucial energy regime for certain class of galactic and extragalactic gamma-ray source populations. A sub-array consisting of several very large telescopes foreseen in CTA (see Fig. 12) will indeed significantly broaden the topics and scientific objectives of CTA.

The above three sub-arrays of Cherenkov telescopes are characterized by specific energy intervals in which the best energy flux sensitivity is achieved. On the other hand, each of these arrays covers at least 2 decades in energy. thus the energy domains of the sub-arrays largely overlap. Since these arrays contain the same basic elements, and generally have overlaps in scientific motivations, an ideal arrangement would be the combination of these sub-arrays in a single facility which would have a sensitive and homogeneous coverage throughout the energy region from approximately 30 GeV–300 TeV. The conceptual design of CTA is motivated, to a large extent, by this objective [23]. The high detection rates, coupled with good angular and energy resolutions over four energy decades will make CTA a powerful multi-function and multi-purpose gamma-ray observatory with a great capability for spectrometric, morphological and temporal studies of a diverse range of persistent and transient high energy phenomena in the Universe.

Multi-GeV Regime: *Gamma-Ray Timing Explorers*

Despite the recent great achievements of high energy gamma-ray astronomy, there are obvious shortcomings in the performance of the current so-called "pair-conversion" tracking detection technique—the most effective approach used in the satellite-borne instruments for detection of gamma-rays at energies above several tens of MeV. First of all this concerns the low angular resolution, especially at energies below 1 GeV and poor photon statistics at energies above 10 GeV. While there is a room for improvement of the angular resolution at low energies, e.g. reducing it to 1° around 100 MeV (by using very thin $e^{\pm}$ pair-converters above the tracking detectors), the increase of the gamma-ray photon statistics at high energies is a more difficult task. One should note in this regard that the flux sensitivity of *Fermi* LAT at 1 GeV of about 10^{-12} erg/cm^2 s can be achieved only after one year all-sky survey. While for persistent gamma-ray sources this seems to be an adequate sensitivity (given that a huge number of sources are simultaneously monitored within the large, almost $\sim 2\pi$ steradian homogeneous FoV), the small, $\approx 1\,\mathrm{m}^2$ detection area limits significantly the potential of this instrument for detailed studies of the temporal and spectral characteristics of highly variable sources like blazars or solitary events like gamma-ray bursts (GRBs). The sensitivity achieved by *Fermi* at high energies will be not easy to improve significantly by any future space-based mission, unless the Moon would be used in (far) future as a possible platform for installation of very large ($\gg 10\,\mathrm{m}^2$) area pair-conversion tracking detectors. It is clear that the space-based resources of GeV gamma-ray astronomy have achieved a point where any further progress would appear extremely difficult and very expensive. In any case, for the next decades to come there is no space-based mission planned for the exploration of the high-energy gamma-ray sky. On the other hand, the principal possibility of extension of the Imaging Atmospheric Cherenkov Telescope (IACT) technique towards 10 GeV promises a new breakthrough in gamma-ray astronomy [27]. The (relatively) large gamma-ray fluxes in this energy interval, together with the huge detection areas offered by the IACT technique, can provide the highest gamma-ray photon statistics compared to any other energy band of cosmic gamma-radiation. Thus, in the case of realization of 10 GeV threshold IACT arrays, the presently poorly explored interval between 10 and 100 GeV could become one of the most advanced domains of gamma-ray astronomy with a great potential for studies of highly variable phenomena.

The reduction of the energy threshold down to 10 GeV or even less is principally possible within the basic concept of the IACT technique, but it requires an extreme approach—25 to 30 m diameter class telescopes equipped with very high ($\geq 40\,\%$) quantum efficiency focal plane imagers, operating in a robotic regime at very high (5 km or) mountain altitudes [66].

The energy range from several GeV to 30 GeV has very specific astrophysical and cosmological objectives: exploration of highly variable non thermal phenomena, in particular in the remote universe at redshifts of z = 5, as well as in compact galactic objects like pulsars and microquasars. The successful realization of such a gamma-ray timing explorer, hopefully during the lifetime of the *Fermi* observatory would be a great achievement for gamma-ray astronomy.

The remarkable success of observational gamma-ray astronomy, together with recent intensive theoretical and phenomenological studies of acceleration and radiation processes in astrophysical objects, supply a strong rationale for the further exploration of the sky at high and very-high energies. Although generally the main motivations of gamma-ray astronomy remain unchanged, the recent observational results have introduced important corrections to our understanding of many relevant phenomena. They revealed new features which in many cases require revisions of current theoretical models and formulations of new concepts. It is expected that over the next decade the ongoing operation of *Fermi* will be accompanied by observations with the next generation ground-based detectors. The data obtained in the enormous energy range from 100 MeV to 1 PeV will provide very deep insight into a number of problems of high energy astrophysics and fundamental physics.

3 Radiation Mechanisms

3.1 General Comments

Basic Processes

Any interpretation of an astrophysical observation starts with identification of relevant radiation mechanism(s). Therefore the comprehensive knowledge on principal radiation and absorption processes is a key issue in astrophysics. With some exceptions, all basic processes of gamma-ray production and absorption have been studied in great detail using the methods and tools of modern experimental and theoretical physics. Each gamma-ray domain has its specifics and is characterized by several radiation processes. Generally, the most important mechanisms in the high and very-high energy domains are

- *electron bremsstrahlung*
- *decay of neutral π^0-mesons* produced at p–p and p–A interactions
- *inverse Compton scattering*

The first two mechanisms operate effectively in (relatively) dense and extended gaseous environments. While the second mechanism is important in all energy bands above 100 MeV, the electron bremsstrahlung contributes mostly to sub-TeV gamma-rays. The reason is the synchrotron and inverse Compton (IC) cooling of electrons which typically at very high energies starts to dominate over the bremsstrahlung losses. For the same reason the inverse Compton scattering is a very important process of gamma-ray production at TeV energies, but it could be a dominant radiation mechanism also at lower energies. It works effectively almost everywhere, from compact objects like pulsars and AGN to extended sources like supernova remnants and clusters of galaxies. Since the Compton cooling time decreases linearly with energy, the process becomes especially effective at very high energies.

Because of universal presence of dense radiation fields, ultrarelativistic electrons always release a significant fraction of their energy through IC gamma-ray, except for regions where the energy density of the magnetic field significantly exceeds the energy density of the surrounding radiation fields. In such cases, the energy of electrons is released in synchrotron radiation which typically does not extend beyond X-ray energies. Therefore, from the point of view of production of high energy gamma rays, the synchrotron cooling generally plays a negative (dissipative) role. However, in the so-called *extreme accelerators*, where the particle acceleration proceeds at the maximum possible rate allowed by classical electrodynamics [61], the *synchrotron radiation* of both electrons and protons can become the dominant channel of gamma-radiation, especially in highly magnetized environments moving with relativistic speeds towards the observer [51]. In radiation dominated environments the synchrotron radiation of secondary charged μ-mesons and π-mesons also may contribute significantly to very high energy gamma-radiation [198].

In strong magnetic fields, when particles move along the field lines, they radiate through the so-called *curvature* radiation (see e.g. [201]). This mechanism is generally believed to be the most important process responsible for gamma-radiation of pulsars, but it can effectively work also in other astronomical objects, in particular in the vicinity of rotating supermassive black holes [61, 181]. Thus, in compact and highly magnetized objects, high energy gamma-rays can be effectively produced through

- *magnetobremsstrahlung of electrons and protons*

under which one should understand the radiation of electrons moving in magnetic fields in a general sense, i.e. both in the synchrotron and curvature radiation regimes.

In the case of combination of certain favourable conditions regarding, in particular, the existence of dense surrounding radiation fields and an effective confinement of relativistic particles, gamma-rays can be effectively produced at interactions of protons with low-energy photons, directly through decays of π^0-mesons, the products of

- *photomeson reactions*: $p + \gamma \rightarrow \pi^0 \rightarrow \gamma$

or indirectly, through synchrotron radiation of secondary electrons, the products of photomeson reactions and

- *Bethe-Heitler pair production*: $p + \gamma \rightarrow e^+ e^-$

Another mechanism of gamma-ray production related to interactions of accelerated particles with low-frequency photons is (e.g. Ref. [219]) the de-excitation of nuclei following

- *photo–disintegration interactions*: $A + \gamma \rightarrow A'^* + X; A^* \rightarrow A' + \gamma$

However, despite some recent optimistic claims (see e.g. Refs. [77, 199]), the efficiency of conversion of energy of accelerated nuclei to gamma-rays in optically thin objects appear very small, $\eta \ll 1$. This is a rather generic and robust limit [50] which makes negligible the role of this mechanism in almost all realistic models of

cosmic gamma-ray production. In any case, the photo-disintegration reaction rates are significantly below the rates of photomeson and Bethe-Heitler pair-production (BHPP) processes.

Generally, because of small cross-sections, the interactions of ultrarelativistic protons and nuclei with photon fields are rather slow processes. Of course, the rates can be higher in compact objects at the presence of dense radiation fields. However, the rates are always limited because of the requirement to the source to be optically thin for gamma-rays regarding the

- *photon–photon pair production*: $\gamma + \gamma \rightarrow e^+ e^-$

The cross-section of this process exceeds by orders of magnitude the cross-sections of the photomeson and pair-production cross-sections, therefore the condition $\tau_{\gamma\gamma} \leq 1$ implies very low efficiency of photomeson and BHPP processes unless the protons are very effectively confined in the source. And, vice versa, the high efficiency of the photomeson and BHPP processes implies large optical depth of the source. In this case, the products of these reactions, electrons, positrons and gamma-rays, initiate electromagnetic cascades in the low-energy photon gas, which wash out the signatures of primary particles, thus the energy spectrum of gamma-rays that emerge the source is basically determined by the specifics of the cascade development, rather than the initial spectra of the first generation particles.

Gamma-rays are also absorbed due to pair production at interactions with the gas and magnetic fields. While the first process generally has a negligible effect in most of the astrophysical environments, the pair production of very high energy gamma-rays is a very important process in objects with very strong magnetic field—first of all in pulsars and perhaps also in the magnetospheres of rotating black holes.

Relativistic Electron–Photon Cascades

The interactions of relativistic electrons, either accelerated directly or being secondary products of various hadronic processes, with ambient targets in forms of *matter, radiation* and *magnetic fields* result in gamma-rays. Generally, the gamma-ray production is effective when the radiative cooling time does not significantly exceed (i) the source age, (ii) the time of non-radiative losses caused by adiabatic expansion or by particle escape, and (iii) the cooling time of competing radiation mechanisms resulting in low-energy photons *outside* the gamma-ray domain. Note that although an effective gamma-ray production requires dense targets, if the charged particles are effectively confined to the gamma-ray production region, these conditions could be fulfilled even in environments with relatively low gas and photon densities or a weak magnetic field. For example, the gamma-ray production efficiency could be close to 1 even in very extended objects, when $t_{\text{rad}} \gg R/c$; R is the characteristic linear size of the production region, c is the speed of light.

The major gamma-ray production mechanisms associated with relativistic electrons have their 'counterparts'—gamma-ray *absorption* mechanisms of the same electromagnetic origin, which result in electron–positron pair production in matter

(the counterpart of bremsstrahlung), in photon gas (the counterpart of inverse Compton scattering), and in magnetic field (the counterpart of synchrotron radiation). The gamma-ray production mechanisms and their absorption counterparts have similar cross-sections, therefore the condition for radiation $t_{\rm rad} \geq R/c$ generally implies small optical depths for the corresponding gamma-ray absorption process, $\tau_{\rm abs} \leq 1$.

But in many astrophysical scenarios, in particular in compact galactic and extragalactic objects with favourable conditions for particle acceleration, the radiation processes can be very fast, so that $t_{\rm rad} \leq R/c$. Since the energy threshold of pair production implies that the corresponding 'counterpart' (production) processes proceed in relativistic regimes, namely (i) $E_{\gamma,\rm e} \gg 2m_e c^2$ in matter, (ii) $E_{\gamma,\rm e}\omega_0 \gg m_{\rm e}^2 c^4$ in photon gas (Klein-Nishina regime), or (iii) $(E_{,\rm e}/m_{\rm e}c^2)(B/B_{\rm crit}) \gg 1$ in the magnetic field (quantum regime), when a large fraction of the electron energy is transferred to the secondary photon, the problem cannot be reduced to the absorption effect. In this regime, the secondary electrons produce a new generation of high energy gamma-ray, these photons again produce electron–positron pairs, etc., so an electromagnetic cascade develops.

The characteristics of electromagnetic cascades in matter have been comprehensively studied in literature, in particular the context of interactions of cosmic rays with the Earth's atmosphere. These results can be applied to a broad class of the so-called "hidden sources", objects of high energy radiation with dense surrounding gas: super-massive black holes, compact X-ray binaries, young pulsars inside dense shells of recent supernovae explosions, etc. (see e.g. Ref. [96]). If the thickness of the surrounding gas significantly exceeds $100\,{\rm g/cm^2}$, the protons accelerated in the central source would initiate electron–photon cascades While "hidden sources" are considered as potential neutrino sources, they are less attractive targets for high energy gamma-ray astronomy. However, the gamma-ray emission in these objects is not fully suppressed. The recycled radiation with spectral features determined by the thickness ('grammage') of the gas shell, should be seen in gamma-rays, unless the synchrotron radiation of secondary electrons dominates over the bremsstrahlung, and thus channels the main fraction of the non thermal energy into the sub-gamma-ray domain.

The development of electromagnetic cascades in photon gas and magnetic fields is a more common phenomenon in astrophysics. In photon fields, such cascades are created on broad astronomical scales, from compact objects like gamma-ray bursts to the intergalactic medium as a whole. Very high energy gamma-rays emitted by astronomical objects and interacting with diffuse extragalactic photon fields initiate electromagnetic cascades in the entire Universe. The superposition of contributions of gamma-rays from these cascades can constitute a significant fraction of the observed diffuse extragalactic background.

The importance of electron–photon cascades in astrophysical objects supported by $\gamma-\gamma$ pair-production and inverse Compton scattering has been first realized in Ref. [106]. The development of these cascades is well described by the so-called compactness parameter [147] $l = L\sigma_{\rm T}/Rm_{\rm e}c^3$, where L is the luminosity and R is the radius of the source. When l is less than 10, the cascade develops in the linear

regime, i.e. when the soft radiation produced by cascade electrons does not have a significant feedback effect on the cascade development [67].

The development of cascades in magnetic fields are best studied for pulsar magnetospheres (see e.g. Refs. [89, 148, 236]. Such cascades could be triggered in some other sites as well, e.g. in the Earth's geomagnetic field [208] and in blazars [91], etc. The pair cascades in magnetic fields are effective when the product of the particle (photon or electron) energy and the strength of the B-field becomes close to the "quantum threshold" of about $EB \geq B_{\text{crit}} m_e c^2 \simeq 2 \times 10^7$ TeV Gauss. A simple approach, similar to the so-called approximation A used for cascades in matter, in general can satisfactorily describe the basic features of cascades in magnetic field [70], but it does not provide adequate accuracy for a quantitative description of the cascade characteristics [78].

As long as we are interested in the one-dimensional cascade development (which seems to be quite sufficient for many astrophysical purposes), all 3 types of cascades can be described by the same integro-differential equations as the ones derived by Landau and Rumer [180], but in each case specifying the cross-sections of the relevant interaction processes. These equations over a broad range of energies has been studied using numerical solutions of the so-called adjoint cascade equations has been recently conducted by [68].

Although for certain astrophysical scenarios the development of cascades in 'pure' environments can be considered as an appropriate and fair approximation, in some conditions the interference of the processes associated with interactions of cascade electrons and gamma-rays with both the ambient photon gas and magnetic field (or matter) can significantly change the character of cascade development, and consequently the spectra of observed gamma-rays. The impact of such interference is very complex and quite sensitive to the choice of the principal parameters. Therefore each practical case is a subject to independent studies.

Basic Definitions

Calculations of emissivity of gamma rays resulting from interactions of accelerated particles require two functions—the differential cross section of the process, $\frac{d\sigma}{dEd\Omega}$, and the energy distribution of parent particles, $f(E)$. While for given energy of primary particles E, the cross-sections are generally well know functions of the energy and emission angle of secondary products, the distribution functions of parent particles depends on the acceleration mechanism (e.g. diffuse shock acceleration or stochastic acceleration) and the local conditions of the ambient medium—the strength of magnetic field, the shock speeds, the level of turbulence, etc.

The general kinetic equation that describes the evolution of the particle energy distribution $f(E, \boldsymbol{r}, t)$ can be written in the form

$$\frac{\partial f}{\partial t} = \nabla \cdot (D_r \nabla f) - \nabla \cdot (\mathbf{u}_r f) + \frac{\partial}{\partial E} (P_r f) - \frac{\partial}{\partial E} (b_r f) + \frac{\partial^2}{\partial E^2} (d_r f), \quad (1)$$

which includes the terms responsible for the diffusion, acceleration and energy losses of particles. For the derivation of this equations, its approximate solutions under specific assumptions, and for implications of these solutions the reader is referred to the original source [140].

Strictly speaking, this equation should be invoked in any detailed treatment of gamma-ray production. However, this implies that one has to make an assumption about the acceleration mechanism which in many cases is highly unknown. Thus the inclusion of acceleration aspect in the overall treatment of the problem not only would introduce additional uncertainties, but in some cases could be redundant. In such cases, it would be reasonable to assume an injection of particles, protons and/or electrons, into the gamma-ray production rate with a given history of rate and energy spectrum, $Q(E, t)$. Such an approach could be quite appropriate, for example when the accelerator of particles and the target (the gamma-ray production region) are separated, or if the acceleration zone is much smaller than the gamma-ray production zone. In this case Eq. (1) integrated over the volume, can be reduced to the following simple form

$$\frac{\partial N}{\partial t} = \frac{\partial}{\partial E}(PN) - \frac{N}{\tau_{\text{esc}}} + Q(E, t) \tag{2}$$

where

$$N(E, t) \equiv \int f \, d^3 r. \tag{3}$$

Here $P(E, t) = -\partial E/\partial t$ is the energy loss rate for a particle with energy E, and the parameter $\tau_{\text{esc}}(E, t)$ is the characteristic escape time of particles from the source due to diffusion and convection, $\tau_{\text{esc}} = (1/\tau_{\text{diff}} + 1/\tau_{\text{conv}})^{-1}$. The injection function generally is assumed to be a power-law with a high-energy cut-off, $Q(E, t) \propto E^{-\alpha} \exp\left[-(E/E_0)^{\beta}\right]$.

Here β characterizes the sharpness of the spectrum in the region of the cut-off. The cut-off energy can be estimated equating the acceleration rate to the cooling rate, $t_{\text{cool}} = t_{\text{acc}}$. Due to energy losses, the particle spectrum in the emission region is modified, and can significantly deviate from the original injection spectrum. The steady-state particle distribution can be obtained by setting $\partial N/\partial t = 0$ in Eq. (2). If we further assume that the term proportional to τ_{esc}^{-1} can be neglected compared to the term containing the energy loss rate, the solution for $N(E)$ is very simple,

$$N(E) = \left|\frac{dE}{dt}\right|^{-1} \int_E^{\infty} Q(E) \, dE. \tag{4}$$

Several specific realizations for energy losses are of particular interest. For example, if the energy losses are proportional to the particle energy, $dE/dt \propto E$, which is the case of relativistic Bremsstrahlung, proton-proton inelastic collisions and adiabatic losses, the initial (injection) power-law spectrum of particles remains unchanged, $N(E) \propto E^{-\alpha}$. The steady-state distribution is steeper if electrons are cooled due to

synchrotron radiation or inverse Compton scattering (IC) in the Thomson regime, $dE/dt \propto E^2$, namely $N(E) \propto E^{-(\alpha+1)}$. On the other hand, the spectrum becomes harder if dE/dt is constant (as for ionization losses) or $dE/dt \propto E^{-1}$ (the case of IC scattering in the Klein-Nishina limit): $N(E) \propto E^{-(\alpha-1)}$ and $N(E) \propto E^{-(\alpha-2)}$, respectively. These features are reflected in the steady-state gamma-ray spectra.

An order-of-magnitude estimate of gamma-ray luminosity can be obtained using the cooling time ($t_{\text{cool}} = E/|dE/dt|$) of the corresponding radiation mechanism,

$$L_\gamma \approx W_{\text{e}(p)} t_{\text{cool}}^{-1}, \tag{5}$$

where $W_{e(p)}$ is the total energy budget of relativistic particles. If V is the volume of the region, and E_{min} and E_{max} are the minimum and maximum energy of the accelerated particles, respectively, $W_{\text{e}(p)}$ can be calculated as

$$W_{\text{e}(p)} = \int_V \int_{E_{\text{min}}}^{E_{\text{max}}} E \, N_{\text{e}(p)}(E) \, dE \, d^3x. \tag{6}$$

The shape of the gamma-ray spectral energy distribution can be obtained, in some cases with a good accuracy, by applying the so-called δ-function approximation which assumes that the same fixed fraction of the energy of parent particle goes to secondary gamma-rays. However, the δ-function formalism should be applied with care, especially near the cut-off of distribution. Otherwise one may arrive at wrong conclusions (see below) about the distribution of parent particles derived from the energy distribution of the radiation using the δ-function approximation.

3.2 Brief Overview of Important Processes

Interactions with Matter

The electron bremsstrahlung and the decay of neutral π-mesons produced at inelastic *pp* collisions are the most effective high energy gamma-ray production processes in astrophysical environments related to interactions with matter.

Electron Bremsstrahlung

Often the electron bremsstrahlung is treated together with its counterpart process— the Bethe-Heitler pair production (BHPP) by gamma-rays at interactions with electrons and nuclei. Comprehensive description of cross-sections of these two processes can be found in many monographs, in particular in Ref. [149]. Although these two processes have many similarities, the electron bremsstrahlung has much broader
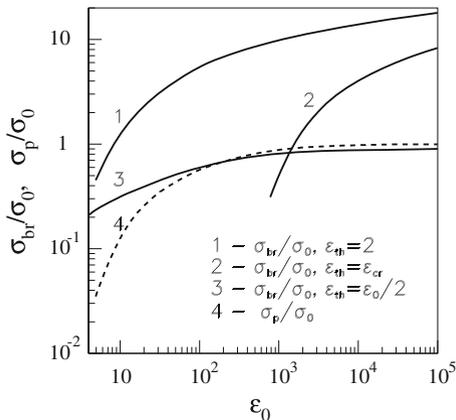
Fig. 13 Total cross-sections of the bremsstrahlung (σ_{br}) and pair production (σ_p) processes in hydrogen normalised to the asymptotic value (σ_0) of the pair production cross-section at $\varepsilon_0 \to \infty$. The bremsstrahlung cross-sections are calculated for secondary gamma-ray produced with energies exceeding (*1*) the pair-production threshold, $\varepsilon_{th} = 2$; (*2*) the critical energy, $\varepsilon_{th} \simeq 700$; (*3*) half of the energy of the primary electron, $\varepsilon_{th} = \varepsilon_0/2$

astrophysical applications than BHPP. The latter becomes effective only in rare astrophysical environments, e.g. in very dense shell surrounding particle accelerators.

The integral cross-sections of the bremsstrahlung and pair production processes in hydrogen gas are shown in Fig. 13. The energies of electrons and γ-rays are expressed in units of $m_e c^2$. The cross-sections are normalised to the asymptotic value of the pair production cross-section at $\varepsilon_0 \to \infty$:

$$\sigma_0 = 7/9 \times 4\alpha_f r_e^2 Z(Z+1) \frac{\ln(183Z^{-1/3})}{1 + 0.12(Z/82)^2} \tag{7}$$

where Z is the charge of the target nucleus, and r_e is the classical electron radius. Based on σ_0, a new parameter called *radiation length* is introduced

$$X_0^{(m)} = 7/9(n\sigma_0)^{-1}, \tag{8}$$

the physical meaning of which is the average distance over which the ultrarelativistic electron loses the main fraction of its energy due to bremsstrahlung. The same parameter also implies the mean free path of γ-rays. This parameter is widely used to describe the cascade development in optically thick sources. The cascade effectively develops at depths exceeding the radiation length. Usually the radiation length is expressed in units of g/cm^2. For hydrogen gas $X_0^{(m)} \simeq 60$ g/cm^2. The second important parameter that characterises the cascade development is the critical energy below which ionization energy losses dominate over bremsstrahlung losses. In hydrogen gas, $\varepsilon_{cr} \simeq 700$. Effective multiplication of particles in a cascade is possible only at
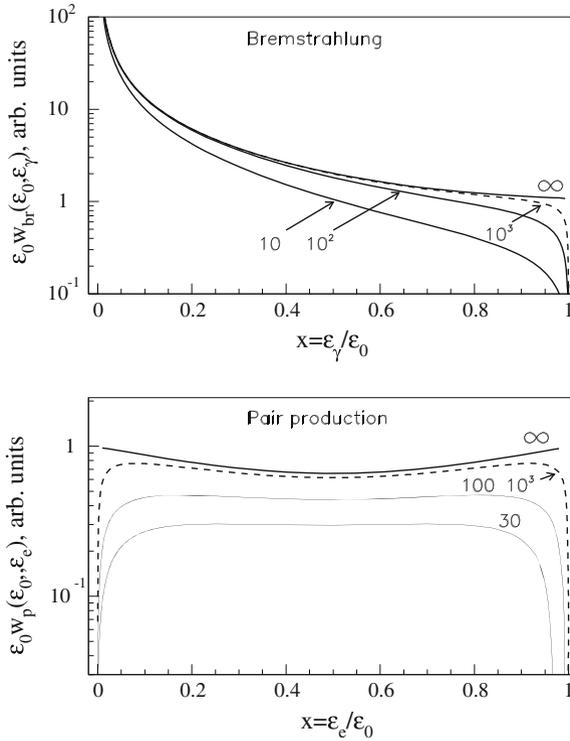
Fig. 14 Differential cross-sections of the bremsstrahlung (*upper panel*) and pair production (*bottom panel*) processes in hydrogen. The cross-sections are normalised to one radiation length. The energies of primary electrons and gamma-ray ε_0 (in units of $m_e c^2$) are indicated at the *curves*

energies $\varepsilon \geq \varepsilon_{cr}$. At lower energies electrons dissipate their energy by ionization rather than producing more high energy gamma-ray which would support further development of the electron–photon shower.

In Fig. 13 the bremsstrahlung total cross-sections are shown calculated for 3 different values of minimum energy of emitted gamma-ray: $\varepsilon_{th} = 2$, ε_{cr} and $\varepsilon_e/2$. It is seen from Fig. 13 that while for $\varepsilon_{th} = 2$ the pair-production cross-section is an order of magnitude smaller compared to the bremsstrahlung cross-section, for $\varepsilon_{th} = \varepsilon_e/2$ the cross-sections of two processes become almost identical at $\varepsilon \geq 100$.

The differential cross-sections of bremsstrahlung and pair production are presented in Fig. 14. The pair-production cross-section obviously is a symmetric function around the point $x = \varepsilon_e/\varepsilon_0 = 0.5$. The bremsstrahlung differential cross-section has a $1/\varepsilon_\gamma$ type singularity at $\varepsilon_\gamma \to 0$, but because of the hard spectrum of bremsstrahlung photons the energy losses of electrons contribute mainly to high energy γ-rays. Thus bremsstrahlung should be treated as an essentially catastrophic process. Nevertheless, it is convenient to introduce the so-called average energy loss-rate,

$$-\left(\frac{\mathrm{d}E_e}{\mathrm{d}t}\right)_{\mathrm{br}} = \left(\frac{cm_p n}{X_0}\right) E_e. \qquad (9)$$

Correspondingly, the lifetime of electrons is

$$t_{\mathrm{br}} = \frac{E_e}{-\mathrm{d}E_e/\mathrm{d}t} \simeq 4 \times 10^7 (n/1\,\mathrm{cm}^{-3})^{-1}\,\mathrm{year}, \qquad (10)$$

where n is the number density of the ambient gas.

Note that the electron energy loss rate given by Eq. (9) is proportional to the electron energy, and, correspondingly, the lifetime given by Eq. (10) is energy independent. This implies that for a initial (acceleration) power-law spectrum $Q(E_e)$, bremsstrahlung losses do not change the original electron spectrum, taking into account that the cooled steady-state spectrum $N(E_e)$ is given by Eq. (4).

Interestingly, in the case of power-law spectrum of electrons $N(E_e)$, the spectrum of bremsstrahlung gamma-rays is also power low with the same power-law index (this is a result of $1/E_\gamma$ dependence of the differential cross-section). Thus, the bremsstrahlung gamma-ray spectrum simply repeats the shape of the electron acceleration spectrum $Q(E_e) \propto E_e^{-\Gamma}$.

This is true, however, only when the energy losses are dominated by bremsstrahlung. In hydrogen gas, at energies below $\sim 700 m_e c^2 \simeq 350\,\mathrm{MeV}$ ionization dominates over the bremsstrahlung. Because both the ionization and bremsstrahlung loss rates are proportional to n, this condition does not depend on the ambient gas density. On the other hand, in the relativistic regime, the ionization loss rate does not depend on the electron energy. Thus, in accordance with Eq. (4), the steady-state electron spectrum becomes flatter, $N(E_e) \propto E_e^{-\Gamma+1}$, and correspondingly at energies below several hundred MeV we should expect a very hard bremsstrahlung gamma-ray spectrum with power-law photon index $\Gamma - 1$.

In many astrophysical scenarios the inverse Compton and synchrotron losses may well dominate, especially at very high energies, over bremsstrahlung, depending on the ratio of the energy density of the radiation and magnetic fields to the number density of the ambient gas. The synchrotron and inverse Compton energy loss rates are proportional to the electron energy. This makes the steady-state electron spectrum steeper, $N(E_e) \propto E_e^{-\Gamma-1}$, and correspondingly the bremsstrahlung gamma-rays emerge with photon index $\Gamma + 1$.

The positron annihilation is another channel of gamma-ray production related to interactions with matter. The astrophysical significance of this process generally is linked to the annihilation line at energy $m_e c^2 = 0.511\,\mathrm{MeV}$, as well as to the 3-photon positronium continuum, produced by annihilation of thermalized positrons with background cold electrons of the ambient gas/plasma. However, if positrons are injected into the production region with relativistic energies, a significant fraction of positrons (from 10 to 20 %, depending on the ionization state of the ambient plasma) annihilate in flight before they cool down to the temperature of the thermal gas [55]. The energy spectrum of gamma-rays has a symmetric form with two sharp peaks at $E_\gamma \sim m_e c^2$ and $E_\gamma \sim E_+$. This implies that the energy of relativistic positron

is transferred mainly to the leading photon. For a power-law spectrum of positrons, $N_+ \propto E_+^{-\alpha}$ the spectrum of annihilation radiation at $E_\gamma \gg m_c^2$ has a power-law form

$$J_{\text{ann}}(E_\gamma) \propto E_\gamma^{-(\alpha+1)} \left[\ln(2E_\gamma/m_c^2) - 1 \right] . \qquad (11)$$

The total cross-section of annihilation of a relativistic positron of energy ε_+ is given by $\sigma_{\text{ann}} = \frac{3}{8}\sigma_{\text{T}}^2(E_+/m_ec^2)^{-1}[\ln(2E_+/m_ec^2) - 1]$. Correspondingly the annihilation time is

$$t_{\text{ann}} \simeq 4 \times 10^6 \frac{(E_+/m_c^2)}{\ln(2E_+/m_c^2) - 1} (n/1\,\text{cm}^{-3})^{-1}\,\text{year}. \qquad (12)$$

Comparing Eq. (12) with Eq. (10) one finds that for a positron of energy $E_+ \leq 15\,\text{MeV}$, the annihilation times becomes shorter than the bremsstrahlung cooling time, i.e. at these energies the annihilation continuum starts to dominate over the bremsstrahlung spectrum, taking into account that in both processes the leading photon receives a substantial part of the positron kinetic energy.

Gamma Rays and Neutrinos from pp Interactions

Relativistic protons and nuclei produce high energy gamma-rays and neutrinos in inelastic collisions with ambient gas due to the production and decay of secondary pions, kaons and hyperons. The neutral π^0-mesons provide the main channel of conversion of the kinetic energy of protons to high energy gamma-rays. For the production of π^0-mesons the kinetic energy of protons should exceed $E_{\text{th}} = 2m_\pi c^2(1 + m_\pi/4m_p) \approx 280\,\text{MeV}$, where $m_\pi = 134.97\,\text{MeV}$ is the mass of the π^0-meson. This particle immediately decays to two gamma-rays. The mean lifetime of π^0-decay, $t_{\pi^0} = 8.4 \times 10^{-17}$ s, is significantly shorter than the lifetime of charged π-mesons ($\approx 2.6 \times 10^{-8}$ s). At high energies, all three types of pions are produced with comparable probabilities. The spectral form of π-mesons is generally determined by a few (one or two) leading particles (that carry a significant fraction of the nucleon energy) rather than by the large number of low-energy secondaries.

The decays of charged pions lead to neutrinos with spectra quite similar to the spectrum of the accompanying π^0-decay gamma-rays. However, this symmetry can be violated in environments with high gas or radiation densities. In certain conditions, the characteristic time for inelastic interactions of charged pions with nucleons or photons could be shorter than the decay time, so the energy of pions degrades before they decay. At very high energies this would result in significantly smaller fluxes of neutrinos compared to gamma-rays.

The distinct feature of the spectrum of π^0-decay γ-rays is the maximum at $E_\gamma = m_\pi c^2/2 \simeq 67.5\,\text{MeV}$, independent of the energy distribution of π^0 mesons, and consequently of the parent protons. The appearance of such a bump in the gamma-ray spectrum is a result of the $\pi^0 \to 2\gamma$ decay kinematics. It is easy to show (see e.g. Ref. [220] that the spectrum of gamma-rays from decays of mono energetic pions

of energy E_π and velocity v_π is constant $f(E_\pi) = c/(v_\pi E_\pi)$ within the interval between $E_1 = 0.5E_\pi(1 - v_\pi/c)$ and $E_2 = 0.5E_\pi(1 + v_\pi/c)$. The spectrum of gamma-rays for an arbitrary distribution of π^0-mesons $\Pi(E_\pi)$ can be presented as superposition of rectangles for which only one point at $m_\pi c^2/2$ is always presented. Obviously this should result in a spectral maximum independent of the distribution of parent pions.

Although precise calculations of gamma-ray spectra require quite heavy integrations over differential cross-sections obtained experimentally at particle accelerators, the emissivity of γ-rays for an arbitrary *broad* energy distribution of protons without sharp features can be derived within a simple formalism of δ-function approximation which nevertheless provides surprisingly good over a broad gamma-ray energy range [58].

The gamma-ray emissivity $q_\gamma(E_\gamma)$ is directly defined by $q_\pi(E_\pi)$ as

$$q_\gamma(E_\gamma) = 2 \int_{E_{\min}}^{\infty} \frac{q_\pi(E_\pi)}{\sqrt{E_\pi^2 - m_\pi^2 c^4}} dE_\pi, \tag{13}$$

where $E_{\min} = E_\gamma + m_\pi^2 c^4/4E_\gamma$. The emissivity of secondary pions q_π from inelastic proton–proton interactions can be calculated with high accuracy using accelerator measurements of the inclusive cross-sections $\sigma(E_i, E_p)$. The emissivity of π^0-mesons calculated in the δ-function approximation for the cross-section $\sigma(E_\pi, E_p)$ then becomes

$$q_\pi(E_\pi) = c\, n_{\rm H} \int \delta(E_\pi - \kappa_\pi E_{\rm kin})\sigma_{\rm pp}(E_p)n_{\rm p}(E_p){\rm d}E_p$$
$$= \frac{c\, n_{\rm H}}{\kappa_\pi}\sigma_{\rm pp}\left(m_p c^2 + \frac{E_\pi}{\kappa_\pi}\right) n_{\rm p}\left(m_p c^2 + \frac{E_\pi}{\kappa_\pi}\right) \tag{14}$$

where $\sigma_{\rm pp}(E_p)$ is the total cross section of inelastic pp collisions, and κ_π is the mean fraction of the kinetic energy $E_{\rm kin} = E_p - m_p c^2$ of the proton transferred to the secondary π^0-meson per collision; $n_{\rm p}(E_p)$ is the energy distribution of the protons. In a broad region from GeV to TeV energies $\kappa_\pi \approx 0.17$ which includes a $\sim 6\%$ contribution from η-meson production [139]. From the threshold at $E_{\rm kin} \simeq 0.3$ GeV, $\sigma_{\rm pp}$ rises rapidly to about 30 mb. But after $E_{\rm kin} \sim 2$ GeV, $\sigma_{\rm pp}$ increases only logarithmically. The measurements of the total inelastic cross-section are shown in Fig. 15 together with a simple analytical approximation [165]:

$$\sigma_{\rm inel}(E_p) = (34.3 + 1.88\, L + 0.25\, L^2) \times [1 - (E_{\rm th}/E_p)^4]^2 \, {\rm mb}, \tag{15}$$

where $E_{\rm th} = m_p + 2m_\pi + m_\pi^2/2m_p = 1.22$ GeV is the threshold energy of the proton for π^0 production. This expression correctly describes $\sigma_{\rm inel}(E_p)$ near the threshold, and fits the experimental data and SYBILL simulations up to at least $E_p \sim 10^4$ TeV.

The figure shows a plot with σ_{inel}, mb on the vertical axis and E_p, TeV on the horizontal axis. The equation inside the figure reads:

$$\sigma_{\text{inel}} = (34.3 + 1.88\,L + 0.25\,L^2) \times$$

$$\left[1 - \left(\frac{E_{th}}{E_p}\right)^4\right]^2 \text{, mb}$$
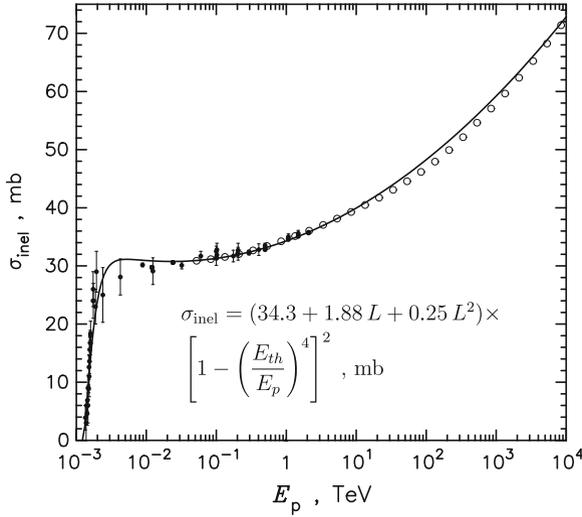
Fig. 15 Inelastic cross-section of *pp* interactions approximated by Eq. (15). The experimental data are from http://wwwppds.ihep.su:8001/c5-5A.HTML, the open points correspond to the cross-sections which are used in the SIBYLL code

The δ-function approach yields the gamma-ray emissivity with good accuracy for a *broad* energy distribution of protons without sharp features. Indeed, in the δ-function formalism, the shape of the photon spectrum is similar to the shape of the parent proton spectrum, shifted in energy by a factor κ_π. However, this approximation does not give correct results if the proton energy distribution contains sharp features like a pile-up or a cut-off [165].

The energy spectra of gamma-rays from *pp* interactions have been comprehensively studied recently by Kelner et al. [165] and Kamae et al. [162] who also developed convenient procedures for calculations of gamma-ray spectra for arbitrary energy distributions of protons. In particular, in Ref. [165] simple analytical parametrizations of energy spectra of secondary particles—gamma rays, electrons, and neutrinos have been obtained based on simulations of proton-proton interactions using the public available SIBYLL code.

For a proton of energy E_p, the number of gamma-rays per a *pp* collision from the decays of secondary π^0 and η-mesons, in the interval $(x, x + dx)$, where $x = E_\gamma/E_p$, can be presented as

$$F_\gamma\left(x, E_p\right) = B_\gamma \frac{\ln x}{x} \left[\frac{1 - x^{\beta_\gamma}}{1 + k_\gamma x^{\beta_\gamma}\left(1 - x^{\beta_\gamma}\right)}\right]^4$$

$$\times \left[\frac{1}{\ln x} - \frac{4\beta_\gamma x^{\beta_\gamma}}{1 - x^{\beta_\gamma}} - \frac{4k_\gamma \beta_\gamma x^{\beta_\gamma}\left(1 - 2x^{\beta_\gamma}\right)}{1 + k_\gamma x^{\beta_\gamma}\left(1 - x^{\beta_\gamma}\right)}\right] \quad (16)$$
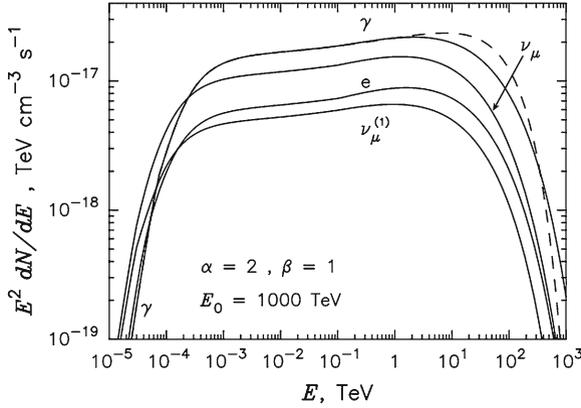
Fig. 16 Spectra of gamma-ray, electrons and neutrinos from inelastic pp collisions, for power-law proton distributions given by Eq. (18) with $\alpha = 2$, $\beta = 1$, and $E_0 = 1000$ TeV. The *dashed curve* is the photon spectrum calculated in the δ-function approximation, with $\kappa_\pi = 0.17$. Two curve for all muonic neutrinos (ν_μ) corresponds to the contributions from two channels—$\nu_{\mu(1)}$: $\pi \to \mu\nu_\mu$, and $\nu_{\mu(2)}$: $\mu \to e\nu_e\nu_\mu$. The spectrum of electronic neutrinos coincides, with an accuracy better than 5 %, with the electron spectrum

with $B_\gamma = 1.30 + 0.14L + 0.011L^2$, $\beta_\gamma = (1.79 + 0.11L + 0.008L^2)^{-1}$, and $k_\gamma = (0.801 + 0.049L + 0.014L^2)^{-1}$, where $L = \ln(E_p/1\,\text{TeV})$.

Equation (16) allows one to derive the gamma-ray emissivity $q_\gamma(E_\gamma)$ for an arbitrary distribution of protons $N_p(E_p)$:

$$
q_\gamma(E_\gamma) = c\,n_\text{H} \int_{E_\gamma}^{\infty} \sigma_\text{inel}(E_p)\,N_p(E_p)\,F_\gamma\left(\frac{E_\gamma}{E_p}, E_p\right)\frac{dE_p}{E_p}
$$
$$
= c\,n_\text{H} \int_0^1 \sigma_\text{inel}\left(\frac{E_\gamma}{x}\right) N_p\left(\frac{E_\gamma}{x}\right) F_\gamma\left(x, \frac{E_\gamma}{x}\right)\frac{dx}{x}, \tag{17}
$$

where the inelastic pp cross section $\sigma_\text{inel}(E_p)$ is given by Eq. (15).

Figure 16 shows the spectra of secondary gamma-ray, electrons and neutrinos obtained for the proton distribution given in the form

$$
N(E) = A\,E^{-\alpha}\exp\left[-\left(\frac{E}{E_0}\right)^\beta\right], \tag{18}
$$

with $\alpha = 2$, $\beta = 1$, and $E_0 = 1000$ TeV. The constant A is set from the condition that the energy density of protons $w = \int EN(E)dE = 1\,\text{erg/cm}^3$.

It should be noted that, the spectrum of gamma-rays deviates, in contrast to commonly used assumptions, from the power-law form of the parent proton spectrum even at energies well below the cut-off energy and well above the threshold of π^0 production. This is explained by the increase of the total cross-section

with energy; the effect becomes significant especially at energies of protons above 1 TeV (see Fig. 15). Another interesting feature is the shape of the gamma-ray spectrum in the cut-off region. The δ-function approximation which agrees quite well with the results of accurate calculations at low energies, fails to describe correctly the spectrum at highest energies around the cut-off (see Fig. 16). In particular, it predicts $\exp[-(E_\gamma/E_{\gamma,0})]$ type cut-off term with $E_{\gamma,0} = k_\pi E_0 \simeq 0.17 E_0$, i.e. the same shape as the proton spectrum but shifted by a factor of 6 towards low energies. The accurate calculations result in a significantly different shape, namely $\exp[-(E_\gamma/E_{0,\gamma})^{1/2}]$ with $E_{0,\gamma} \simeq 0.06 E_0$. This implies that the gamma-ray spectrum displays a smoother cut-off, and the impact of the proton cut-off starts earlier, already around $E_\gamma \sim 0.01 E_0$.

Finally we note that for $\alpha = 2$, the amplitude of the gamma-ray spectrum at energies $E \ll E_0$ exceeds the level of the flux of muonic neutrinos. The difference is due to the contribution of η-mesons in the production of gamma-rays; If one takes into account only the decays of π^0-mesons, the spectra of gamma-rays and muonic neutrinos appear almost identical at energies well below the cut-off region. This conclusion does not contradict the fact that in each pp interaction the number of muonic neutrinos is a factor of 2 larger, on average, than the number of gamma-rays. The imbalance is compensated, in fact, at low energies (see Fig. 16).

Processes in Magnetic Fields

Synchrotron Radiation

Many important results of the theory of synchrotron radiation can be obtained within the framework of classical electrodynamics (see e.g. Ref. [140]) which is limited by the condition

$$\chi = \frac{E}{m_e c^2} \frac{B}{B_{cr}} \ll 1, \tag{19}$$

where $B_{cr} = m_e^2 c^3/e\hbar \approx 4.4 \times 10^{13}$ G is the critical value of the magnetic field.

The synchrotron cooling time for a particle of mass m and energy E in a magnetic field of strength B is

$$t_{synchr} = 6\pi \left(\frac{m}{m_e}\right)^3 \frac{m_e c^2}{c\sigma_T B^2} \frac{mc^2}{E}. \tag{20}$$

The cooling time of a protons is much longer, by a factor of $\left(m_p/m_e\right)^4 \approx 10^{13}$, compared to the synchrotron cooling time of an electron of same energy. Therefore in astrophysics the electron synchrotron radiation is far more important process compared to the proton synchrotron radiation, although the latter in some specific conditions can also be an effective mechanism of gamma-radiation.

The energy distribution of photons radiated by an electron of energy E is given by equation (see e.g. Ref. [140])

$$\frac{dN_\gamma}{dE_\gamma dt} = \frac{\sqrt{3}}{2\pi} \frac{e^3 B}{m_e c^2 \hbar E_\gamma} F\left(\frac{E_\gamma}{E_c}\right), \tag{21}$$

where

$$F(x) = x \int_x^\infty K_{5/3}(\tau) d\tau. \tag{22}$$

Here $K_{5/3}$ is the modified Bessel function of order $5/3$, and

$$E_c = \frac{3e\hbar B}{2m_e c} \left(\frac{E}{m_e c^2}\right)^2 \tag{23}$$

is the so-called characteristic energy of synchrotron radiation.

In a regular magnetic field, the strength B in these equations should be considered as the perpendicular component of the field, $B_\perp = B\sin\theta$, where θ is the pitch-angle. In chaotic magnetic fields, Eq. (21) should be averaged over directions of magnetic field, and the function $F(x)$ should be replaced by the function $G(x)$ which can be presented in the form [65]

$$G(x) = \frac{x}{20}[(8 + 3x^2)(\kappa_{1/3})^2 + x\kappa_{2/3}(2\kappa_{1/3} - 3x\kappa_{2/3})], \tag{24}$$

where $\kappa_{1/3} = K_{1/3}(x/2)$, $\kappa_{2/3} = K_{2/3}(x/2)$ (for a different presentation of this function in terms of Whittaker's function see Ref. [120]). Note that while the function $F(x)$ has a maximum at $x = 0.2858$ (max $F(x) = 0.9180$), the maximum of the function $G(x)$ is shifted towards smaller values: $x = 0.2292$ (max $G(x) = 0.7126$). The functions $F(x)$ and $G(x)$, as well as the ratio $G(x)/F(x)$ are shown in Fig. 17.

Although the function $G(x)$ in Eq. (24) has a quite compact form, for practical purposes often it is desirable to have an approximation which does not contain special functions. In particular, the following convenient approximations for functions $F(x)$ and $G(x)$ obtained in Ref. [65]

$$F(x) \approx 2.15x^{1/3}(1 + 3.06x)^{1/6} \times \frac{1 + 0.884x^{2/3} + 0.471x^{4/3}}{1 + 1.64x^{2/3} + 0.974x^{4/3}} e^{-x}, \tag{25}$$

$$G(x) \approx \frac{1.808x^{1/3}}{\sqrt{1 + 3.4x^{2/3}}} \frac{1 + 2.21x^{2/3} + 0.347x^{4/3}}{1 + 1.353x^{2/3} + 0.217x^{4/3}} e^{-x}. \tag{26}$$

provide an accuracy better than 0.2% over the entire range of variable x.

The spectral energy distributions (SED) of synchrotron photons corresponding to distributions of electrons presented in the form of Eq. (18), can be described by a simple analytical expression

$$L_\gamma \propto E_\gamma^{-\frac{\alpha-1}{2}} \exp\left[-\left(E_\gamma/E_{\gamma,0}\right)^\lambda\right], \tag{27}$$

Fig. 17 The emissivity
functions for synchrotron
radiation $F(x)$ and $G(x)$. The
dashed line shows the ratio
$G(x)/F(x)$
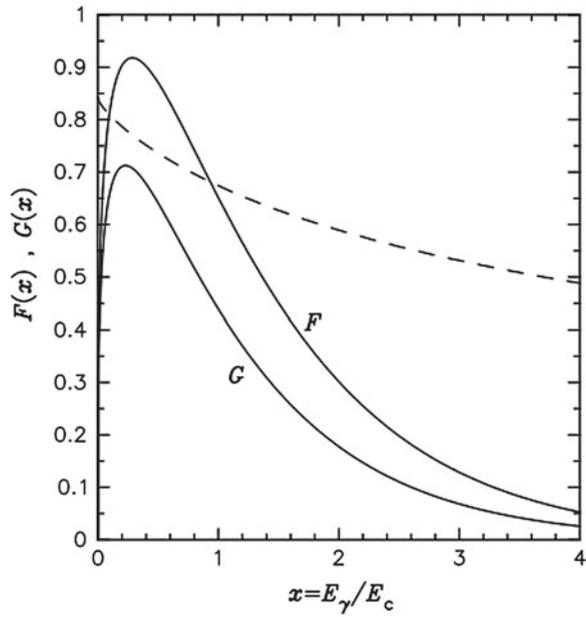
In Fig. 18 the results of accurate calculations are compared with the ones obtained
within the δ-function approximation. It is seen that while the latter describes correctly
the power-law part of the spectrum of synchrotron radiation, it fails to reproduce
correctly the radiation spectrum in the cut-off region. Indeed, this approximation
for the index in the exponential cut-off in Eq. (27) predicts $\lambda = \beta/2$ which is quite
different from the results of exact analytical calculations, $\lambda = \beta/(\beta+2)$ [135, 253].
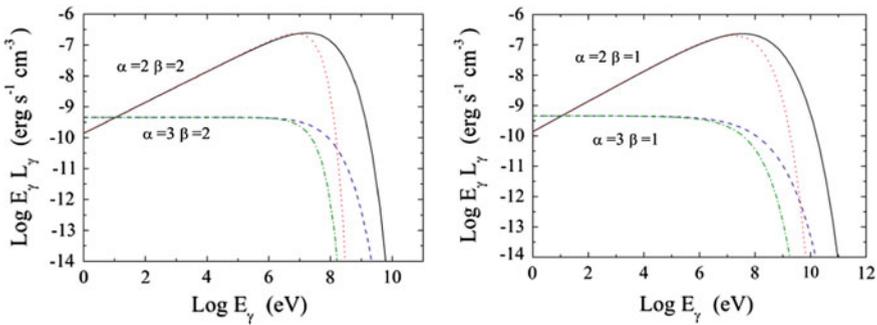
Fig. 18 Spectral energy distributions of synchrotron radiation from a distribution of electrons as
in Eq. (18), with indices $\alpha = 2, 3$ and $\beta = 1$ (*left panel*) and $\beta = 2$ (*right panel*). In all cases
$E_0 = 10^{15}$ eV and $B = 1$ mG. The *solid* and *dashed curves* are the result of exact calculations,
whereas the *dotted* and *dash-dotted curves* were obtained applying the δ-function approximation

The spectra of synchrotron radiation in the cut-off region appear much smoother compared to the ones calculated in the δ-function approximation.

The shape of the particle spectrum in the cut-off region depends, among other factors, on the specific mechanisms of acceleration and energy losses. The cut-off energy itself can be estimated from the balance between the rates of acceleration and energy losses. The particle acceleration time can be written as

$$t_{\text{acc}} = \eta\,(E)\,\frac{E}{ecB}. \tag{28}$$

The parameter $\eta\,(E) \geq 1$ characterizes the acceleration rate, and generally (in most models) is treated as a free parameter. Note that $E_{\text{eff}} = \eta^{-1}B$ can be treated as the projection of an electric field E on the particle trajectory, effectively averaged as the particle moves along the trajectory. This representation is useful, in particular, in a formal treatment of the diffusive shock acceleration [61]. In the majority of cosmic accelerators $B \gg E$, and correspondingly $\eta \gg 1$. However in certain situations η can be close or even less than 1, namely when $E \geq B$. When such a condition is satisfied, the particles are unbound in the acceleration region, and the energy gain is determined by the difference of the electric potential across the acceleration region.

If energy losses are dominated by synchrotron radiation, the condition $t_{\text{acc}} = t_{\text{synchr}}$ yields a maximum energy

$$E_0 = \left(\frac{3}{2}\right)^{3/4} \frac{m^2 c^4}{\sqrt{\eta\,e^3 B}}. \tag{29}$$

The corresponding cut-off energy of the synchrotron spectrum appears independent of the magnetic field B (see e.g. Ref. [51]).

$$E_{\gamma,0} = \frac{9}{4}\frac{mc^2}{\alpha_f \eta}. \tag{30}$$

Here $\alpha_f = e^2/\hbar c \approx 1/137$ is the fine-structure constant. For $\eta = 1$, the cut-offs appear at energies $\approx 160\,\text{MeV}$ and $\approx 300\,\text{GeV}$ for synchrotron radiation by electrons and protons, respectively. Any violation of these limits would imply acceleration of particles at conditions when $E > B$. Realization of this condition is difficult, but *a priori* cannot be excluded in compact relativistic objects.

In more standard scenarios, $B \gg E$, i.e. $\eta \gg 1$. For example in the case of non-relativistic diffusive shock acceleration,

$$\eta \approx 10\frac{D(E)}{r_{\text{g}}c}\left(\frac{c}{v_{\text{s}}}\right)^2, \tag{31}$$

where $r_{\text{g}} = E/eB$ is the particle gyroradius, $D(E)$ is the diffusion coefficient, and v_{s} is the shock speed. If diffusion proceeds in the Bohm regime, $D(E) = r_{\text{g}}c/3$, thus

$$E_{\gamma,0} \approx \frac{27}{40} \frac{mc^2}{\alpha_f} \left(\frac{v_s}{c}\right)^2.$$ (32)

For young supernova remnant in the Sedov phase, the shock speeds are typically of order of several thousand km/s, therefore the synchrotron peaks appear in the range of a few keV.

Magnetic Pair Production

In general, the energy of the synchrotron photons is much less than the energy of the parent particle. But in certain astrophysical environments, synchrotron radiation can take place near the quantum threshold. In this limit, the production of electron–positron pairs in the magnetic field by high-energy photons,

$$\gamma + B \longrightarrow e^+ + e^-.$$ (33)

becomes equally important. In the quantum regime, synchrotron radiation and magnetic pair production are tightly coupled. The operation of these two processes together may initiate and support an effective electron–photon cascade development in environments with very large magnetic field, provided that the mean free path of gamma-rays or electrons are smaller than the size of the source.

The probabilities of interaction per unit length (also called differential cross sections) of an electron or a photon with the magnetic field are [78]

$$\omega_{\text{synchr}}\left(E_e, E_\gamma\right) = \frac{\alpha_f m_e c^3}{\pi\sqrt{3}\hbar} \frac{1}{E^2} \left[\left(\frac{E_e - E_\gamma}{E} + \frac{E_e}{E_e - E_\gamma}\right) K_{\frac{2}{3}}\left(\frac{2u}{3\chi}\right)\right.$$
$$\left. - \int_{\frac{2u}{3\chi}}^{\infty} K_{\frac{1}{3}}(y)\, dy \right],$$ (34)

and

$$\omega_p\left(E_\gamma, E_e\right) = \frac{\alpha_f m_e c^3}{\pi\sqrt{3}\hbar} \frac{1}{E_\gamma^2} \left[\left(\frac{E_\gamma - E_e}{E_e} + \frac{E_e}{E_\gamma - E_e}\right) K_{\frac{2}{3}}\left(\frac{2\tilde{u}}{3\chi}\right)\right.$$
$$\left. + \int_{\frac{2\tilde{u}}{3\chi}}^{\infty} K_{\frac{1}{3}}(y)\, dy \right],$$ (35)

respectively. Here $u = E_\gamma/(E_e - E_\gamma)$ and $\tilde{u} = E_\gamma^2/E_e(E_\gamma - E_e)$. The parameter χ is given by Eq. (19), with $E = E_e$ or $E = E_\gamma$ depending on the process.

Figure 19 shows the total cross sections σ_{synchr} and σ_p as a function of χ. These are obtained integrating Eqs. (34) and (35) over E_γ and E_e, respectively. For $\chi \ll 1$, the probability of synchrotron radiation is approximately constant, but the probability
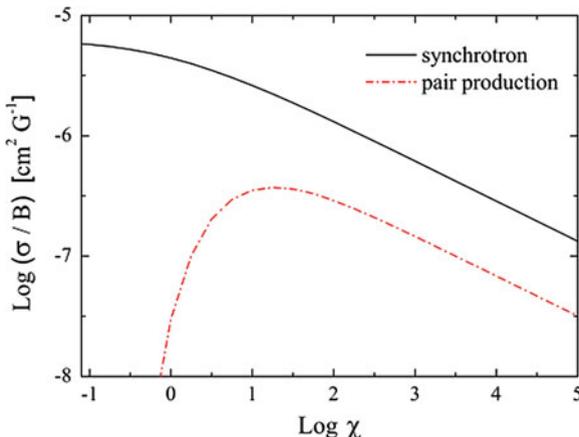
Fig. 19 Interaction probabilities of synchrotron radiation and magnetic pair production as a function of χ
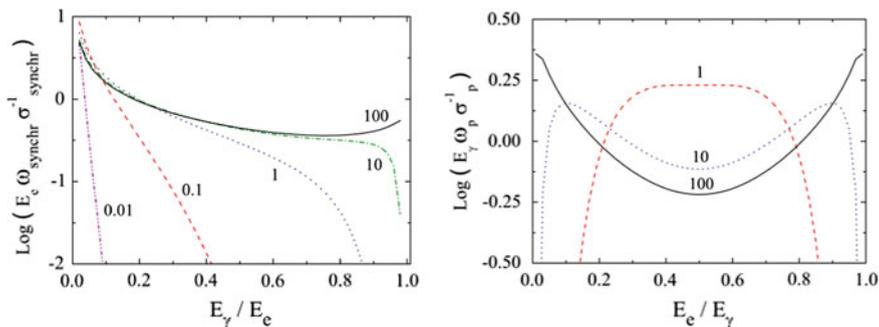
Fig. 20 Differential interaction probabilities (differential cross sections) of synchrotron radiation (*left panel*) and pair production in a magnetic field (*right panel*). Both quantities are normalised to the corresponding total cross section. The values of χ are indicated at each curve

of pair production drops abruptly as $\exp(-8/3\chi)$. For $\chi \gtrsim 10$, both cross sections decrease with χ as $\chi^{-1/3}$, but the probability of synchrotron radiation is several times larger.

Figure 20 shows the differential cross sections for both processes for different values of χ. The distribution of pairs created through interaction of gamma rays with the magnetic field is symmetric around $E_e = 0.5E_\gamma$. The pair spectrum is flat, and the photon transfers most of its energy to one of the created particles. The energy spectrum of synchrotron photons is very steep for $\chi \ll 1$, but becomes flat at $\chi \geq 1$.

Interactions with Photon Fields

The interaction of relativistic electrons with radiation fields through inverse Compton scattering provides one of the principal gamma-ray production processes in astrophysics. It works effectively almost everywhere, from compact objects like pulsars and AGN to extended sources like supernova remnants and clusters of galaxies. Since the Compton cooling time decreases linearly with energy, the process becomes especially effective at very high energies.

The electron–positron pair production in photon–photon collisions is tightly coupled with the inverse Compton scattering. First of all, it is an absorption process that prevents the escape of energetic gamma-rays from compact objects, and determines the "gamma-ray horizon" of the Universe. At the same time, in an environment where the radiation pressure dominates over the magnetic field pressure, the photon–photon pair production and the inverse Compton scattering "work" together supporting an effective transport of high energy radiation via electromagnetic "Klein-Nishina" cascades.

Although the IC scattering of protons is suppressed, compared to the scattering of electrons, by a factor of $(m_e/m_p)^4$, very high energy protons effectively interact with the ambient photon fields through electron–positron pair production and photomeson processes. While in the first process gamma-rays are produced indirectly, via inverse Compton scattering of the secondary electrons, photo-meson reactions result in the direct production of π^0-mesons and their subsequent decay to γ-rays. Typically, at extremely high energies these interactions proceed effectively both in compact objects and large scale structures.

Inverse Compton Scattering

The total cross-section of inverse Compton scattering depends only on the product of energies of the interacting electron ε and photon ω_0, $\kappa_0 = \omega_0 \varepsilon_e$ (all energies are in units of $m_e c^2$). In the non relativistic regime ($\kappa_0 \ll 1$) it approaches the classical (Thomson) cross-section $\sigma_{IC} \approx \sigma_T(1 - 2\kappa_0)$, while in the ultrarelativistic regime ($\kappa_0 \gg 1$) it decreases with κ_0 as $\sigma_{IC} \approx (3/8)\sigma_T \kappa_0^{-1} \ln(4\kappa_0)$. With an accuracy of better than 10 % in a very broad range of κ_0, the cross-section can be represented in the following simple form [118].

$$\sigma_{IC} = \frac{3\sigma_T}{8\kappa_0}\left[\left(1 - \frac{2}{\kappa_0} - \frac{2}{\kappa_0^2}\right)\ln(1 + 2\kappa_0) + \frac{1}{2} + \frac{4}{\kappa_0} - \frac{1}{2(1 + 2\kappa_0)^2}\right] \quad (36)$$

The total cross-section of Compton scattering as a function of κ_0 is shown in Fig. 21.

The energy distribution of up-scattered gamma-rays is determined by the differential cross-section of the process. Assuming that a mono energetic beam of low energy photons ω_0 penetrates an isotropic and homogeneous region filled with relativistic electrons of energy ε_e, the spectrum of radiation scattered at the angle θ relative to

Fig. 21 Total cross-sections of inverse Compton scattering and photon–photon pair production in isotropic radiation fields. Two spectral distributions for the ambient photon gas are assumed: (i) mono energetic with energy ω_0 (*curves 1 and 3*), and (ii) Planckian with the same mean photon energy $\omega_0 \simeq 3kT/m_ec^2$ (*curves 2 and 4*)
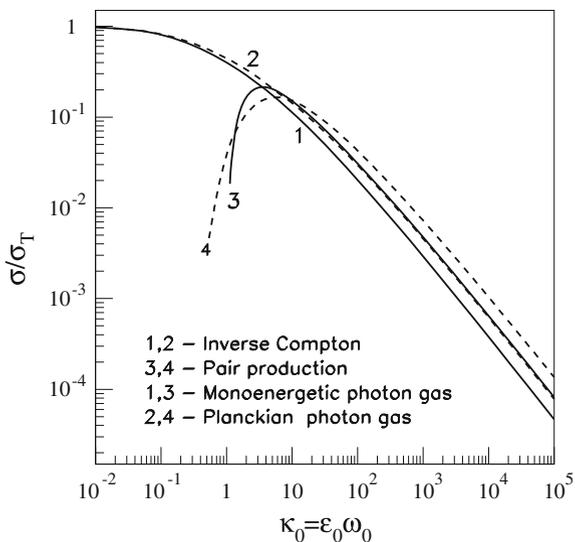
the initial photon beam is written as [54].

$$\frac{\mathrm{d}^2 N(\theta, \varepsilon_\gamma)}{\mathrm{d}\varepsilon_\gamma \mathrm{d}\Omega} = \frac{3\sigma_\mathrm{T}}{16\pi\omega_0\varepsilon_\mathrm{e}^2} \left[1 + \frac{z^2}{2(1-z)} - \frac{2z}{b_\theta(1-z)} + \frac{2z^2}{b_\theta^2(1-z)^2} \right], \quad (37)$$

where $b_\theta = 2(1-\cos\theta)\omega_0\varepsilon_\mathrm{e}$, $z = \varepsilon_\gamma/\varepsilon_\mathrm{e}$. The energy of the high energy γ-ray photon ε_γ varies in the limits $\omega_0 \ll \varepsilon_\gamma \ll \varepsilon_{\gamma,\max}$, where $\varepsilon_{\gamma,\max} = \varepsilon_\mathrm{e}b_\theta/(1+b_\theta)$, $b = 4\kappa_0$.

In the case of isotropically distributed electrons and photons, the integration of Eq. (37) over the angle θ gives [54, 98, 161].

$$\frac{\mathrm{d}N(\varepsilon_\gamma)}{\mathrm{d}\varepsilon_\gamma} = \frac{3\sigma_\mathrm{T}}{4\omega_0\varepsilon_\mathrm{e}^2} \left[1 + \frac{z^2}{2(1-z)} + \frac{z}{b(1-z)} - \frac{2z^2}{b^2(1-z)^2} \right.$$
$$\left. + \frac{z^3}{2b(1-z)^2} - \frac{2z}{b(1-z)} \ln\frac{b(z-1)}{z} \right]. \quad (38)$$

The differential energy spectra of gamma-rays for several fixed values of κ_0 are shown in Fig. 22. In the deep Klein-Nishina regime ($\kappa_0 \gg 1$) the spectrum grows sharply towards the maximum at $\varepsilon_{\gamma,\max}$. This implies that in this regime just one interaction is sufficient to transfer a substantial fraction of the electron energy to the upscattered photon. In the Thomson regime ($\kappa_0 \ll 1$) the average energy of the upscattered photon is $\varepsilon_\gamma \approx \omega_0\varepsilon_\mathrm{e}^2$, thus only a fraction $\varepsilon_\gamma/\varepsilon_\mathrm{e} \sim \kappa_0 \ll 1$ of the primary electron energy is released in the upscattered photon.

For a power-law distribution of electrons, $\mathrm{d}N_\mathrm{e}/\mathrm{d}\varepsilon_\mathrm{e} \propto \varepsilon_\mathrm{e}^{-\Gamma}$, the resulting gamma-ray spectrum in the non relativistic regime ($a = 4\omega_0\varepsilon_\gamma \ll 1$) has a power-law form
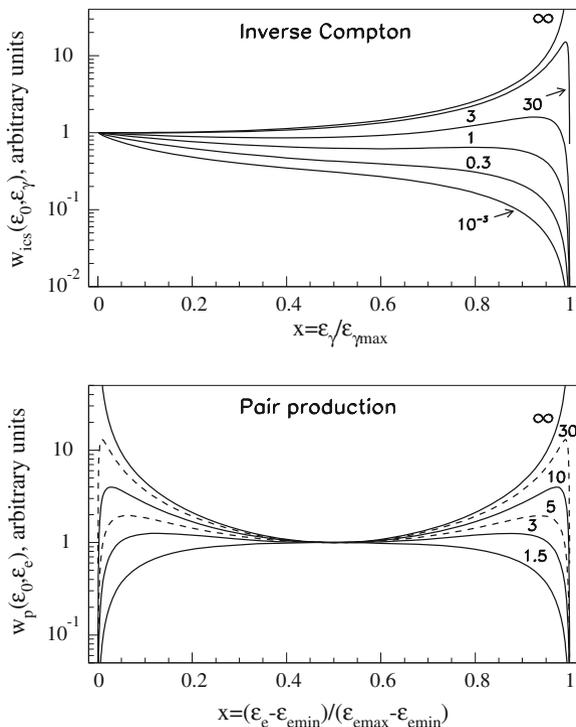
Fig. 22 Differential spectra of gamma-rays from inverse Compton scattering (*upper panel*) and electrons from photon–photon pair production (*bottom panel*) in an isotropic and mono-energetic photon field. The parameters $\varepsilon_{\gamma,max}$, $\varepsilon_{e,min}$ and $\varepsilon_{e,max}$ are defined as $\varepsilon_{\gamma\,max} = 4\varepsilon_0(\kappa_0/1+4\kappa_0)$ and $\varepsilon_{e\,min,e\,max} = 0.5\varepsilon_0(1 \mp \sqrt{1 - 1/\kappa_0})$. The same values of the parameters $\kappa_0 = \varepsilon_e\omega_0$ and $s_0 = \varepsilon_\gamma\omega_0$ are indicated by the *curves*

with photon index $\alpha = (\Gamma + 1)/2$ [140]. In the ultrarelativistic ($a \gg 1$) regime the gamma-ray spectrum is noticeably steeper, $\propto \varepsilon_\gamma^{-\alpha}(\ln a + \text{const})$ with $\alpha = (\Gamma + 1)$ [98]. Useful analytical approximations for gamma-ray spectra over a broad energy interval have been obtained in Refs. [54, 118].

The energy-loss rate of relativistic electrons in a mono-energetic field of photons with energy ω_0 and number density n_{ph} is given by the following equation [54]

$$\frac{d\varepsilon_e}{dt} = \frac{3\sigma_T c n_{ph}}{4\omega_0 b}\left[\left(6 + \frac{b}{2} + \frac{6}{b}\right)\ln(1 + b) - \ln^2(1 + b)\right.$$
$$\left. - 2\text{Li}\left(\frac{1}{1+b}\right) - \frac{(11/12)b^3 + 8b^2 + 13b + 6}{(1 + b)^2}\right]. \quad (39)$$

where $\text{Li}(x) = \int_x^1 (1 - y)^{-1}\ln(y)\,dy$.

In the Thomson and Klein-Nishina regimes Eq. (39) reduces to the well known expressions [98].

$$\frac{d\varepsilon_e}{dt} = \frac{4}{3}\sigma_T c\omega_0 n_{ph}\varepsilon_e^2 \qquad \text{at} \quad b \ll 1, \tag{40}$$

and

$$\frac{d\varepsilon_e}{dt} = \frac{3}{8}\frac{\sigma_T cn_{ph}}{\omega_0} (\ln b - 11/6) \qquad \text{at} \quad b \gg 1. \tag{41}$$

The energy losses in these two regimes have quite a different dependence on the electron energy. While in the Thomson regime the loss rate is proportional to ε_e^2, in the Klein-Nishina regime it is almost energy independent. This implies that in the first case the energy losses make the electron spectrum steeper, whereas the in the Klein-Nishina regime the electron spectrum becomes harder.

Photon–Photon Pair Production

Photon–photon pair production is the inverse process to pair annihilation. Therefore the differential cross-section is identical to the pair annihilation cross-section, except for a different phase-space volume. In the relativistic regime this process is quite similar also to inverse Compton scattering. However, unlike the pair annihilation and Compton scattering, the photon–photon pair production has a strict kinematic threshold given by

$$\varepsilon_{\gamma 1}\varepsilon_{\gamma 2}(1 - \cos\theta) \geq 2, \tag{42}$$

where $\varepsilon_{\gamma 1}$ and $\varepsilon_{\gamma 2}$ are the energies of two photons in units of $m_e c^2$ colliding at an angle θ (in the laboratory frame).

The large cross-section makes the photon–photon pair production one of the most relevant elementary processes in high energy astrophysics. Several convenient approximations for the total cross-section of this process in the isotropic radiation field have been proposed in Refs. [60, 118, 145]. With an accuracy of better than 3 %, the total cross-section in the monoenergetic isotropic photon field can be represented in the following analytical form

$$\sigma_{\gamma\gamma} = \frac{3\sigma_T}{2s_0^2}\left[\left(s_0 + \frac{1}{2}\ln s_0 - \frac{1}{6} + \frac{1}{2s_0}\right)\ln(\sqrt{s_0} + \sqrt{s_0 - 1})\right.$$
$$\left. - \left(s_0 + \frac{4}{9} - \frac{1}{9s_0}\right)\sqrt{1 - \frac{1}{s_0}}\right]. \tag{43}$$

The total cross-sections of inverse Compton scattering and pair production in an isotropic monoenergetic photon field of energy ω_0 are shown in Fig. 21 (curves 1 and 3, respectively). Both cross sections depend only on the product of the primary

(ε_e or ε_γ) and target photon (ω_0) energies, $\kappa_0 = \varepsilon_e\omega_0$ and $s_0 = \varepsilon_\gamma\omega_0$. While as $\kappa_0 \to 0$, the inverse Compton cross-section approaches the Thomson cross-section, $\sigma_{IC} \approx \sigma_T(1 - 2\kappa_0)$, as $s_0 \to 1$ the pair production cross-section approaches zero, $\sigma_{\gamma\gamma} \approx (1/2)\sigma_T(s_0-1)^{3/2}$. For $\kappa_0, s_0 \gg 1$ the two cross-sections are quite similar and decrease with κ_0 and s_0: $\sigma_{IC} \approx (3/8)\sigma_T\kappa_0^{-1}\ln(\kappa_0)$, $\sigma_{\gamma\gamma} \approx (2/3)\sigma_T s_0^{-1}\ln(s_0)$. The pair-production cross-section has a maximum at the level of $\sigma_{\gamma\gamma} \approx 0.2\sigma_T$ achieved at $s_0 \approx 3.5$–4.

The parameter that characterises gamma-ray absorption at photon–photon inter-actions in a source of size R is the so-called optical depth

$$\tau(\varepsilon_\gamma) = \int_0^R \int_{\omega_1}^{\omega_2} \sigma(\varepsilon_\gamma, \omega)n_{ph}(\omega, r)\,d\omega\,dr, \tag{44}$$

where $n_{ph}(\omega, r)$ describes the spectral and spatial distribution of the target photon field in the source.

Because of narrowness of the pair-production cross-section, for a large class of broad band target photon energy distributions $n_{ph}(\omega)$, the optical depth at given gamma-ray energy ε_γ is essentially determined by a relatively narrow band of tar-get photons with energy centred on $\omega_* = 4/\epsilon_\gamma$. Therefore, the optical depth can be written in the form $\tau(\varepsilon_\gamma) = \eta(\sigma_T/4)\omega_* n_{ph}(\omega_*)R$, where the normalization factor η depends on the spectral shape of the background radiation. For a power-law tar-get photon spectrum, $n_{ph}(\omega) = n_0\omega^{-\alpha}$, the parameter η is calculated analytically, $\eta = (7/6)4^\alpha\alpha^{-5/3}(1 + \alpha)^{-1}$ [224].

The energy spectrum of electrons produced at photon–photon pair production has been studied in Refs. [60, 99, 118]. For a low-energy monoenergetic photon field ($\omega_0 \ll 1$), and correspondingly $\varepsilon_\gamma \gg 1$, the spectrum of electron–positron pairs can be represented, with an accuracy of better than a few percent, in the following analytical form [60]:

$$\frac{dN(\varepsilon_e)}{d\varepsilon_e} = \frac{3\sigma_T}{32\omega_0^2\varepsilon^3}\left[\frac{4\varepsilon_\gamma^2}{(\varepsilon_\gamma - \varepsilon_e)\varepsilon_e} \ln\frac{4\omega_0(\varepsilon_\gamma - \varepsilon_e)\varepsilon_e}{\varepsilon_\gamma} - 8\omega_0\varepsilon_\gamma \right.$$
$$\left. + \frac{2(2\omega_0\varepsilon_\gamma - 1)\varepsilon_\gamma^2}{(\varepsilon_\gamma - \varepsilon_e)\varepsilon_e} - \left(1 - \frac{1}{\omega_0\varepsilon_\gamma}\right)\frac{\varepsilon_\gamma^4}{(\varepsilon_\gamma - \varepsilon_e)^2\varepsilon_e^2} \right]. \tag{45}$$

The kinematic range of variation of ε_e is

$$\frac{\varepsilon_\gamma}{2}\left(1 - \sqrt{1 - \frac{1}{\omega_0\varepsilon_\gamma}}\right) \leq \varepsilon_e \leq \frac{\varepsilon_\gamma}{2}\left(1 + \sqrt{1 - \frac{1}{\omega_0\varepsilon_\gamma}}\right). \tag{46}$$

The differential energy spectra of gamma-rays for several fixed values of the parameters $s_0 = \omega_0\varepsilon_\gamma$ are shown in Fig. 22. The spectra are symmetric around the point $x = \varepsilon_e/\varepsilon_\gamma$. Although the average energy of the secondary electrons is $\varepsilon_\gamma/2$,

for very large s_0 the interaction has a catastrophic character—the major fraction of the energy of the primary photon is transferred to the leading electron.

These two processes—inverse Compton scattering and photon–photon pair production—determine the basic features of interactions of electrons and gamma-rays in the radiation dominated environments. At extremely high energies, the higher order QED processes may compete with these basic channels. Namely, when the product of the energies of colliding cascade particles (electrons or photons) E and the background photons ω significantly exceed $10^5 m_e^2 c^4$, the processes $\gamma\gamma \to e^+e^-e^+e^-$ and $e\gamma \to e\gamma e^+e^-$ may dominate over the single (e^+, e^-) pair production and Compton scattering, respectively. Because the $\gamma\gamma \to 2e^+2e^-$ and $e\gamma \to e\gamma e^+e^-$ channels result in production of 2 additional electrons, they substantially change the character of interactions. Note, however, that an effective realization of these processes is possible only under very specific conditions with an extremely low magnetic field and narrow energy distribution of the background photons.

Interactions of Hadrons with Radiation Fields

In astrophysical environments the radiation density often exceeds the density of gas component. In these conditions the interactions of high energy hadrons with radiation can dominate over interactions with matter, albeit the relevant cross-sections are relatively small. The main processes of hadron–photon interactions include (i) inverse Compton scattering: $p+\gamma \to p+\gamma'$, (ii) electron–positron pair production: $p+\gamma \to pe^+e^-$, (iii) photo-disintegration of nuclei: $A + \gamma \to A' + kN$, (iv) photomeson production: $N + \gamma \to N + k\pi$. In extremely dense radiation fields the secondary $\pi^\pm$-mesons may effectively interact with photons before they decay.

Except for the inverse Compton scattering, all other processes take place only above certain kinematic thresholds: ~ 1, 10, and 140 MeV (in the rest frame of projectile particles) for the pair production, photo-disintegration, and pion production, respectively.

The process of inverse Compton scattering of protons is identical to the inverse Compton scattering of electrons, but the energy loss rate of protons is suppressed, for the fixed energy of both particles, by a factor of $(m_e/m_p)^4 \approx 10^{-13}$. Generally, this process does not have noticeable astrophysical applications. At energies above the pair production threshold, the inverse Compton energy loss rate is significantly (by a factor of $\alpha(m_p/m_e)^2 \sim 10^4$) slower compared to the losses caused by pair-production.

In certain conditions the pair-production may result in significant spectral distortions of highest energy protons propagating through dense photon fields. The cross-section of this process, which in fact is the same Bethe-Heitler pair production, is quite large, but in each interaction only a small fraction of the proton energy is transferred to the secondary electrons. Therefore the energy loss rate of protons remains relatively slow. Moreover, the energy region where this process dominates is quite narrow. When the proton energy exceeds the pion production threshold, the photomeson interactions immediately start to dominate over the pair production.

The photo-disintegration of nuclei may have a strong impact on the formation of the chemical composition of very high energy cosmic rays in compact astrophysical objects as well as in the intergalactic medium. However, this process is an inefficient mechanism for gamma-ray production [50].

Photomeson production is much more important channel for transformation of the kinetic energy of protons into high energy gamma-rays, electrons and neutrinos. Close to the energy threshold, $\epsilon_{\text{th}} = m_\pi c^2 (1 + m_\pi / 2m_p) \approx 145$ MeV, the process proceeds through single-pion production, $p + \gamma \to p + \pi^0$, and $p + \gamma \to n + \pi^+$. At higher energies, multi-pion production channels begin to dominate. The cross-sections of these processes are basically well known from particle accelerator experiments. For astrophysical applications the data obtained with gamma-ray beams at energies from 140 MeV to 10 GeV are quite sufficient, if one takes into account the fact that for typical broad-band target photon spectra the hadron–photon interactions are contributed mainly from the region not far from the energy threshold, i.e. $E_\gamma \leq 1$ GeV.

The cross-sections of interactions of secondary electrons and gamma-rays with the ambient photons exceed by three orders of magnitude the photomeson cross-sections. Therefore the electrons and gamma-rays cannot leave the active region of pion production, but rather initiate electromagnetic cascades in the surrounding photon and magnetic fields. The standard spectra of the low-energy cascade gamma-rays that eventually escape the source are not sensitive to the initial spectral distributions, and thus contain information only about the total hadronic power of the source. On the other hand the secondary neutrinos freely escape the production region, and thus carry direct information about the energy spectra of accelerated protons.

Another interesting feature of the mixed hadronic/electromagnetic cascades in radiation dominated environments is the effective transport of primary nonthermal energy released in accelerated protons further away from a central engine through production and escape of secondary neutrons [129].

The presence of dense photon fields in the compact particle accelerators may have an even more fundamental impact. In relativistic flows, the multiple conversions of relativistic particle from charged to neutral state (proton→neutron→proton..., e → γ → e...) may allow a strong (up to the bulk Lorenz factor squared) energy gain in each cycle, whereas in the standard relativistic shock acceleration scenario the energy gain $\sim \Gamma^2$ occurs only in the first circle. This acceleration mechanism [123] is capable to boost protons in GRBs and AGN jets to maximum available energies, thus could be a key to the solution of the problem of the highest energy, $E \geq 10^{20}$ eV, particles observed in cosmic rays.

A very useful tool for detailed calculations of photomeson processes in radiation fields is the SOPHIA code [197]. For the given distributions of relativistic protons $N_p(E_p)$ and target photons $n_{\text{ph}}(\epsilon)$, the gamma-ray emissivity can be written as

$$\frac{dN_\gamma}{dE_\gamma} = \int N_p(E_p)\, n_{\text{ph}}(\epsilon)\, \Phi\,(\kappa,\ x)\, \frac{dE_p}{E_p}\, d\epsilon. \tag{47}$$

Here $\kappa = 4\epsilon E_p/m_p^2 c^4$ and $x = E_\gamma/E_p$.

Using numerical results obtained with the code SOPHIA, the function $\Phi(\kappa, x)$ can be approximated with an accuracy better than 10 % by a simple analytical expressions [164]. In the region $x_- < x < x_+$, the function $\Phi_\gamma(\kappa, x)$ can be presented in the form

$$\Phi_\gamma(\kappa, x) = B_\gamma \exp\left\{-s_\gamma \left[\ln\left(\frac{x}{x_-}\right)\right]^{\delta_\gamma}\right\} \left[\ln\left(\frac{2}{1+y^2}\right)\right]^{2.5+0.4\ln(\kappa/\kappa_0)}, \quad (48)$$

where

$$y = \frac{x - x_-}{x_+ - x_-}. \quad (49)$$

and

$$x_\pm = \frac{1}{2(1+\kappa)}\left[\kappa + r^2 \pm \sqrt{(\kappa - r^2 - 2r)(\kappa - r^2 + 2r)}\right]. \quad (50)$$

For $x < x_-$,

$$\Phi_\gamma(\kappa, x) = B_\gamma \left[\ln 2\right]^{2.5+0.4\ln(\kappa/\kappa_0)}, \quad (51)$$

and, finally, for $x > x_+$ $\Phi_\gamma(\kappa, x) = 0$. The parameters B_γ, s_γ and δ_γ as functions of κ are tabulated in Ref. [164].

Similar approximations have been obtained in Ref. [164] also for other secondary products. Figure 23 shows an example of energy spectra of the secondary gamma rays, neutrinos, and electrons produced in photomeson collisions of protons of energy $E_p = 10^{20}$ eV with 2.7 K CMBR. The chosen value of E_p is of particular interest in the study of the origin and propagation of cosmic rays. Due to interaction with the CMBR, only the most energetic gamma rays created at distances less than 1 Mpc can reach the observer. Pairs also cool rapidly in the intergalactic medium due to syn-
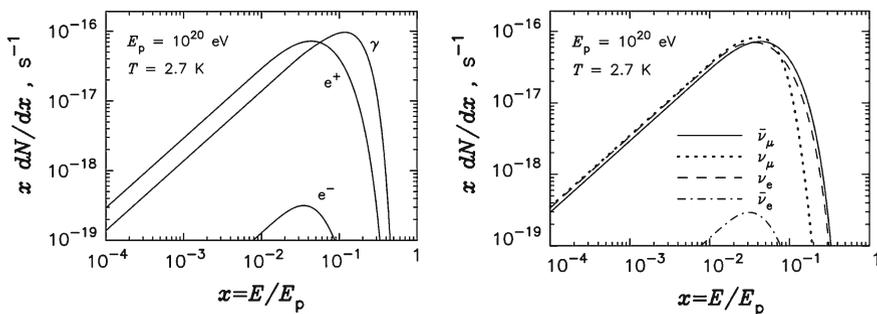
Fig. 23 Energy spectra of secondary products from photomeson interactions of protons of energy $E_p = 10^{20}$ eV with the photons of 2.7 K CMBR. *Left panel* Gamma-rays and electrons; *Right panel* Electronic and muonic neutrinos and anti-neutrinos
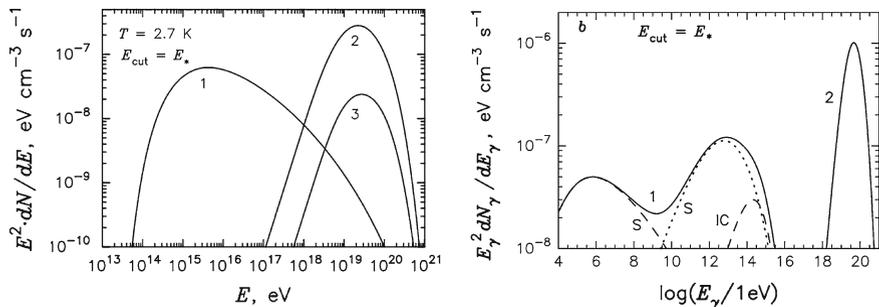
Fig. 24 *Left panel* Production spectra of electrons and positrons from direct Bethe-Heitler pair production (*curve 1*) and from photomeson interactions—*curve 2* (positrons) and *curve 3* (electrons). The parent proton distribution is a power-law with an exponential cutoff as described in the text. *Right panel* Synchrotron (S) and IC radiation (IC) radiation spectra of cooled electrons and positrons in a magnetic field of strength $B = 1\,\mu G$. *Curve 1* is the sum of the contributions of secondary electrons, *curve 2* is the direct gamma-ray spectrum from the decay of π^0-mesons

chrotron and inverse Compton losses. And only neutrinos can propagate through the intergalactic medium without distortion of their original energy spectra and direction.

At energies below the pion production threshold, protons continue to interact with background photons through Bethe-Heitler pair production. Although these interactions are not accompanied by emission of gamma rays, the secondary electrons quickly lose their energy via inverse Compton scattering and synchrotron radiation, and in this way result in gamma-rays. In this regard, the inverse Bethe-Heitler pair production often operates as an indirect process of gamma-ray emission. The broad-band gamma-ray spectra produced at interactions of ultrarelativistic protons with 2.7 K CMBR are shown in Fig. 24. It is assumed that protons are distributed as $N_p(E_p) = A E_p^{-2} \exp(-E_p/E_0)$ with the cutoff energy $E_0 = E_* = m_p^2 c^4/4kT = 3 \times 10^{20}$ eV (E_* corresponds to the energy of protons when π-mesons are effectively produced at interactions with 2.7 K CMBR). The production spectrum of pairs drops rapidly below $E_e' \sim 10^{15}$ eV, however this spectral energy distribution is modified by energy losses, to give a steady-state pair distribution $N_e \propto E_e^{-2}$. For the adopted magnetic field $B = 1\,\mu G$, the synchrotron radiation spectrum of cooled electrons has two maximums around 1 MeV and 10 TeV corresponding to the Bethe-Heitler and photomeson electrons, respectively. For the assumed magnetic field the contribution of IC gamma-rays is important only above 100 TeV.

4 SNRs and Origin of Galactic Cosmic Rays

SNRs are believed to be the major contributors to the locally measured cosmic ray (CR) flux up to the so-called "knee", a distinct spectral feature around 10^{15} eV. The main (phenomenological) argument in favour of this hypothesis is the CR production

rate in the Galaxy, $\dot{W}_{CR} \approx (0.3-1) \times 10^{41}$ erg/s. It can be supported by SNRs if approximately 10 % of the kinetic energy of galactic SN explosions is released in CRs (see e.g. Refs. [96, 139, 140]). The second argument has more *theoretical* background; it is linked to the potential of the so-called *diffusive shock acceleration* (DSA) mechanism which may convert with a high efficiency the available kinetic energy of bulk motion to relativistic particles.

Yet, despite the recent advances in the field, the SNR paradigm of the origin of galactic CRs remains a *hypothesis*. Moreover, it is not free of 'nasty problems' (see e.g. Ref. [152]). The direct measurements of cosmic rays are very important, but they alone hardly can address all outstanding issues, in particular the ones related to the localization and identification of the sites of particle accelerators. It is believed that gamma-ray astronomy should play a key role in the solution of this long-standing problem. Over the last decade, the space- and ground-based gamma-ray observations have significantly contributed to the understanding of particle accelerators in our Galaxy, and thus have proved the early predictions on the potential of the field in this regard.

4.1 Gamma-Ray Signatures of SNRs

The effective acceleration of cosmic ray protons and nuclei in supernova remnants makes them potentially detectable sources of GeV and TeV gamma-rays resulting from hadronic interactions through the production and decay of π^0-mesons. Thus, a straightforward test of acceleration of cosmic rays in SNRs would be detection of hadronic gamma-rays—directly from young remnants [124] and/or from dense clouds overtaken by the expanding shells [63].

The theory of DSA allows definite observational predictions. In particular, the distinct feature of shock acceleration in the non-linear regime is the concave shape of the particle energy distribution. At low (GeV) energies it is relatively steep with differential spectral index larger than 2, but at highest energies the spectrum becomes very hard. In the case of strongly modified shocks, the proton spectrum, just before the high energy cutoff, can be as hard as $E^{-1.5}$ (see Ref. [190]). Generally, these features are reflected in the spectrum of secondary gamma-rays [95, 131, 195, 254]. However, the energy-dependent *propagation effects* may introduce significant modifications in the proton spectra, in particular in the dense regions where the major fraction of gamma-rays is produced. This concerns massive molecular clouds located outside the mid-age SNRs [130, 137], as well as possible dense compact condensations inside young SNR shells [254]. Therefore, the observed gamma-ray spectra could significantly deviate from the acceleration spectrum of protons.

Presently seven shell type SNRs—Cas A [19, 28, 73], Tycho [16], SN 1006 [20], RX J1713.7-3946 [64, 132], RX J0852-4622 [33], RCW 86 [45], and G353.6-0.7 [14] are identified as TeV gamma-ray emitters. Remarkably, while the first six sources are well established young SNRs, the object G353.6-0.7 is the first SNR discovered

serendipitously in TeV gamma-ray, and only later confirmed by radio and X-ray observations [14].

While TeV gamma-rays from young SNRs prove the effective acceleration of CRs to energies up to 100 TeV, the relative contributions of accelerated protons and electrons to the gamma-ray production remain unknown. The problem is that the ratio of gamma-ray produced by accelerated protons interacting with the surrounding gas and by ultra-relativistic electrons upscattering the 2.7 K MBR, is very sensitive to generally unknown parameters, in particular to the gas density and the magnetic field of the ambient medium. The efficiency of the inverse Compton (IC) scattering is especially high at TeV energies (up to $E_e \sim 100$ TeV, it proceeds in the Thomson regime, and therefore the cooling time $t_{\rm cool}^{\rm IC} \propto 1/E_e \propto 1/E_\gamma^{1/2}$). For example, the typical production times of an 1 TeV photon by an electron and a proton of the same characteristic energy of about 20 TeV, are $\approx 5 \times 10^4$ year and $5 \times 10^7 (n/1\,{\rm cm}^{-3})^{-1}$ year, respectively (see e.g. Ref. [52]). Correspondingly, at 1 TeV the ratio of production rates of IC gamma-rays to π^0-decay gamma-rays, is approximately $10^3 (W_e/W_p)(n/1\,{\rm cm}^{-3})^{-1}$, where W_e and W_e are the total energies in 20 TeV electrons and protons, respectively. Thus even for a very small electron-to-proton ratio (at the stage of acceleration), $e/p = 10^{-3}$, the contribution of the IC component will dominate over the π^0 decay gamma-rays (in the shell with a typical gas density $n \leq 1\,{\rm cm}^{-3}$), unless the magnetic field in the shell significantly exceeds $10\,\mu$G. In this case, the accelerated electrons are cooled predominantly via synchrotron radiation, thus only a small fraction, $w_{MBR}/w_B \approx 0.1(B/10\,\mu{\rm G})^{-2}$, is released in IC gamma-rays. Alternatively, the proton-to-electron acceleration ratio should exceed 10^3 which, in principle, cannot be excluded given the uncertainty in one of the key aspects of DSA related to the so-called 'injection problem' (see Ref. [190]).

Young SNRs are less prominent at low energies. Only Cas A [7], Tycho [141], RX J1713.7-3946 [12], and RX J0852-4622 [228] have been detected by *Fermi*, and all of them are characterized by modest GeV gamma-ray fluxes. On the other hand, strong GeV gamma-ray emission has been reported from a number of mid-age SNRs (see Fig. 25), in particular from W28 [6, 142], IC 433 [9, 234] and W44 [8, 143] Remarkably, W28 [44] and IC 433 [15, 72] have been reported as VHE gamma-ray emitters as well. In this regard one should note also that for any reasonable model parameters, the multi-TeV particles cannot be confined in the shell of a mid-age remnant. Therefore the gamma-ray emission is likely to be the result of cosmic rays which already have left the remnants and interact with the nearby dense gaseous complexes [56]. In the case of mid-age SNRs we should expect both GeV and TeV gamma-rays from such regions [137, 237]. The detection of GeV gamma-ray emission by *Fermi* LAT from the molecular cloud complex that surrounds the supernova remnant W44 is a strong evidence of realization such a scenario [239].

One should note also that while for any reasonable model parameters the TeV particles cannot be confined in the shell of a mid-age remnant, the latter can still contain low-energy particles. Therefore one we should expect MeV/GeV gamma-rays not only from the regions outside the remnant, but also from the shell itself. Generally, in the case of energy-dependent escape, the proton spectrum should be
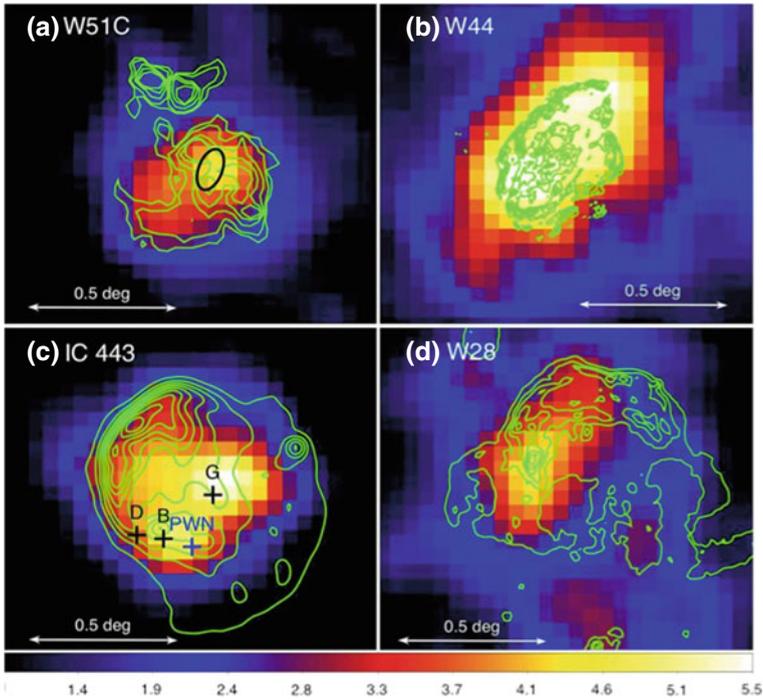
Fig. 25 2–10 GeV gamma-ray images of four mid-age SNRs reported by *Fermi* LAT collaboration: **a** W51C; **b** W44; **c** IC443; and **d** W28. Superposed are the contours from the VLA radio maps (from Ref. [240])

modified, namely become steeper. This seems to be the case of W44. The gamma-ray spectrum of 50 MeV–10 GeV gamma-rays from the shell of W44 reported by the *AGILE* team hardly can be explained by bremsstrahlung or inverse Compton scattering of electrons, but it agrees quite well with gamma-rays from *pp* interactions for a steep, E^{-3} type proton spectrum [143].

The detection and identification of gamma-ray components produced in the shell and in nearby clouds provide an important insight into the understanding of *acceleration* and *escape* mechanisms of cosmic rays in SNRs. Although GeV gamma-rays tell us only about low-energy particles, they in fact serve unique carriers of information about the sites of "ancient" PeVatrons. It is important to note in this regard that no TeV gamma-rays are expected from the shells of old and mid-age SNRs. On the other hand, this belief should not prevent us from future searches for TeV gamma-rays from the shells of old SNRs. Although so far the DSA mechanism in general, and in the context of CR acceleration in SNRs in particular, seems to work without a major problem, yet it is worth to keep in mind that DSA remains a theoretical paradigm the predictions of which should not be overestimated as long as it concerns the interpretation of gamma-rays from SNRs.

Finally, several galactic TeV gamma-ray sources spatially coincide with the so-called composite SNRs, objects with characteristic features of both standard shell-type SNRs and pulsar wind nebulae. At least in one case, the association of a TeV source with the composite SNR G0.9+0.1 is clearly established [37]. The point-like gamma-ray image with an angular size less than 1.3′ indicates that TeV gamma-rays originate in the plerionic core of the remnant, rather than in the 4′ radius shell.

5 TeV Emission of Young SNRs

Both the particle acceleration and radiation processes are very sensitive to the initial conditions of SN explosions, as well as to the parameters characterizing the surrounding regions. This explains the considerable diversity in the multi wavelength radiation properties of young SNRs reported as TeV gamma-ray sources. So far, the most unusual representative of this class of objects is RX J1713.7-3946.

5.1 RX J1713.7-3946: An Atypical SNR

While the synchrotron radio emission and thermal X-rays are two distinct components of shell type SNRs in general, RX J1713.7-3946 shows weak radio emission, and no thermal X-radiation at all. On the other hand, this object is a powerful non thermal X-ray and TeV gamma-ray emitter. The X- and VHE gamma-ray images of this remnant are shown in Fig. 26a. The overall shell type structure and its correlation with the non thermal X-ray image is visible, although the correlation is less evident on smaller scales [217].

The broad-band gamma-ray spectrum of the entire remnant based on the *Fermi* LAT [12] and HESS [43] measurements is shown in Fig. 26b. It extends over five decades in energy, from 1 GeV to 100 TeV. The theoretical curves correspond to the leptonic (IC) and hadronic (π^0-decay) model predictions calculated within a simple one-zone model, assuming that the GeV and TeV gamma-ray regions fully overlap. It is seen that although both hadronic and leptonic models do satisfactorily explain the spectral points above 1 TeV, the one-zone leptonic model fails to explain the GeV fluxes reported by *Fermi*. The problem here is related to the cooling break in the electron spectrum, and correspondingly to the position of the Compton peak which in the spectral energy distribution (SED) appears above 1 TeV [229]. Thus, the reduction of the break energy down to 200 GeV could in principle solve the problem. Since the magnetic field in this model cannot significantly exceed $10\,\mu G$, the only possibility to shift the Compton peak to sub-TeV energies is to assume that the remnant is much older than 10^3 year, which however is not supported by multi wavelength data. On the other hand, the constraints on the strength of the magnetic field are less robust, if the IC and synchrotron components of radiation are formed in different zones [57]. Such a scenario in young SNRs is not only possible, but, in fact,
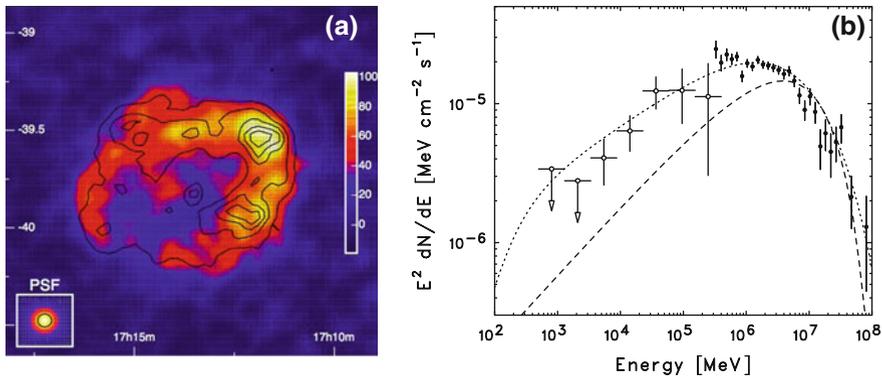
Fig. 26 Spatial and spectral characteristics of RX J1713.7-3946 at very high energies. **a** (*left panel*) The X- and VHE gamma-ray images of RX J1713.7-3946 obtained with the ASCA and HESS telescope array, respectively. **b** (*right panel*) The spectral energy distribution of RX J1713.7-3946 based on the *Fermi* [12] and *HESS* [43] data. The theoretical "hadronic" and "leptonic" gamma-ray spectra calculated within a simple one-zone model are from Ref. [229]. The IC curve is obtained for the electron spectrum derived from the synchrotron X-ray flux assuming magnetic field of about 14 μG. The π^0-decay gamma-ray spectrum corresponds to the spectrum of protons with the power-law index $\Gamma = 1.7$ and exponential cutoff at 25 TeV (from Ref. [43])

can be naturally realized in the forward and inverse shocks in which the magnetic fields are essentially different [254].

The agreement of the spectrum of hadronic gamma-rays with the measurements over the entire GeV to TeV region can achieved assuming a very hard spectrum of protons with power-law index 1.7 and an exponential cutoff at 25 TeV. Although this spectrum is harder than the nominal E^{-2} type acceleration spectrum predicted by the models applied to this source [95, 131, 195, 254], such a hard proton distribution cannot be excluded. Moreover, in the case of inhomogeneous distribution of gas in the shell, the proton spectrum in the densest regions, where the major fraction of gamma-rays is produced, can significantly deviate, due to the propagation effects, from the acceleration spectrum [189, 254] (see below).

The total energetics in accelerated electrons and protons in the relevant leptonic and hadronic models of gamma-rays can be estimated by invoking minimum model parameters. For the given distance to the source of about 1 kpc, the required budget in electrons is determined only by the reported gamma-ray fluxes, $W_e \simeq 3 \times 10^{47}$ erg, while the total energy budget of protons in hadronic models depends on the ambient gas density, $W_p \simeq 10^{50}(n/1\,\mathrm{cm}^{-3})^{-1}$ erg [229]. The lack of the thermal X-ray emission from this source requires gas density as low as 0.1 cm^{-3} which makes the realization of *standard* hadronic scenarios rather problematic [131, 163, 254]. Still, even in the case of very low gas density of the shell, the contribution of hadronic gamma-rays can be significant, if accelerated protons interact with the dense cores of molecular clouds embedded in the shell [136]. The corresponding gamma-emission may exceed the gamma-emission from the shell by a factor that is the ratio of the

total mass of clouds to the mass swept up by the forward shock. Thus, in the scenario with very low gas density of the shell, in addition to the IC radiation by electrons, one may expect significant contribution of hadronic gamma-rays produced in dense gas condensations [254], provided that all particles freely enter the dense clouds. This however could not be the case, especially for the low energy particles. Because of slow diffusion, the penetration of low energy particles into the dense cores of these condensations can take longer than the age of the SNR. Correspondingly, the low energy gamma-ray emission can be suppressed. This effect offers a possible explanation [136, 159, 254] for the hard gamma-ray spectrum below 100 GeV reported by the *Fermi* collaboration from RX J1713.7-3946 [12].

The production of gamma-rays at *pp* interactions in *dense* gas condensations embedded in the *low density* shell is an interesting scenario which keeps the hadronic origin of radiation as a viable option with several attractive features. The increase of photon statistic in future observations with *Fermi* and HESS should help, but hardly could be sufficient to distinguish unambiguously the contributions of leptonic and hadronic interactions to different bands of gamma-ray spectrum. In this regard, CTA has a great role to play. This concerns, first of all, to the precise measurements of the energy spectrum below 1 TeV down to tens of GeV and above 10 Tev up to 100 TeV. The morphological studies provide an independent and complementary information about the radiation mechanism. The low magnetic field, which is a key element of any IC model, allows the multi-TeV electrons to propagate to large distances, and thus to fill a quite large volume. Because of homogeneous distribution of the target photon fields, the spatial distribution of resulting IC gamma-rays appears quite broad. The hadronic model predicts narrower and sharper spatial distribution, mainly due to the enhanced emission in the compressed region of the shock, as it is seen in Fig. 27. However, because of limited angular resolution of gamma-ray telescopes, it is hard to distinguish between the radial distributions predicted by two models. It is demonstrated in Fig. 27 where the radial profiles are smoothed with a typical for the current Cherenkov telescope arrays point spread function of $\delta\psi = 3$ arcmin. Both smoothed profiles reasonably agree with the reported angular distribution of TeV gamma-rays. For decisive conclusions about the sharpness of the shell emission, and therefore on the nature of parent particles, the detector's angular resolution should be around 1–2 arcmin. This seems to be at the edge of capability of the atmospheric Cherenkov imaging technique [156], but its practical realization by CTA would greatly contribute to the understanding of the role of SNRs in the origin of galactic CRs.

5.2 SN1006, Tycho and Cas A

SN 1006 was the first SNR from which the non thermal component of X-rays had been unambiguously identified with synchrotron radiation of multi-TeV electrons. Like the X-ray image, gamma-ray emission is concentrated in two extended regions in North-East and South-West (see Fig. 28a). In this object we deal with a quite homogeneous
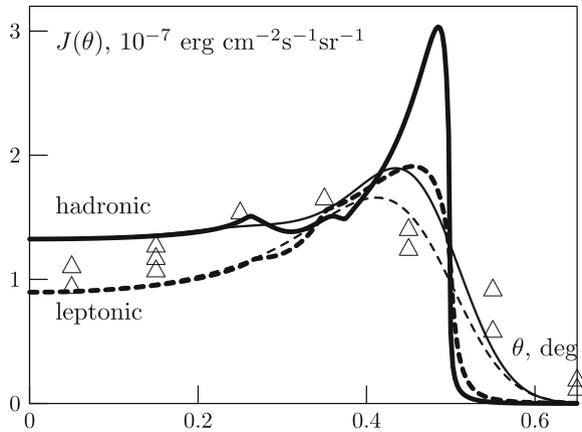
Fig. 27 Radial Profiles of 1 TeV gamma-rays calculated for the hadronic and electronic scenarios [254] in the uniform medium (*solid*) and for the leptonic scenario with the unmodified forward shock (*dashed*). The profiles smoothed with a Gaussian point spread function with $\sigma = 0.05°$ are also shown (*thin lines*). The *triangles* correspond to the azimuthally averaged TeV gamma-ray radial profile as observed by HESS
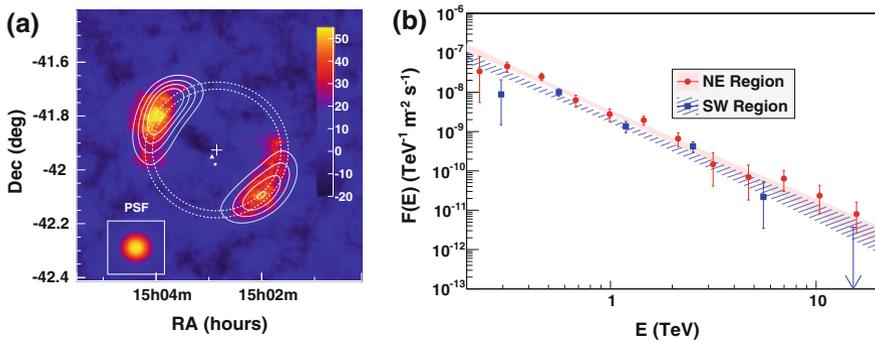
Fig. 28 The spatial and spectral distributions of VHE gamma-rays in SN 1006. **a** (*right panel*) TeV gamma-ray image [20]. The *white contours* correspond to a constant X-ray intensity as derived from the XMM-Newton flux map and smoothed to the HESS point spread function. **b** Differential energy spectra extracted from the NE and SW regions. The *shaded bands* correspond to the range of the power-law fit

environment, therefore such a morphology indicates the sites of concentration of highest energy particles. It has been argued that a possible reason for the bipolar picture could be the dependence of efficiency of injection of supra-thermal particles on the angle between the ambient magnetic field and the shock normal [94].

The gamma-ray emission of SN 1006 is very weak; the integral flux above 1 TeV is of the order of 1 % of the Crab flux. The energy spectra of the NE and SW regions are similar and compatible with a power law distribution, $F(E) \propto E^{-\Gamma}$, $\Gamma \approx 2.3$

(see Fig. 28b), i.e. significantly different from the energy spectrum of RX J1713.7-3946 which is flatter around 1 TeV ($\Gamma \simeq 2$) and steeper above 10 TeV ($\Gamma \simeq 3$). This might imply that in this case we have stronger evidence of hadronic origin of gamma-rays than for RX J1713.7-3946. Indeed, the relatively hard power-law spectrum of SN 1006 extending over two decades, from 0.2 to 20 TeV, significantly deviates from general predictions of IC models. The so called *one-zone* model which assumes a common region of production of the IC and synchrotron components of radiation, satisfactorily explains the gamma-ray flux around 1 TeV (for a reasonable magnetic field of order of $30\,\mu G$), but fails to explain the hard power-law spectrum at multi-TeV energies [20]. However, the one zone-model is a simplification [57] which, in fact, hardly can be realized in SNRs. The difference of the magnetic field in the upstream and downstream regions results in the positional shift of production regions of synchrotron X-rays and IC gamma-rays. While X-rays are predominantly produced in the downstream region, IC gamma-rays are contributed by both the downstream and upstream regions, with a strong dominance of the latter at multi-TeV energies [253]. This leads to a shift of the overall gamma-ray spectrum towards higher energies.

Alternatively, the existing VHE gamma-ray data can be explained by interactions of accelerated protons with power-law distribution $E^{-2.3}$ between 1 and 100 TeV. Below 1 TeV, the proton spectrum should become flatter than E^{-2}. This is a rather robust conclusion opposed by the available energy budget, given that for the ambient gas density $n \sim 0.1\,cm^{-3}$, the required acceleration efficiency of protons in the energy interval 1–100 TeV already exceeds 20 % of the total mechanical energy of the SN explosion, $E_{SN} = 1.4 \times 10^{51}$ erg [20]. The proton spectrum can be described, for example, by a power low $E^{-2.3}$ with a break below 1 TeV, and a high energy cutoff at $E \gg 100$ TeV. This is an interesting option because it requires effective acceleration of protons to extremely high energies, formally to 1 PeV. However, there is an alternative option when the proton spectrum is described by a flat E^{-2} type spectrum with an exponential cutoff at $E \simeq 80$ TeV [20]. The extension of gamma-ray measurements by CTA to energies well beyond 10 TeV should allow us to distinguish between these two realizations, and thus to answer to a principle question as whether SN 1006 acts as a PeVatron.

Gamma-ray signals have been reported, both at high and very high energies (see Fig. 29), from another prominent object, the Type Ia supernova remnant Tycho [16, 141]. While the *Fermi* data between 0.4 and 100 GeV can be described by a power-law with a photon index $\Gamma_{GeV} \approx 2.3$, the spectrum reported by the VERITAS collaboration between 1 and 10 TeV is somewhat harder, $\Gamma_{TeV} \approx 2$. Within the uncertainties of the ambient gas density and the distance to the source, the reported fluxes at both GeV and TeV energies agree with the early phenomenological predictions [124], as well as with the recent theoretical studies [196, 244] of production of hadronic gamma-rays in Tycho. The flat gamma-ray spectrum up to 10 TeV implies that the corresponding spectrum of parent protons continues without a significant steepening or a cutoff to at least several hundred TeV [165]. The extension of the flat gamma-ray spectrum just by a factor of two or three beyond 10 TeV, would be sufficient to claim the source as a PeVatron. Such observations also would robustly

Fig. 29 The reported gamma-ray emission of Tycho from GeV [141] to TeV [16] energies. The expected contributions from the pp-interactions, (*dot-dashed line*), relativistic bremsstrahlung (*dot-dot-dashed*) and Inverse Compton scattering (on three different radiation fields) are also shown (from Ref. [196])
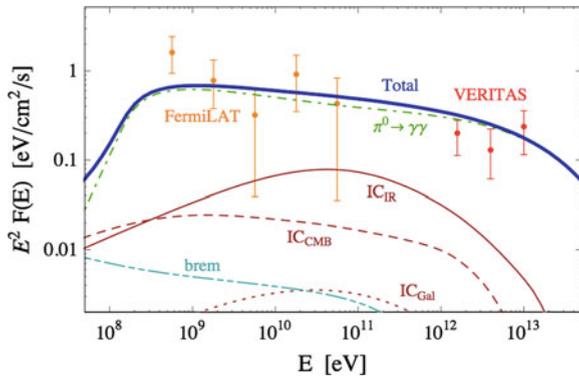
exclude the IC origin of radiation. Note that although the present data give a preference to hadronic models [196] (see Fig. 29), yet a specifically designed two-zone leptonic model cannot be discarded as a possible scenario for explanation of the current gamma-ray data from Tycho [80].

The shell type supernova remnant Cas A is one of the best studied non thermal objects in our Galaxy. Although some of its general features are common for young SNRs, Cas A is a rather unique representative of the remnants of recent supernovae explosions. Its synchrotron emission spans from radio to hard X-rays indicating the presence of relativistic electrons from sub-GeV to multi-TeV energies. For any reasonable assumption on the nebular magnetic field, this object contains enormous total energy in the form of relativistic electrons, $W_e \simeq 3 \times 10^{48}$ erg [85]. The average rate of accumulation of this energy over a short time period, $t_{acc} \leq t_{age} \sim 300$ year, is even more striking, $\dot{W}_e = W_e/t_{acc} \approx 3 \times 10^{38}$ erg/s. It is larger, by at least an order of magnitude, than the electron production rate in any other supernova remnant. On the other hand, the content of protons in this object could be relatively modest. As discussed below, the amount of relativistic protons and nuclei is constrained by gamma-ray fluxes detected in the GeV and TeV energy bands, $W_p \leq 3 \times 10^{49}(n/10\,\mathrm{cm}^{-3})^{-1}$ erg [7]. This constitutes less than 2 % of the total explosion energy, if gamma-rays are produced in the reverse shock where the plasma could be quite dense, $n \geq 10\,\mathrm{cm}^{-3}$. In this case, the ratio of relativistic protons to electrons in Cas A is less than 10, i.e. an order of magnitude below the level observed in cosmic rays. This is in sharp contrast to the hadronic models of γ-radiation of other SNRs, e.g. SN 1006 and RX J1713.7-3946, which require $p/e \geq 10^3$. However, if the reported GeV and TeV gamma-ray fluxes are produced at hadronic interactions in the forward shock, which propagates through a low density circumstellar medium, the total energy in accelerated protons can exceed 10^{50} erg.

Another unique feature of this object is the reported gamma-ray [160] and X-ray [213] emission lines which associate with ^{44}Ti. This radiation component provides direct information about the ejected mass of radioactive titanium-44, $M_{44\mathrm{Ti}} \approx 2 \times 10^{-4} M_\odot$. It has been recently suggested [255] that the unusually

high content of relativistic electrons in Cas A could have a link to the ejection of large amount of radioactive material, first of all ^{44}Ti and ^{56}Ni. The decay products of these nuclei provide a vast pool of supra-thermal positrons and electrons which can be further accelerated, by both the reverse and forward shocks, to multi-TeV energies.

The flux measured by *Fermi* in the 0.3–30 GeV interval is flat with a slope $\Gamma = 2.0 - 2.2$ [7]. The VHE gamma-ray signal from Cas A has been discovered in the 1–10 TeV energy interval by the *HEGRA* collaboration [28], and confirmed by the *MAGIC* [73] and *VERITAS* [19] collaborations. The gamma-ray spectrum is compatible with power low, $dN/dE \propto E^{-\Gamma}$ with a photon index $\Gamma = 2.6$ and total flux above 1 TeV $F_\gamma(\geq 1\,\text{TeV}) = 7.7 \times 10^{-13}$ ph/cm^2 s [19]. The GeV and TeV electrons which are responsible for the broad-band synchrotron radiation of Cas A, inevitably produce also gamma-rays—through bremsstrahlung and inverse Compton scattering. The calculations of gamma-ray fluxes based on the radio data are straightforward, but strongly depend on distributions of the gas and magnetic field in the nebula. Since the average magnetic field in Cas A cannot be significantly less than 0.3 mG (otherwise the contribution of bremsstrahlung would lead to overproduction of low-energy gamma-rays [85]), the TeV gamma-ray emission of IC origin its expected to be very low, unless one invokes regions with low magnetic field but yet with adequate conditions for effective acceleration of electrons to multi-TeV energies. Although at first glance this sounds a rather superficial assumption, the regions with very low magnetic field in SNRs cannot be excluded. Moreover, in the case of Cas A this can be realized in a quite natural way, through the operation of the reverse shock.

On the other hand, the overall GeV–TeV gamma-ray spectrum can be readily explained by interactions of accelerated protons and nuclei with the ambient gas, assuming a power-law spectrum of protons extending to 100 TeV with a power-law index $\alpha = 2.3$, or by a harder spectrum with $\alpha = 2.1$ but suffering an exponential cutoff at 10 TeV [7] (see Fig. 30). This ambiguity, which is a result of large statistical uncertainties of TeV gamma-ray fluxes, leaves open the question whether Cas A accelerate particles to PeV energies. As discussed before, this is true also for other SNRs. Meanwhile, without a clear answer to this question, we cannot be sure that SNRs are the major contributors to the galactic cosmic rays up to the "knee" around 1 PeV. The "hunt" for galactic PeVatrons continues.

5.3 Radiation Signatures of Proton PeVatrons

The most straightforward search for galactic PeVatrons can be conducted by gamma-ray detectors designed for operation in the energy regime between 10 and 100 TeV. In SNR shocks with relatively low acceleration rate, the synchrotron losses prevent acceleration of electrons to energies beyond 100 TeV. Also, at such high energies the contribution of the IC component is suppressed because of the Klein-Nishina effect. Therefore, the contribution of the IC gamma-rays to the radiation above 10 TeV is
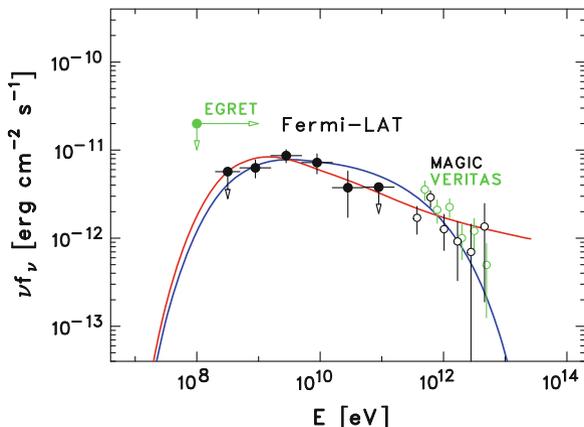
Fig. 30 The differential energy spectra of gamma-rays reported from Cas A by the *Fermi*, *MAGIC* and *VERITAS* collaborations (from [7]). Two *curves* correspond to calculations of theoretical gamma-ray spectra from *pp* interactions assuming for protons (i) a single power low distribution with an index $\alpha = 2.3$ (*red line*) (ii) a power-law distribution with $\alpha = 2.1$ an exponential cutoff at $E_0 = 10$ TeV

expected to gradually fade out. Thus, in the case of detection of gamma-rays up to 100 TeV, the hadronic origin of radiation would be unambiguously established.

Figure 31 shows the X-ray and gamma-ray luminosities from a 1000 year old proton PeVatron calculated for three different distributions of accelerated protons. Both radiation components are initiated by interactions of accelerated protons for the ambient gas density $n = 1 \, \text{cm}^{-3}$ and magnetic field $B = 300 \, \mu\text{G}$. While gamma-rays
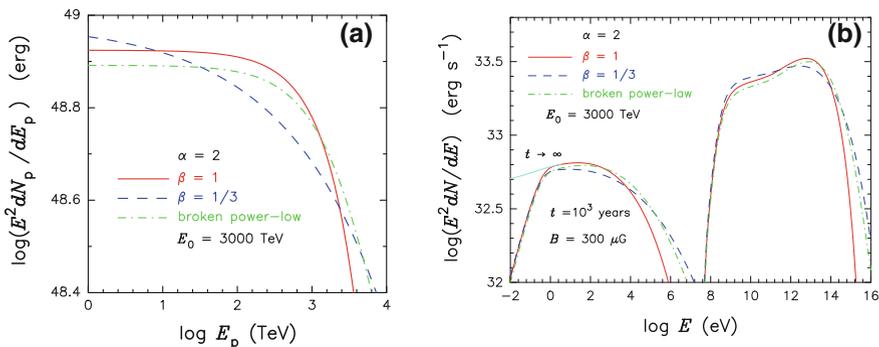
Fig. 31 The broad-bad radiation of a PeVatron initiated by interactions of protons with the ambient gas. **a** (*left panel*) Three different distributions of protons: 1,2—"power-law with an exponential cutoff", $E^{-\alpha} \exp(-E/E_0)^\beta$ with $\alpha = 2$, $E_0 = 3$ PeV, and $\beta = 1$ (*solid curve*), $\beta = 1/3$ (*dashed curve*), and 3—"broken power-law" when the spectral index is changed at $E = 1$ PeV from $\alpha = 2$ to $\alpha = 3$. **b** (*right panel*) Luminosities of brand-band emission produced at *pp* interactions. The gas density $n = 1 \, \text{cm}^{-3}$, magnetic field $B = 300 \, \mu\text{G}$, and the age of the source $t = 10^3$ years

arise directly from decays of π^0-mesons, X-rays are result of synchrotron radiation of secondary electrons, the products of $\pi^{\pm}$-decays. The lifetime of electrons producing X-rays, $t_{\text{synch}} \simeq 1.5 B_{\text{mG}}^{-3/2} (E_{\text{X}}/1\,\text{keV})^{-1/2}$ year, is very short (≤ 50 year) compared to the age of the source. Therefore, X-rays could be treated as a "prompt" radiation emitted simultaneously with gamma-rays.

The X- and gamma-ray fluxes depend on the present content of total energy of accelerated protons accumulated in the source, and on the density of the ambient matter. Approximately the same fraction of energy of the parent protons is transferred to secondary electrons and gamma-rays. However, since the energy of sub-TeV electrons is not radiated away effectively, the direct (π^0-decay) gamma-ray luminosity exceeds the synchrotron luminosity. The L_{X}/L_{γ} ratio depends on the proton spectrum as well as on the particle injection history; typically it does not exceed 0.2–0.3.

The spectrum of highest energy gamma-rays contains an important information about the shape of the proton spectrum around the cutoff E_0 which is crucial for identification of acceleration mechanisms in SNRs, as well as for understanding of the role of different processes responsible for the formation of the knee in the CR spectrum. The X- and gamma-ray luminosities in Fig. 31 are calculated for a proton accelerator operating during 10^3 year with a constant rate $L_{\text{p}} = 10^{39}$ erg/s; the total energy in protons is $W_{\text{p}} = L_{\text{p}} \cdot T \simeq 3 \times 10^{49}$ erg. To estimate the X-and gamma-ray energy fluxes (in units of erg/cm^2 s) from an arbitrary PeVatron, one should multiply the luminosities in Fig. 31b by the factor $\kappa \approx 10^{-44}(nW_{\text{p}}/3 \times 10^{49}\text{erg/cm}^3)(d/1\,\text{kpc})^{-2}$. Thus, all galactic PeVatrons up to distances of 10 kpc and with $nW_{\text{p}} \geq 10^{49}$ erg/cm^3, can be probed by CTA, the sensitivity of which above 10 TeV is expected to be better than 10^{-13} erg/cm^2 s [23]. However, because of tiny gamma-ray fluxes, detailed spectroscopic measurements in the cutoff region could be possible only for powerful and/or nearby objects. Note that the extension of the proton spectrum to 1 PeV is crucial for effective production of neutrinos in the 10–100 TeV range, the most optimal energy interval for the TeV neutrino detectors like IceCube or KM3NeT [243]. However, the sensitivity of neutrino detectors is quite limited, and even the brightest in gamma-ray SNRs can be only marginally detected by these instruments.

In such circumstances, the search for the PeVatrons via synchrotron X-radiation of secondary ($\pi^{\pm}$-decay) electrons has been suggested as an alternative tool [52]. In principle, *Chandra* and *XMM-Newton* have sufficient sensitivity to perform such studies. However, the limited effective energy domain (≤ 10 keV) of these instruments is not optimal for detection of synchrotron radiation of secondary electrons. Indeed the major challenge of this method is the extraction of the "hadronic" component of X-rays from the synchrotron radiation of directly accelerated electrons. These two components can be separated if the magnetic field in the SNR exceeds 100 μG and the proton spectrum extends to 1 PeV. These two conditions are, in fact, connected since the acceleration of protons in SNRs to PeV energies is possible only at the presence of large magnetic fields. The second key condition for operation of SNRs as PeVatrons is the diffusion in the *Bohm limit*. In this case, the proton cutoff energy is proportional to the strength of the magnetic field. Thus, the corresponding energy in the spectrum of

secondary synchrotron radiation $h\nu \propto BE_0^2 \propto B^3$. On the other hand, the position of the cutoff of synchrotron radiation of directly accelerated electrons does not depend on the magnetic field and typically appears in the soft X-ray domain, $h\nu \leq 1\,\text{keV}$ (see e.g. [57]). Thus, if the spectrum of synchrotron radiation of secondary electrons extends well beyond 10 keV, the background caused by the synchrotron radiation of directly accelerated electrons is dramatically reduced.

The spectrum of the secondary synchrotron radiation in the cutoff region is smoother and broader than the gamma-ray spectrum around the cutoff. For the proton spectrum written in a general form $(dN/dE)_\text{p} \propto \exp[-(E/E_0)^{\beta_p}]$, the distributions of secondary gamma-rays and electrons in the cutoff region is $(dN/dE)_{\gamma/e} \propto \exp[-(E/E_{0,\gamma/e})^{\beta_{\gamma/e}}]$, with $\beta_{\gamma/e} \approx 0.5\beta_p$ [165]. The spectrum of the synchrotron radiation in the cutoff region is described by a function proportional to $\exp[-(\varepsilon/\varepsilon_0)^{\beta_s}]$, with $\beta_s = \beta_e/(2 + \beta_e)$ [253]. For example, in the Bohm diffusion regime, $\beta_p = 1$, thus in the cutoff region the distribution of gamma-rays is proportional to $\exp[-(E_\gamma/E_{\gamma/e})^{1/2}]$, while the spectrum of the secondary synchrotron radiation behaves as $\exp[-(\varepsilon/\varepsilon_0)^{1/5}]$. This is an important feature (see Fig. 31), which should allow detailed studies of the proton spectra around and beyond the cutoff with the planned hard X-ray imagers of the NuSTAR and ASTRO-H X-ray satellites.

5.4 Expectations from Future Studies

As discussed above, the gamma-ray emission from several famous representatives of young SNRs can be explained by interactions of shock accelerated protons and nuclei. However, this cannot be taken as a proof of the major contribution of SNRs to the production of galactic cosmic rays because the data can be explained also by the competing leptonic processes. Moreover, we do not have a single example which would be considered as an evidence of accelerations of protons up to 1 PeV. A natural reason for the deficit of such energetic particles could be their leakage from the shell [114, 254]. Indeed, the acceleration and confinement of multi-TeV particles in the remnants can last less than several hundred years after the explosion, so one should be lucky to "catch" a SNR while it is still accelerating particles to multi-TeV energies. This significantly constraints the number of SNRs emitting very high energy gamma-rays. Moreover, the sensitivity of current IACT arrays limits the detection of TeV gamma-rays by young SNRs located within 1–2 kpc distances. An increase of the 'detectability distance' by an order magnitude would be a great achievement, since it would make visible most of young SNRs in our Galaxy. The inner parts of the Galaxy with enhanced star formation rate are of special interest. Because of reduction of gamma-ray fluxes proportional to the distance squared, the increase of the distance by an order of magnitude formally demands an improvement of the sensitivity for point-like sources by two orders of magnitude. But, fortunately, in the case of young SNRs an order of magnitude improvement of the sensitivity would be sufficient. Indeed, the sensitivity of IACT arrays for extended sources is proportional

to the angular size of the source, $J_{\min} \propto \Psi$, as long as Ψ is larger than the angular resolution of the IACT arrays, $\delta\theta \approx 0.1°$. On the other hand, the angular sizes of young SNRs scales with the distance approximately as $\Psi \sim 1°(d/1\,\mathrm{kpc})^{-1}$. Thus, the expected improvement of the sensitivity of CTA by an order of magnitude compared to the current IACT arrays, will dramatically increase the number of potentially detectable SNRs. This should allow compelling population studies—a key issue for the proof of the SNR origin of galactic cosmic rays.

The acceleration of protons in SNRs to multi-TeV and PeV energies can be studied also in an indirect way using the gamma-ray "echos" of old, faded out accelerators. In certain environments, TeV gamma-rays can be detected even ten thousand years after the so-called Sedov phase when all particles already have left the remnant. Indeed, the particles, after they escape the SNR shell, interact with the surrounding atomic and molecular gas. Before being fully diffused away and integrated into the "sea" of galactic CRs, these particles produce gamma-rays the spectrum of which can significantly differ from both the radiation of the SNR shell and the diffuse galactic gamma-ray emission [56]. The massive molecular clouds located in the vicinity of the supernova remnant, provide dense targets for hadronic interactions, and thus dramatically increase the chances of tracing the run-away protons via the secondary gamma-rays. For parameters of a typical SNR at a distance of 1 kpc, a molecular cloud of mass $10^4 M_\odot$ can emit multi-TeV emission at a detectable level, if the cloud is located within 100 pc from the SNR [137]

The location of molecular clouds close to SNRs could be accidental, but in general there is a deep link between SNRs and MCs, especially in the star-forming regions [194]. Depending on the location of massive clouds, the time of particle injection into the interstellar medium, as well as on the diffusion coefficient, we might expect a broad variety of energy distributions of gamma-rays—from very hard spectra (much harder than the spectrum of the SNR itself) to very steep ones [56]. Correspondingly, the ratio of GeV to TeV gamma-ray fluxes can significantly vary from site to site.

One should note that the role of giant molecular clouds is not limited by SNR studies. These massive objects are intimately connected with the star formation regions that are strongly believed to be most probable sites of CR production (with or without SNRs) in our Galaxy. They serve also as a unique "barometers" for measurements of the energy density of cosmic rays in remote parts of the Galaxy [113]. It is generally believed that the local CR flux directly measured at the Earth, gives a correct estimate for the level of the "sea" of galactic CRs. However, one cannot exclude that the flux of local CRs could be dominated by a single or few local sources, especially given that the Solar system is located in a rather extraordinary region—inside active star formation complexes which constitute the so-called Gould Belt. The recent anomalies discovered in cosmic rays, such as very high content of positrons in the leptonic component of cosmic rays [24], or the significant differences between energy spectra of protons and alpha particles [25], tell us that the generally adopted picture of homogeneously distributed galactic CRs contributed by a single class of accelerators (SNRs) could be an oversimplification. In fact, we may deal with a diverse variety of cosmic ray accelerators. The fortunate location of giant

molecular clouds in the vicinity of particle accelerators provides us with a unique tool to probe cosmic rays in these environments.

Despite the remarkable advances of recent years, we do not have yet a robust evidence of acceleration of nucleonic component of cosmic rays in supernova remnants. It is expected that explorations of these objects with the major future ground-based gamma-ray detector, the CTA, with significantly improved performance regarding the minimum detectable flux, the angular resolution, and broader (almost five decades!) energy coverage, will remove, to a large extent, many current uncertainties and ambiguities concerning the origin of gamma-rays from SNRs. A breakthrough in this regard is especially expected in the unexplored so far energy interval above 30 TeV. Indeed, the contribution of the inverse Compton component of radiation at these energies is largely suppressed, thus there will be little doubt that we deal with gamma-rays of hadronic origin which would carry direct information about protons accelerated up to 1 PeV. Also, the best angular resolution of CTA (as good as 1–2 arcmin; J. Hinton, private communication) is expected at energies above 10 TeV. This is very important for the study of several key, but so far poorly understood issues related, for example to the escape of highest energy particles from the acceleration zone. In general, the discovery of the first cosmic *PeVatrons*, can be considered as one of the highest priority objective of CTA. The galactic *PeVatrons* are guaranteed sources of neutrinos in the 10–100 TeV range, the most optimal energy interval for their registration by the km^3 scale detectors like IceCube and KM3NeT. Unfortunately, because of the limited sensitivity, even the most promising young SNRs suspected to be the brightest neutrino sources in our Galaxy in the multi-TeV band, can be robustly detected only after 5–10 years of continuous exposure. Together with studies of ultra-high energy gamma-rays and neutrinos, the future hard X-ray imaging detectors like *NuSTAR* and *ASTRO-H*, should be able to conduct an effective search for the currently active *PeVatrons* through the synchrotron radiation of secondary electrons produced at hadronic interactions. Finally, the coverage of the GeV energy domain by the *Fermi* space telescope is very important, especially for the search of 'smoking guns' of former *PeVatrons*, as well as for exploration of the environments which harbour these mysterious cosmic ray "factories".

Of course, the search for *PeVatrons* should not necessarily be linked to SNRs. Some other classes of sources, e.g. superbubbles [151, 205] or remnants of gamma-ray bursts in our Galaxy [79] have been proposed as effective accelerators of particles to PeV energies and potential contributors to galactic cosmic rays. Moreover, the galactic plane survey by CTA might result in an exciting discovery of a new class of $E \gg 10$ TeV gamma-ray sources of unknown origin. The outcome could be a dramatic revision of the concept of origin of galactic cosmic rays without a major contribution from SNRs.
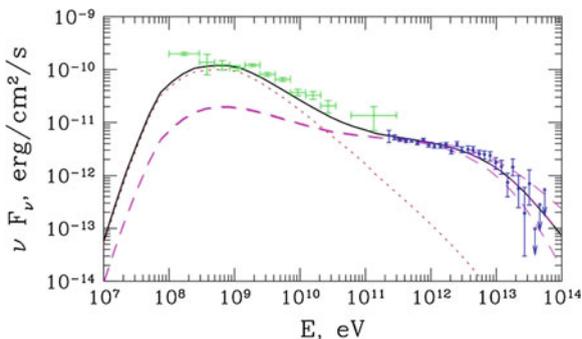
Fig. 32 Energy spectrum of the source in the GC based on the combined *Fermi* (*green points*) and H.E.S.S. (*blue points*) data. The *curves* correspond to the calculated contributions from a 10-day proton flare that occurred 300 years ago (*dashed line*) and from a constant proton source that switched on 10^4 years ago (*dotted line*). The *solid curve* is the superposition of these two components. The *upper* and *lower dashed lines* correspond to the 100 and 50 TeV cut-offs in the proton spectrum while the *solid curve* assumes $E_{cut} = 75$ TeV

6 Galactic Center

The Galactic Center (GC) is an extraordinary site that harbours many remarkable objects, in particular the compact radio source Sgr A*, a suspected super-massive black hole (SMBH) located at the dynamical center of the Galaxy. Over the last four decades this regions has been repeatedly observed by different gamma-ray detectors. GC contains a bright gamma-ray source with a broad-band spectrum that spans from 100 MeV [117] to 30 TeV [46]. Spatially it coincides with the position of Sgr A* (see Fig. 8). However the upper limit on the angular size of a few arcminutes is still large to exclude any other sources located within the central ≤ 10 pc region. Unlike the radio and X-ray bands, no variability has been observed both at GeV and TeV energies. This disfavours but still cannot discard Sgr A* as a possible gamma-ray source. Further constraints on the gamma-ray emission models come from the energy spectrum which has an interesting shape; at low energies it is hard with a photon index $\Gamma \approx 2.2$. Above 2 GeV the spectrum becomes stepper by $\Delta\Gamma \approx 0.5$ [117], but at TeV energies it hardens again with a photon index $\Gamma \simeq 2.1$ and an apparent break or a cutoff above 10 TeV (see Fig. 32).

6.1 Sgr A*

The temporal and spectral features of Sgr A* are unusual and, as a whole, very different from other compact galactic and extragalactic black hole candidates. This concerns, in particular, the extraordinary low luminosity of Sgr A*. In addition

to other important astrophysical implications, the low luminosity of Sgr A* has a dramatic effect on the visibility of the source in gamma-rays [48]. Because of internal photon–photon pair production, the high energy gamma-ray emission of black holes is strongly suppressed. However this is not the case of Sgr A* which makes it (together with the black hole in the radio galaxy M87) a unique laboratory for study of possible particle acceleration processes in the accretion flows near the event horizon of black holes. High energy gamma-rays from compact regions close to the SMBH can be produced in various ways.

- *Synchrotron and curvature radiation of protons*

Gamma-ray production at interactions of protons with magnetic field are effective only in the so-called extreme accelerators [61] where particles are accelerated at the maximum possible rate, $\dot{E} = eB$. Such an effective acceleration can be realized in the vicinity of SMBHs [181]. Even for an "ideal" combination of parameters allowing the most favourable acceleration/cooling regime, the characteristic energy of synchrotron radiation is limited to $\epsilon_{max} = (9/4)\alpha_f^{-1}m_p c^2 \simeq 0.3$ TeV. This implies that the proton-synchrotron radiation cannot explain the flux observed from the direction of GC since it would require a source moving with unrealistically large Doppler-factor, $D \geq 30$. More promising seems to be the curvature radiation [181] when the protons emit while moving along the regular magnetic field lines. For the SMBH in GC, $3 \times 10^6 M_\odot$, $\epsilon_{max} = 3E_p^3/2m^3 R \simeq 0.2(B/10^4)^{3/4}$ TeV. Formally, this allows extension of the spectrum to 10 TeV, provided that the magnetic field exceeds $B \simeq 10^6$ G. However, such a strong field would make the source opaque for $E \geq 1$ TeV gamma rays [48].

- *Photomeson interactions of protons*

Close to the event horizon of SMBH in GC protons can be accelerated to energies $E \sim 10^{18}$ eV [61, 181] and interact effectively with the radiation of the compact infrared source. The mean free path of protons through the photon field is estimated $\Lambda_{p\gamma} \sim (\sigma_{p\gamma} f n_{ph})^{-1} \simeq 10^{15}(R_{IR}/10^{13}$ cm$)^2$ cm. This means that approximately $R/\Lambda_{p\gamma} \sim 0.01$ fraction of the energy of protons is converted into secondary particles [48]. However, while neutrinos and neutrons, as well as gamma-rays of energies below 10^{12} eV escape freely the emission region, TeV gamma-rays effectively interact with the ambient photon and magnetic fields, and thus initiate IC and/or (depending on the strength of the B-field) synchrotron cascades. Gamma-rays produced in this way can explain the observed absolute TeV flux, if the acceleration power of 10^{18} eV protons is about 10^{37} erg/s. However the energy spectrum of the cascade gamma-rays does not agree with the observed overall GeV–TeV energy spectrum.

- *Proton–proton interactions*

While acceleration of protons to energies $E \sim 10^{18}$ eV requires extreme conditions which not always are satisfied, the boost to TeV energies can be more routinely realized. In these cases, interactions of protons with the ambient thermal gas become the main source of gamma-ray production. Protons can be accelerated to TeV energies also in the accretion disk, e.g. through strong shocks developed in the accretion flow. However, the efficiency of gamma-ray production, which is determined by the ratio

of accretion time to the *pp* cooling time, appears less than 10^{-4}. Correspondingly, the required *current* acceleration rate of high energy protons $L_p \geq 10^{39}$ erg/s is unacceptably high for GC. This makes unlikely the gamma-ray production via *pp* interactions.

- *Gamma rays of leptonic origin*

The models of gamma-ray emission associated with accelerated protons provide rather modest efficiencies of conversion of the energy of accelerated protons to gamma-rays. The radiative energy loss rate of electrons is much higher, and therefore the models that link gamma-rays to accelerated electrons provide more economic ways of gamma-ray production. Obviously, these electrons should be accelerated to at least $E_{max} \simeq 10$ TeV. This immediately constrains the strength of the chaotic component of the magnetic field; even under an extreme assumption that the acceleration proceeds at the maximum possible rate, $(dE/dt)_{acc} \simeq eB$, one gets $B \leq 10(E_{max}/10\,\mathrm{TeV})^{-2}$ G.

The requirement of particle acceleration at the maximum rate imposes strong restrictions on the geometry of magnetic field and possible acceleration mechanisms. In this regard, acceleration in ordered electric and magnetic fields, may provide certain advantages. Moreover, in the ordered field, the energy dissipation of electrons is reduced to the curvature radiation loses. This allows significant gain in the energy of accelerated electrons, up to $E_{e,max} \simeq 100(B/10\mathrm{G})^{1/4}$ TeV. In this scenario, electrons radiated through curvature and inverse Compton mechanisms. The curvature radiation peaks at $\epsilon_{curv} \simeq 200(E_e/10^{14}\,\mathrm{eV})^3$ MeV. The Compton scattering of same electrons leads to the second peak at much higher energies, $E_\gamma \sim E_e \simeq 100$ TeV (because the scattering proceeds in the Klein-Nishina limit). However, because of interactions with infrared photons, gamma-rays of energy exceeding 10 TeV can not freely escape the source. Synchrotron radiation and Compton scattering of the secondary (pair-produced) electrons lead to re-distribution of the initial gamma-ray spectrum. The results of calculations strongly depend on the maximum energy of electrons, therefore the interpretation of the observed gamma-ray spectrum from GeV to TeV energies requires fine-tuning of the model parameters. What concerns the acceleration rate of electrons, it could be quite reasonable—as low as 10^{37} erg/s. In this regard, the major problem of this interpretation is the compact production region and very short acceleration and radiative cooling times. They naturally imply fast variability which has not been yet detected in both GeV and TeV energy bands. Although this cannot yet rule out the gamma-ray production in the vicinity of SMBH, the lack of gamma-ray variability of the source is, of course, a serious argument against Sgr A* as the source of gamma-radiation of GC. However, Sgr A* remains a potential source to be (indirectly) responsible for the gamma-ray signal through interactions of runaway particles accelerated in Sgr A*, but later injected into the surrounding dense gas environment [49, 183, 184].
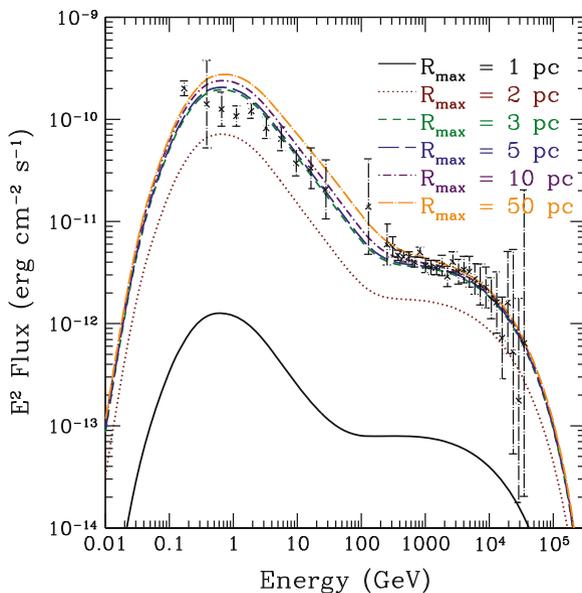
Fig. 33 Energy spectra of gamma-ray emission from GC. The *Fermi* and HESS data are shown together with calculations of gamma-rays from *pp* interactions within radial cones of various size up to 50 pc [183]. The flux falls off rapidly after 3 pc because the main contribution comes from the 1.2–3 pc circum-nuclear ring

6.2 Diffuse Gamma-Ray Emission from the Central 10 pc Region

In the scenario of acceleration of protons in Sgr A* and their subsequent injection into the interstellar medium, the total gamma-ray flux, as well as the spectral and angular characteristics of π^0-decay gamma-rays significantly depend on the proton injection rate and the regime of their diffusion. In particular, it has been realized [49] that for a certain combination of model parameters the energy spectrum of resulting gamma-rays can contain several spectral features related to the effects of proton propagation. The analysis of the combined *Fermi* LAT and HESS data show that the complex shape of the GeV–TeV radiation can be indeed naturally explained by the propagation effects of protons interacting with the dense gas within the central 10 pc region [117, 183]. The good agreement between the data and calculations are shown in Fig. 33 where the radial profile of the gas density has been carefully taken into account. The flat spectra in the segments of the proton spectrum around 1 GeV, and at TeV energies (below 10 TeV) have different explanations. While at GeV energies the protons are diffusively trapped, so they lose a large fraction of their energy before they leave the dense 3 pc region, at TeV energies they propagate rectilinearly. At intermediate energies the protons start to effectively leave the inner

3 pc region, and the steepening of the energy spectrum can be naturally referred to the energy dependent diffusion coefficient. What concerns the proton injection spectrum, it should be hard power-law, close to E^{-2}, with an intrinsic cutoff around 100 TeV. The required total energy of protons currently trapped in the gamma-ray production region, $W_p \simeq L_\gamma t_{pp\rightarrow\gamma} \simeq 10^{49}(n/10^{-3}\,\mathrm{cm}^{-3})^{-1}$ erg is quite modest, given that the density in the circum-nuclear ring could be as large as $10^5\,\mathrm{cm}^{-3}$ [183]. The total injection power of protons depends on the gas density and the diffusion coefficient; in the case of most economic realization of the scenario, the average rate of proton injection over the 10^3 year could be close to 10^{37} erg/s.

It is remarkable that this simple scenario allows a successful interpretation of the complex spectrum of gamma-rays extending over 5 decades from 0.3 GeV to 30 TeV, based only on an assumption of a single component of protons and a small number of reasonable model parameters. This, however, does not exclude other possible explanations, e.g. by two different components of protons, e.g. by a short flare that took place some 300 years ago on top of a long-term quasi-continuous proton injection (see Fig. 32). Moreover, diffuse gamma-ray emission can be produced also by high energy electrons.

In this regard, one should mention the leptonic models of extended (although not resolvable) sources of gamma-ray emission in GC. In particular, an interesting idea of formation of a hypothetical "black-hole plerion" powered by the advection-dominated accretion flow (ADAF) has been suggested in Ref. [81]. It has been shown that the electrons accelerated to Lorentz factor $\gamma \geq 10^8$ at the termination shock at 0.01 pc from the SMBH in GC can in principle reproduce the HESS data. However, the GeV emission predicted by this model appears well below the flux detected by *Fermi* LAT.

The recently discovered synchrotron nebula G359.95-0.04 located at 8 arcsecond from Sgr A* is another interesting site where TeV gamma-rays can be produced [153, 246]. Because of the high radiation density, the radiative cooling of electrons in this nebula is dominated by intense Compton scattering in the Klein-Nishina regime; this results in formation of an electron energy distribution which differs dramatically from those produced in PWNe located in the galactic plane [153]. Correspondingly, the broad-band radiation of electrons consisting of synchrotron and IC components also is characterized by a quite non-standard spectrum. It can reproduce the TeV data, however fails to explain the flux at GeV energies, unless we postulate existence of an additional, low energy component of electrons, or assume that the GeV gamma-ray emission belongs to a different source. In this regard, one should mention that recently a new leptonic model has been suggested [179] in which the electrons accelerated during flares in Sgr A*, are accumulated in the central ≤ 1 pc region around the super massive black hole. Since multi-TeV electrons cannot escape Sgr A*, the IC radiation of this region cannot be extended beyond the GeV domain.

Note that although all these leptonic models imply gamma-ray production in extended regions, they are still too compact to be resolved by current gamma-ray detectors. On the other hand, in the hadronic models the bulk of gamma-ray emission is produced in the circum-nuclear region of radius 3 pc, which for the distance to the galactic center of 8 kpc implies an angular size slightly more than 1 arcmin, i.e.

quite close to the angular resolution of CTA expected at energies above 10 TeV. The detection of a 1 arcmin TeV gamma-ray source in GC would prove the hadronic origin of gamma-ray emission and its connection to protons accelerated in Sgr A*. An independent argument in favour of Sgr A* as a source of multi-TeV protons is provided by the HESS observations of the huge molecular clouds located within 100 pc around GC.

Gamma Rays from the Central Molecular Zone

The interpretation of the spatially unresolved gamma-ray emission towards Sgr A* by interactions of runaway protons with the dense gas in the central several pc ring, implies no-sharp edges of the gamma-ray production region but rather a smooth transition to a new radiation component formed in more extended regions of GC. The energy and spatial distributions of this radiation depend on the injection history of protons and the character of their diffusion. For the given injection rate, slower diffusion allows more effective accumulation of protons, and correspondingly higher gamma-ray luminosity. And vice versa—fast diffusion may dramatically reduce the gamma-ray luminosity down to the level below the detection threshold even for very high injection rate of protons.

The diffusion time of protons increases with the distance as $t \propto r^2/D$ (D is the diffusion coefficient), thus the time-scales that characterize the processes of CR accumulation in the so-called Central Molecular Zone (CMZ) of radius ≈ 200 pc are much longer than in the inner 3 pc ring. Generally, if the accelerator is located in Sgr A*, the proton density is expected to fall with the distance from the galactic center. However, in the case of rectilinear propagation of particles (as expected at highest energies), at certain epochs after strong flares in Sgr A* we may expect, in principle, higher CR fluxes at larger distances compared to the inner parts of GC. In any case, independent of the spatial distribution of CRs, the large-scale diffuse gamma-ray emission from CMZ is expected to be highly inhomogeneous, because it largely reflects the distribution of the gas concentrated primarily in giant molecular clouds.

The HESS observations of the galactic center region indeed revealed an extended TeV gamma-ray emission from CMZ [39] with a clear indication of correlation with the most prominent giant molecular clouds (see Fig. 8). Using the maps of TeV gamma-ray emission obtained with HESS, and the maps of CS ($J = 1 - 0$) emission which contains information about the column density in dense cores of molecular clouds, we can "measure" the density of CRs in CMZ. In particular, assuming that the locally measured CR spectrum is valid in the GC, we can calculate the expected π^0-decay γ-ray flux from CMZ. It is seen from Fig. 34 that the gamma-ray spectra measured by HESS from the overall 200 pc region, as well as from individual clouds like the Sgr B2 complex, are significantly harder than the predicted one. The absolute CR flux derived from gamma-ray observations, also appears to be significantly enhanced (by an order of magnitude at multi-TeV energies) in this region relative to the local cosmic ray flux in the solar neighbourhood.
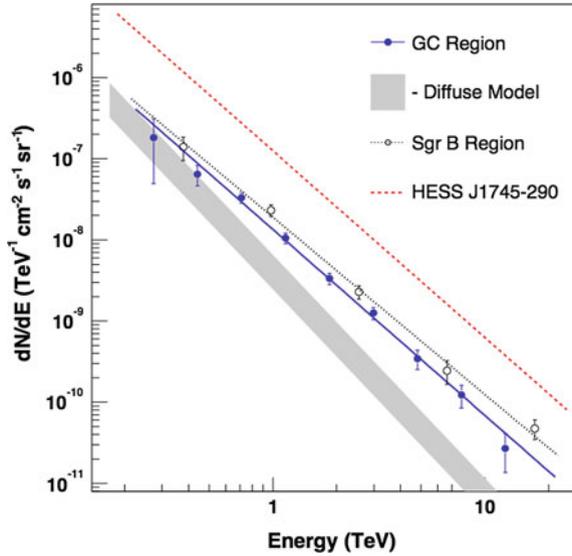
Fig. 34 Gamma ray flux per unit solid angle from different parts of the GC region (data points), compared with the contribution of gamma-rays expected from interactions of galactic cosmic rays [39]. The *grey shaded zone* shows the flux expected from CMZ within 200 pc due to hadronic interactions assuming for cosmic rays flux the local one as measured in the solar neighbourhood. The overall gamma-ray flux from the region of CMZ (*full circles*) consists of contributions from several clouds, one of which, the Sgr B2 complex, is shown separately (*open circles*). The flux of the central source (HESS J1745-290) towards Sgr A* is also shown (using an integration radius of 0.148)

This indicates to a strong nonthermal activity accompanied with proton acceleration which perhaps in the past was higher than at the present epoch. The HESS observations show that the ratio of gamma-ray flux to the molecular gas column density varies with galactic longitude, with a noticeable "deficit" of gamma rays at $l \approx 1.3°$. This interesting feature implies a non-uniform spatial distribution of cosmic rays in the ≤ 200 pc region of GC. Amongst other possible considerations, an interesting explanation of this effect could be the high activity of the particle accelerator in the recent past. Indeed, in this case protons accelerated in Sgr A* or in a nearby object, have not yet had time to diffuse out to the periphery of the 200 pc region. The epoch of the high activity of the accelerator depends on the proton diffusion coefficient. Assuming, for example, that the propagation of multi-TeV protons in the galactic center proceeds with a speed similar to the one in the Galactic Disk, the epoch of the high activity of the accelerator and the total energy release in relativistic particles during the outburst are estimated 10^4 year and 10^{50} ergs, respectively [39]. Two potential cosmic ray accelerators are located within the error box of HESS J1745-290. These are the supernova remnant Sgr A East the age of which is estimated between 1,000 and 10,000 years, and Sgr A* assuming that it was more active in the past.

Fermi Bubbles

Recently, remarkable evidence has emerged for two enormous gamma-ray structures centred on the core of the Galaxy and extending to approximately 10 kpc above and below the Galactic plane. Since these spectacular radiation features have been found in the *Fermi* LAT dataset [222], they received a nickname *Fermi* Bubbles. At lower galactic latitudes they coincide with somewhat smaller structures found earlier in the X-ray data of ROSAT [218] and in 20–60 GHz data of *WMAP* (the so-called *WMAP* "haze") [133]. Although currently the origin of these mysterious structures is under intense debates, it is clear that they are connected, in one way or another, to the activity of the galactic nucleus. While X-rays are of thermal origin, the microwave "haze" most likely is a result of synchrotron radiation of relativistic electrons [133], although a significant contribution of the dust to the microwave emission cannot be excluded.

What concerns the *Fermi* Bubbles—they undoubtedly have nonthermal origin. The parent relativistic particles (primarily protons) can be accelerated in the nucleus, and then injected into Fermi Bubbles. Alternatively, protons and electrons can be produced *in situ* through the first and/or second order Fermi acceleration mechanisms supported by hydrodynamical shocks or plasma waves in highly turbulent medium. The processes that create and support these structures can originate either from an AGN type activity related to the central black hole (Sgr A*) or from the ongoing star
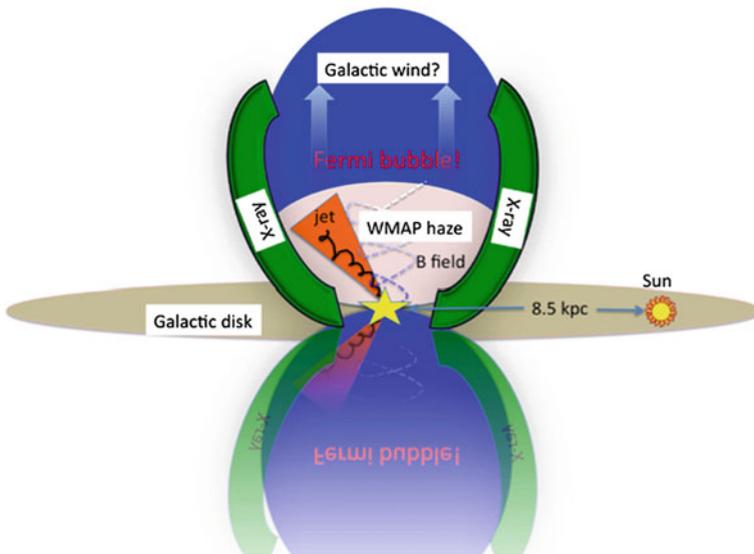
Fig. 35 The location and geometry of the *Fermi* Bubbles shown schematically together with the *WPAP* haze and the X-ray emission structure detected by ROSAT (from Ref. [222]). All three structures may have a common origin (although produced in different ways) related to much higher activity of the *central black hole* (Sgr A*) in the past or to the nuclear starburst in GC

formation in the galactic nucleus. The Dark Matter is another alternative for production of the *WMAP* "haze' and *Fermi* Bubbles, either directly or due to synchrotron radiation and IC scattering of electrons (the products of Dark Matter annihilation or decay), but the morphology of *Fermi* Bubbles disfavours the Dark Matter origin of these structures [222].

Figure 35 illustrates the location, geometry and scales of *Fermi* Bubbles, and their relation to the *WPAP* "haze" and the large-scale X-ray structures. The luminosity of gamma-rays with hard, E^{-2} type spectrum in the energy interval 1–100 GeV is quite impressive, $L_\gamma \approx 4 \times 10^{37}$ erg/s—an order of magnitude larger than the microwave luminosity of the "haze", and only an order of magnitude less than the thermal X-ray luminosity [222]. Clearly, given the overall limited energy budget of GC, the mechanisms of particle acceleration and gamma-ray emission in *Fermi* Bubbles should proceed with very high efficiency.

In this regard, a natural explanation of *Fermi* Bubbles and the *WMAP* "haze" would be that they are produced by the same population of highly relativistic ($\gg$10 GeV) electrons through IC scattering and synchrotron radiation, respectively [222]. Because of severe radiative energy losses, the mean free path of $\geq$100 GeV electrons is significantly shorter than the size of *Fermi* Bubbles. Therefore it is difficult to explain how the electrons could fill these vast structures, unless we assume very fast propagation of electrons with a speed exceeding 10,000 km/s, or postulate *in situ* electron acceleration throughout the entire volume of the bubbles [119, 222].

The requirement for ultra fast propagation of electrons seems to be unrealistic, both for the diffusion and the transport via galactic wind. On the other hand, *in situ* particle acceleration can be realized through stochastic (2nd order Fermi) acceleration [192] or due to series of shocks propagating through the bubbles and accelerating relativistic electrons [115]. It has been argued [192] that stochastic acceleration of electrons may reproduce rather well the energy dependent intensity profiles of gamma-radiation, and predict significant limb brightening above a few hundred GeV. The reason is the presumably higher turbulence due to presence of shocks at the bubble edges which would imply acceleration of electrons to higher energies close to these edges. Importantly, the suggested acceleration mechanism cannot boost the electron energy beyond 1 TeV, thus in order to explain the extension of the observed gamma-ray spectrum up to 100 GeV, one has to invoke FIR and optical/UV background emission supplied by the galactic disk (see Fig. 36). This model provides robust predictions. In particular, since the FIR and optical/UV contributions to the target field for IC scattering decrease quickly with distance from the disk, the spectrum of gamma-rays from high latitudes should contain a cutoff above tens of GeV. The limb brightening at highest energies is another characteristic feature predicted by this model. These spectral and spatial features can be explored in the near future, after the gamma-ray photon statistics in the *Fermi* LAT dataset is achieved an adequate level.

Hadronic origin of gamma-rays is an alternative interpretation suggested for *Fermi* Bubbles [119, 186]. Generally it is claimed to be inefficient channel for gamma-radiation, especially because of the low plasma density, $n \leq 10^{-2}$ cm^{-3}. However, these claims are based on a misleading interpretation of the efficiency of gamma-ray production processes in general, and their specifics in *Fermi Bubbles*, in particular. In
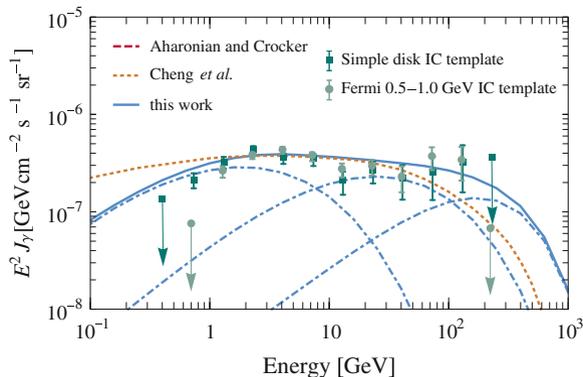
Fig. 36 The spectral energy distribution of gamma-rays from the *Fermi* bubbles compared to the theoretical predictions. (i) IC model of Ref. [192] (*solid line*) assuming stochastic acceleration of electrons in the bubbles (the contributions from the scattering on the CMB, FIR, and optical/UV backgrounds are shown separately); (ii) IC model of Ref. [115] (*dotted line*) assuming diffusive shock acceleration of electrons; (iii) hadronic model of Ref. [119] (*dashed line*). The figure is from Ref. [192]

fact, the *Fermi* Bubbles can be considered as an exemplary case when pp interactions proceed with very high efficiency. For example, the efficiency of this process can dramatically exceed, by several orders of magnitude, the gamma-ray production efficiency in young SNRs. Indeed, the plasma density alone does not tell us much about the efficiency; obviously, this issue should be discussed in the context of time-scales that characterize the gamma-ray production and the confinement of protons in the bubbles. If protons have been continuously injected and trapped in the bubbles over the time-scales of approximately 10^{10} year (the age of the Galaxy), the main energy in accelerated protons would be lost in pp collisions given that the characteristic time of the latter, $t_{pp} = 1/(k_p n \sigma_{pp} c) \approx 5 \times 10^9 (n/10^{-2})^{-1}$ cm^{-3} year, is shorter than the confinement time. This implies that we deal with the so-called "thick target" scenario, when the system is in saturation. The hadronic gamma-ray luminosity is equal to $L_\gamma \approx W_p / t_{pp \to \pi^0}$, where W_p is the total energy of protons in the bubbles, and $t_{pp \to \pi^0}$ is the time-scale for neutral pion production in pp interactions. In the saturation regime, $W_p = \dot{Q}_p t_{pp}$, assuming that the energy dissipation through pp collisions is the dominant loss process. Since $t_{pp} = 1/3\, t_{pp \to \pi^0}$, we have $L_\gamma = \dot{Q}_p/3$, thus about a third of the power injected into relativistic CRs emerges in gamma-rays (of all energies) independent of the local density, the source volume and the injection time. Note that since the time-scale of pp interactions is comparable to the supposed age of the bubbles of 10^{10} year, the efficiency would be somewhat less. Also, one should take into account that at low energies the ionization and adiabatic losses of protons play a non-negligible role, thus the overall efficiency for a broad energy spectrum of protons would be reduced to several percent. The fluxes of hadronic gamma-rays shown in Fig. 36 confirm these simple estimates. The numerical calculations, which have been performed assuming $E^{-2.1}$ type spectrum of protons

up to 10^{15} eV injected into the plasma of density $n = 5 \times 10^{-3}\,\mathrm{cm}^{-3}$, require an average rate of proton injection 10^{39} erg/s. Note that independent of the history of injection of relativistic protons, the current total energy in protons should be as high as $W_\mathrm{p} = L_\gamma t_{\mathrm{pp} \to \pi^0} \simeq 10^{55}$ erg which is comparable to the magnetic field energy in the bubbles.

It is remarkable that the current rate of star formation in GC can supply the required nonthermal power in relativistic protons. The total infrared luminosity of the inner few hundred parsec region of GC, $L_{IR} \simeq 1.6 \times 10^{42}$ erg/s, gives an estimate for the star formation rate of about $0.08\,\mathrm{M}_\odot$/year and, correspondingly, the supernova (SN) rate of 0.04/century. For the average mechanical energy released per SN, 10^{51} erg, the total power injected by supernovae is estimated $\dot{E}_{SN} \simeq 1.3 \times 10^{40}$ erg/s. Thus, for the "standard" 10 % efficiency of transformation of the SN explosion energy to accelerated particles, GC seems to be able to provide the required 10^{39} erg/s power in relativistic protons. This is a strong argument in favour of interpretation of gamma-ray emission of *Fermi* Bubbles by a population of relic cosmic ray protons and ions associated with star-formation in GC. This scenario implies a relatively slow release of energy from supernovae explosions and requires effective confinement of protons in *Fermi* Bubbles on multi-Gyr time-scales. Propagation of protons in the diffusion regime allows an effective confinement on these scales provided that the diffusion coefficient in *Fermi* Bubbles is smaller, by 1 or 2 orders of magnitude, than in the Galactic plane. This sounds a quite reasonable assumption given the very high turbulence of plasma in the *Fermi* Bubbles.

The gamma-ray production in a saturated regime minimizes the dependence of the production rate on the ambient plasma density which implies a rather constant volume emissivity despite possible variations of the plasma density. This is true, of course, for homogeneous distribution of cosmic rays which can be realized, for example, in the case of nearly constant diffusion coefficient throughout the bubbles. However, the spatial variations of the diffusion coefficient could be rather significant. For example, the higher plasma turbulence around the edges of the bubbles, would slow down the particle diffusion speed compared to the central parts of the bubbles, and thus would increase the concentration of protons closer to the edges. Correspondingly, this would lead to more homogeneous *projected* gamma-ray intensity (surface brightness) as it follows from observations [222]. In this regard, the criticism [192] of hadronic model of gamma-ray production in *Fermi* Bubbles, in particular in the context of expected and observed morphologies of gamma-ray emission, can be readily overcome assuming a strong gradient of turbulence towards the edges.

Finally, it should be noted that the hadronic origin of gamma-ray emission of the *Fermi* Bubbles does not exclude other "hadronic" scenarios with faster energy release related, for example to the activity of the central black hole Sgr A*. Fast energy release can be provided, for example, by the star capture processes by Sgr A* over the last 10 Myr with an average capture rate of 3×10^{-5} year^{-1} and energy release of 3×10^{52} erg per capture [116]. It has been argued in Ref. [116] that the quasi-periodic injection of hot plasma can produce a series of shocks in the bubbles which should be able to (re)accelerate protons to energies well beyond the "knee" up to 10^{18} eV. If confirmed by independent detailed hydrodynamical calculations, this
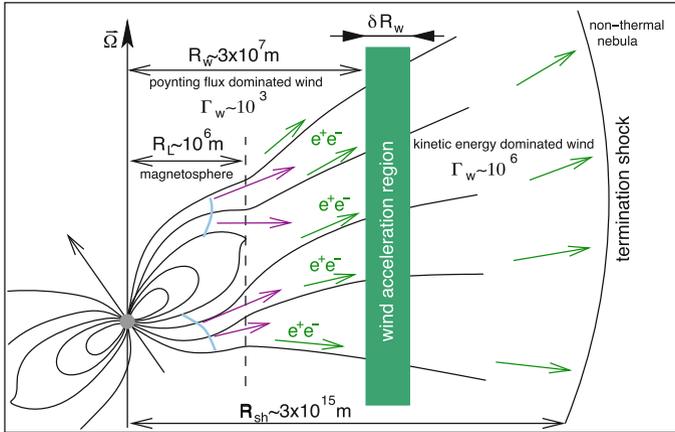
Fig. 37 Three regions of gamma-ray production related to rotation-powered pulsars: Pulsar Magnetosphere, Pulsar Wind, and Pulsar Wind Nebula. The specific numbers shown in the sketch characterize the pulsar in the Crab Nebula

could appear a quite natural solution for the origin of one of the most "problematic" (poorly understood) energy intervals of cosmic rays.

7 Pulsars, Pulsar Winds, Pulsar Wind Nebulae

Pulsars, rapidly rotating neutron stars left over after supernova explosions, are prominent sources of high energy gamma-rays produced in three physically distinct regions: (i) *Pulsar Magnetospheres* (PM), (ii) *Cold Ultrarelativistic Winds* (CRW) which carries almost the entire rotational energy of the pulsar in the form of Poynting flux and/or kinetic energy of the bulk motion, and (iii) the *Pulsar Wind Nebulae* (PWN)—the product of termination of the pulsar wind with subsequent acceleration of relativistic particles and their nonthermal emission. These three zones schematically are shown in Fig. 37.

7.1 Radiation of Pulsar Magnetospheres

In the pre-*Fermi* era only a handful of pulsars have been established as high energy gamma-ray sources. Observations with *Fermi* LAT dramatically increased the number of gamma-ray emitting pulsars to more than 100 from which almost one third have been discovered through their pulsed gamma-ray emission alone. The high quality phased-resolved energy spectra obtained with *Fermi* LAT provide a new insight into the pulsar demographics and physics. In particular, the measured light curves

and energy spectra indicate that gamma-ray emission from the brightest pulsars is produced in the outer magnetosphere with fan-like beams scanning over a large portion of the celestial sphere. The energy spectra of most of gamma-ray pulsars can be described by a simple function of "power-law with exponential cutoff" written in a generalized form

$$E^{-\Gamma} \exp\left[-(E/E_0)^b\right], \tag{52}$$

with $b \leq 1$, and the cutoff energy between 1 and 10 GeV [1]. The detection of gamma-rays beyond a few GeV with no indication of super-exponential attenuation (i.e. $b > 1$) effectively excludes the so-called *polar cap* model and gives a preference to models of gamma-ray production in the outer magnetosphere (to avoid severe pair production in the strong magnetic field in low-altitude zones).

The energy spectra given in the form of Eq. (52) with $b = 1$ generally are well explained by the mechanism of curvature radiation. Although most of the reported spectra can be fitted with the fixed value of $b = 1$ [1], however the extension of spectral measurements of the brightest gamma-ray pulsars towards both higher and lower energies revealed that the spectra beyond the cutoff could be smoother with b closer to 0.5. For example, the phase-averaged spectrum of the Crab pulsar is better fitted with the combination of parameters $b = 0.43$, $\Gamma = 1.59$ and $E_0 = 0.50$ GeV [110], rather than $b = 1$, $\Gamma = 1.97$ and $E_0 = 5.8$ GeV as reported earlier by the *Fermi* collaboration based on smaller gamma-ray statistics [5].

The classical radio pulsars are young astronomical objects of age between 10^3 and 10^5 year. Because of rotational losses, pulsars gradually slow down with time, and typically after 10^5 year become too faint to be detected. However, the pulsars produced in binary systems may have the "second" life. The matter of the companion star falling onto the neutron star spins it up with a rotation period as short as a few milliseconds. As a result, the "recycled" pulsars after $10^8 - 10^9$ years become visible again under a new name of *millisecond pulsars* (MSPs). In this regard, the detection of gamma-rays from tens of MSPs by *Fermi* LAT [2] can be considered as a natural outcome, although not without some surprise. This concerns, in particular, the discovery of gamma-rays from the millisecond pulsar J1823-3021A [134]. The large rate of change of the period of this pulsar derived from its gamma-ray luminosity implies an unexpectedly small age of $2.5 \cdot 10^6$ year, two orders of magnitude shorter than other MSPs, indicating the need for a revision of the current concept of formation of millisecond pulsars. Another interesting (to some extent unexpected) feature of MSPs is that despite significant differences in conditions (e.g. the magnetic field at the neutron star surface in MSPs is smaller than in young pulsars by four orders of magnitude), both the spectral shapes and time profiles of MSPs resemble those of young pulsars [2]. This indicates to the similar conditions formed in the gamma-ray production regions in both type of pulsars.

In general, *no* VHE gamma rays are expected from pulsar magnetospheres, except for realization of some specific conditions in MSPs [111], and perhaps also in some young (Vela-type) pulsars [214], when an additional inverse Compton component of emission can contribute to the radiation of electrons. Such a component seems to be

needed if the recently reported by the VERITAS [76] and MAGIC [75] collaborations pulsed VHE gamma-ray emission from the Crab pulsar has magnetospheric origin. However, the pulsar magnetosphere is not the only region from which we should expect pulsed emission. Periodic gamma-rays can be produced also in the pulsar wind.

7.2 Radiation of Pulsar Winds

The generally accepted paradigm of pulsars and their synchrotron nebulae [166, 212] postulates the existence of a relativistic electron–positron wind, which originates in the pulsar's magnetosphere and terminates in the interstellar medium. The large electric fields in the magnetosphere accelerate electrons, the interactions of which with the surrounding magnetic field initiate electromagnetic cascades. This leads to the formation of a relativistic outflow of a dense electron–positron plasma—the *wind*. Initially, the rotational energy losses are released in the form of Poynting flux (electromagnetic energy), thus the wind's Lorentz factor cannot be very large. The conversion of the Poynting flux leads to the 'acceleration' of the wind with Lorentz factor as large as $\Gamma_{\rm w} \sim 10^4$. Since the observations of the synchrotron nebula require an energy injection into the nebula close to the pulsar's spin-down luminosity [83, 121, 166, 245], the wind should transport the rotational energy of the pulsar to the nebula in the most effective and economical way. In particular, the wind should be 'cold' (in the sense of the low energy of the electrons in the frame of moving plasma) in order to avoid synchrotron losses. The presence of cold ultrarelativistic winds has been argued also in the context of radiation of binary pulsars [170, 173, 232].

Despite the general consensus on the concept, the formation of cold ultrarelativistic pulsar winds remains a mysterious phenomenon. To a certain extent, this is explained by the lack of information about the site(s) and the rate of acceleration of plasma (dissipation of the Poynting flux) in pulsar winds. And, of course, the very existence of pulsar winds still needs to be observationally confirmed. This is not an easy task. It is often argued that the cold ultrarelativistic wind is a 'dark substance' which cannot be directly observed. This statement is true for the synchrotron radiation of the wind—relativistic electrons move together with magnetic field and therefore do not emit synchrotron radiation. However, the cold ultrarelativistic winds of pulsars can be directly detected through their characteristic inverse Compton radiation [100]. Moreover, we already do have a strong evidence in this regard; it is quite plausible that the VHE pulsed gamma-radiation recently reported from the Crab pulsar [75, 76], in fact originates from the pulsar's wind [62].

The phase-averaged spectral points of the Crab pulsar measured by *Fermi* LAT [5] are shown together with VHE points reported by VERITAS [76] and MAGIC [75] in Fig. 38. It is seen that the VHE points are located well above any extrapolation of *Fermi* data fitted by the standard form of the pulsar spectra given by Eq. (52). On the other hand, the extrapolation of GeV fluxes to the VHE-domain as a power-law with photon index $\Gamma \approx 3.8$, and its interpretation as an evidence for the same mag-
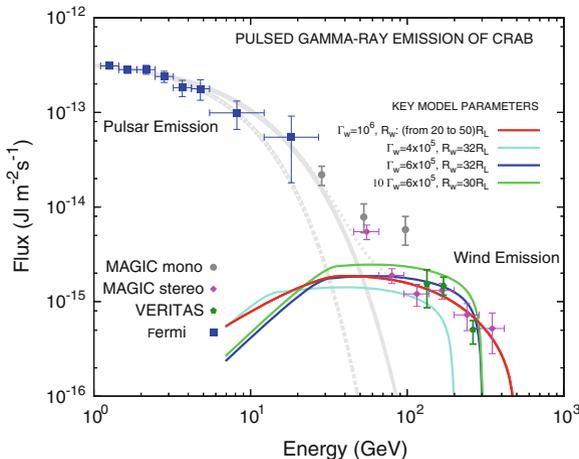
Fig. 38 SED of gamma-radiation produced by the pulsar magnetosphere and by the pulsar wind. The *Fermi* LAT points are best-fitted by Eq. (52) with $b = 1$, $E_0 = 5.8$ GeV and $\Gamma = 1.97$ (*dashed grey line*) or by slightly harder spectrum in the cutoff region with $b = 0.85$, $E_0 = 7$ GeV and $\Gamma = 1.97$. These spectra can be explained by curvature radiation of electrons in the pulsar's magnetosphere. The IC emission of the 'cold' ultrarelativistic wind can explain the pulsed gamma-ray fluxes above 100 GeV. The *solid light blue*, *blue* and *green curves* are calculated under the assumption of an instant acceleration of the wind to the fixed value of the Lorentz factor Γ_w and at the fixed radius R_w (these parameters are indicated in figure). The *solid red curve* corresponds to the case when the Poynting flux transformation takes place within the zone 20–$50R_L$ with linear dependence of the wind's acceleration rate with the distance; the maximum Lorentz factor achieved at $50R$ is set to 10^6 (from Ref. [62])

netospheric origin of GeV and TeV gamma-rays (see e.g. Ref. [75]) requires, in fact, a rather dramatic revision of basic concepts currently employed in magnetospheric models. Also, the assumption of the magnetospheric origin of radiation over the entire gamma-ray domain contradicts the essentially different lightcurves reported at GeV and TeV energies (unless the production sites of these two components are well separated), as well as the apparent tendency of spectral flattening above 100 GeV. A natural and more plausible site of production of pulsed VHE gamma-rays is the ultrarelativistic wind being illuminated by pulsed X-rays arriving from the pulsar's magnetosphere [62]. It is demonstrated in Fig. 38 that the entire gamma-ray region can be considered as a superposition of two separate components—the nominal (magnetospheric) GeV gamma-rays and the VHE component of the Comptonized wind.

Although the inverse Compton gamma-rays are produced by monoenergetic electrons, the SED of gamma-rays in the range of tens to hundreds of GeV is quite flat. This is caused by the combination of effects related to the broad power-law distribution of seed photons and the transition of the Compton cross-section from the Thomson to the Klein-Nishina regime. On the other hand, the spectrum is expected to have a very sharp cutoff at $E = \Gamma_w mc^2$. This not only serves as a distinct feature

for identification of the wind origin of gamma rays, but also allows derivation of wind's Lorentz factor.

Generally, the lightcurve of target photons should be reflected in the time structure of the IC gamma ray signal, however, they cannot be identical because of the effects related to the geometry as well as the specifics of the anisotropic IC scattering. The geometrical effects may lead to significant differences between the arrival times of the target photon and the secondary gamma ray pulses as illustrated in Fig. 39. It is seen that whilst for the wind located close to the light cylinder the gamma-ray signal appears significantly shifted relative to the reported gamma-ray data, $\Delta t \sim 0.1T$, for the wind acceleration at $R_\mathrm{w} = 30 R_\mathrm{L}$ the widths and the positions of the predicted and observed gamma ray peaks P1 and P2 are in very good agreement. On the other hand, in the case of isotropic wind the predicted P1/P2 ratio of the gamma-ray signal mimics the X-ray lightcurve ($\approx$2), while the reported VHE gamma-ray data [75, 76] show P1/P2 < 1. This can be explained by anisotropy of the wind which introduces significant corrections to the shape of the gamma ray lightcurve in general, and in the P1/P2 ratio, in particular (see Fig. 39).

In summary, the interpretation of Crab's pulsed VHE gamma-ray emission in terms of Comptonization of the pulsar wind implies the first direct evidence of cold ultrarelativistic electron–positron wind in a pulsar. The conversion of the Poynting flux to kinetic energy of bulk motion should take place abruptly in the narrow cylindrical zone between 20 and 50 light-cylinder radii and should accelerate the wind to a Lorentz factor of $(5 - 10) \times 10^5$.

7.3 Pulsar Wind Nebulae

Before the discovery of the Crab pulsar, the compact stellar object in the center of the Crab Nebula has been proposed as a potential source of the nebular magnetic fields and relativistic particles [206]. Immediately after the discovery of pulsars [150] and their identification with neutron stars [144, 150], the rotation energy of these compact stars was recognized as the ultimate energy source powering PWNe [203]. A bit later, the existence of relativistic electron–positron pulsar winds has been postulated to explain the link between pulsars and PWNe [212]. The current paradigm of formation of PWNe is based, to a large extent, on the MHD model of interaction of a cold ultrarelativistic electron–positron wind with interstellar medium [166] which, despite a number of introduced simplifications, satisfactorily describes the main features of the Crab Nebula [83, 121]. The recent detailed two-dimensional MHD simulations [101, 245] confirm the basic concept, at least for the Crab Nebula— the most prominent representative of this source population. Moreover, even the fine structures, like the toroidal and jet-like features in the inner part of the Crab Nebula and some other PWNe discovered in X-rays by *Chandra*, can be explained within the framework of the model of Kennel and Coroniti [166], after introducing initial anisotropy of the energy flux in the wind [102, 122, 175]. For a comprehensive review on observational and phenomenological studies of evolution and structure of
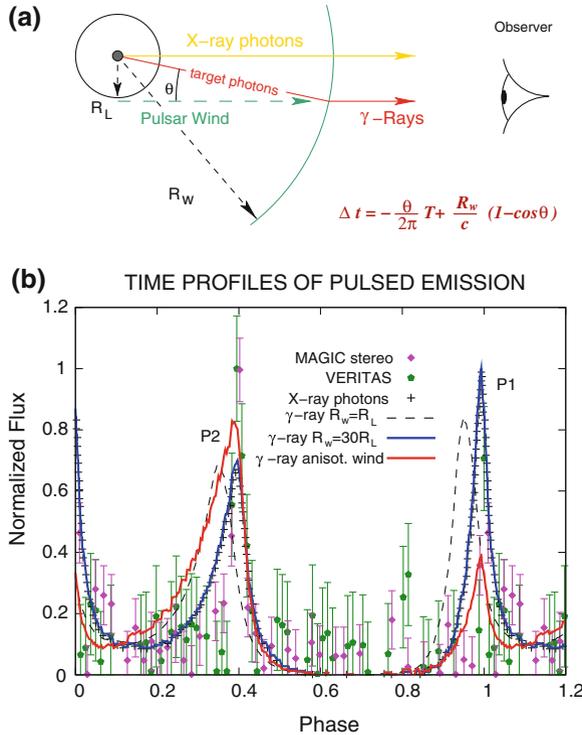
Fig. 39 Formation of the pulsed VHE gamma-ray signal in the wind of the Crab pulsar. **a** Geometry of the IC scattering of magnetospheric X-rays by the ultrarelativistic cold wind; **b** the reported lightcurves of VHE gamma-rays together with theoretical lightcurves of the wind IC emission calculated for different assumptions concerning the acceleration site of the wind and it anisotropy (from Ref. [62])

PWNe the reader is referred to [138], and for the theoretical models and challenges of the physics of pulsar winds and nebulae, to the review article [174].

Crab Nebula

The gamma-ray emission of the Crab Nebula has been studied by many groups using different detection techniques. Presently the reported fluxes cover the energy interval from 1 MeV to 100 TeV (see Fig. 40). Remarkably, even this impressive coverage of 6 decades in energy is only a fraction of the broad band non thermal spectrum that extends over 21 (!) decades of frequencies—from radio wavelengths to very high energy gamma-rays. Within the MHD wind paradigm of the Crab, the ultrarelativistic pulsar wind terminates by a standing reverse shock at which the electrons and positrons of the wind are accelerated and randomized. Although there is no doubt in the important role of the relativistic shock for acceleration
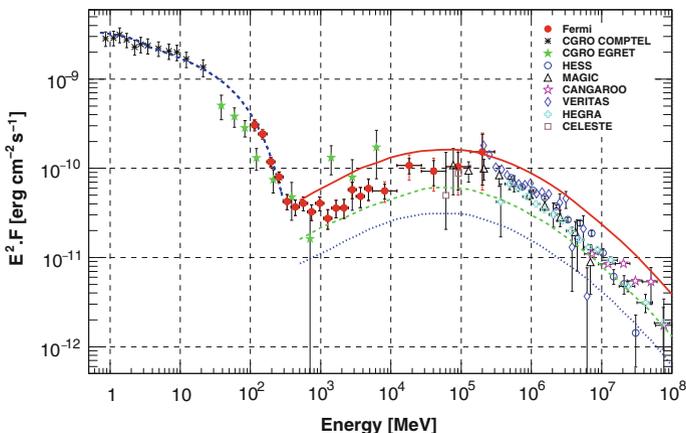
Fig. 40 SED of gamma-rays of the Crab Nebula. The spectral points from low to VHE gamma-rays are shown together with the fit of the synchrotron component (*blue dashed line*) and predictions of IC gamma-rays calculated for three different values of the mean magnetic field: $B = 100\,\mu$G (*solid red line*), $B = 200\,\mu$G, and the equipartition field of the nebula of $300\,\mu$G [83] (from Ref. [5])

of electrons, the acceleration mechanism itself remains unknown (see for a review Ref. [174]). In principle, the electrons can be accelerated in the course of reconnection of the alternating magnetic field at the termination shock, but it is not clear that this mechanism alone would be sufficient to boost the electrons to highest energies. For the hard power-law spectrum of electrons derived from observations of PWNe up to 100 TeV and beyond, the first-order Fermi mechanism seems to be a more feasible scenario [22], although it is not free certain difficulties, in particular it requires significant amplification of turbulence downstream of the shock [202].

The major fraction of the rotational energy of the pulsar transported through the wind is eventually released in non thermal synchrotron radiation observed in radio, optical, X-ray and gamma-rays extending to energies ≥ 100 MeV. The mean energy of synchrotron photons ε_s radiated by an electron of energy $E_{\mathrm{TeV}} = E/1$ TeV in the magnetic field of strength $B_{\mu G} = B/1\,\mu$G is

$$\varepsilon_s \simeq 0.02 E_{\mathrm{TeV}}^2 B_{\mu G} \text{ eV.} \tag{53}$$

For the typical magnetic field of order of $100\,\mu$G at the shock [166], the energy of electrons responsible for ≥ 100 MeV synchrotron radiation should be as large as 10^{15} eV. Thus we deal with a cosmic *PeVatron*. It is remarkable that the Crab Nebula is not simply the first (and so far the only) cosmic accelerator identified as a *PeVatron*, but, more importantly, it operates as an *electron PeVatron*. Because of severe radiative losses of electrons, their acceleration to such high energies is possible only if the acceleration proceeds at a rate close to the theoretical limit given by Eq. (28) with $\eta \sim 1$. Note that the very fact of extension of the synchrotron spectrum up to ≥ 100 MeV is a clear indication, independent of the model parameters,

in particular the strength of the magnetic field, that the acceleration in this object proceeds at a rate close to the theoretical limit (see Sect. 3.2). The recent discovery of gamma-ray flares [11, 110, 221, 235] confirms these early conclusions (see e.g. Ref. [52]) but, at the same time, brings new puzzles and challenges in the understanding of the physics behind the extreme acceleration of electrons and related to them the spectacular gamma-ray flares.

The inverse Compton scattering of the same population of ultrarelativistic electrons provides the second channel for gamma-ray production resulting in formation of very high energy tail of gamma-radiation. Although the energy release in IC constitutes only a small fraction of the synchrotron luminosity of the nebula, the high energy gamma-radiation provides a unique channel of information about conditions in the nebula. In particular, the comparison of X- and TeV gamma-ray fluxes observed from the Crab Nebula leads to a robust estimate of the mean magnetic field between 100 and 200 μG (see Fig. 40), in a remarkably good agreement with predications of the theory of termination of the MHD pulsar wind [166].

Figure 40 demonstrates the homogeneous coverage of the entire gamma-ray domain with high quality data. While the COMPTEL and EGRET results carry information about synchrotron radiation in the cutoff region, the *Fermi LAT* data show the sharp transition from the synchrotron to the IC component of radiation around 1 GeV. Another important feature seen in Fig. 40, is the clear indication of the IC maximum around 100 GeV supported by both satellite (*Fermi* LAT [5]), and ground-based (MAGIC [75] and VERITAS [76]) measurements which show remarkable agreement with each other. The measurements of ground-based observations have almost approached 100 TeV [31, 42, 230], but current instruments are not able to explore the region beyond 100 TeV. Meanwhile, the IC component should extend to the energy region set by the maximum energy of accelerated electrons, i.e. 1 PeV. Although the production of gamma-rays at such energies takes place in the Klein-Nishina regime, and therefore is suppressed, still it can be detected by next generation gamma-ray detectors as it is demonstrated in Fig. 7.

As the brightest persistent point-like TeV gamma-ray source seen effectively from both hemispheres, the Crab Nebula has been considered as a standard candle for cross-calibration of detectors, and correspondingly this technical issue often has been claimed as the major motivation for high energy gamma-ray observations, assuming that the "astrophysical" objectives already have been achieved. However, the Crab flares above 100 MeV [11, 235] indicate that many details remain unresolved, and that we are far from proper understanding of many aspects related to this unique source. The new, exceptional activity of the source in April 2011 [110, 221] provided higher quality data which however only deepened the challenges. Two most striking features of these flares are the short times scales of flares with a rise time as short as 6 h and extension of the energy spectrum to GeV energies (one should note that similar features have been found earlier, in an independent analysis of *Fermi* LAT data on the 2010 flares [87]).

The exceptionally high fluxes during the active state in April 2011 allow detailed spectroscopy for different flux levels (see Fig. 41). In order to study the spectral evolution of the flaring component, a steady-state (constant) background has been
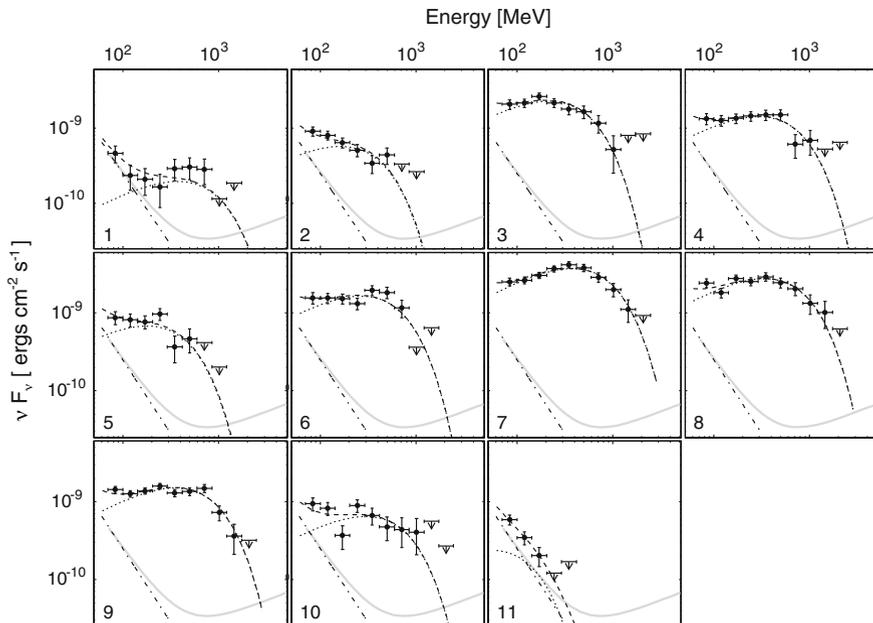
Fig. 41 Time evolution of the Crab SED during the April 2011 flare. The flare duration of approximately 9 days has been divided in 11 time windows of approximately constant flux. The *dot-dashed line* indicates the assumed constant background from the synchrotron nebula. The *dotted lines* show the flaring component, and the *dashed lines* are the sums of the background and flaring components (from Ref. [110])

assumed with a steep power-law spectrum described by a photon index $\Gamma_b = 3.9$. The spectrum of the flaring component has been assumed in the form of power-law with exponential cutoff, $\nu F_\nu = f_0 E^{2-\Gamma_f} \exp[-(E/E_0)^\kappa]$. The spectra during all selected windows can be well described by the same photon index $\Gamma_f = 1.27 \pm 0.12$ and the exponential cutoff index $\kappa = 1$, but with variable total flux f_0 and the cutoff energy E_0. It has been found that while the cutoff energy varies only by a factor of two, the total flux can be changed more than an order of magnitude. If the flares are caused by synchrotron radiation of electrons, the strictly exponential cutoff with $\kappa = 1$ and photon index $\Gamma_f = 1.27$ imply a narrow distribution of electrons with a very sharp high-energy cutoff. Such a distribution hardly can be explained by the diffusive shock acceleration, at least in the standard scheme of its operation. An alternative mechanism of acceleration due to the magnetic reconnection (see e.g. Ref. [252]) seems a viable option. In particular, in the recently suggested scenario of *reconnection-powered linear electric accelerator* [241], the electrons are accelerated by the electric field to highest (1 PeV or so) energies before suffering significant radiative losses, therefore one should expect very narrow distribution of electrons. Another attractive feature of this mechanism is that the synchrotron radiation of electrons can well exceed the "standard" maximum energy of about 100 MeV, indicated

earlier in Ref. [172] for relativistic current sheets. The obvious reason of this effect is the assumption that the electric field exceeds the strength of the magnetic field. This follows from Eq. (30) in Sect. 3.2. The condition $E > B$ implies that the η parameter defined as $\eta = B/E$ is less than 1. Correspondingly the maximum energy of synchrotron radiation is shifted to higher energies by the factor of $E/B > 1$ [231]. The episodic nature of radiation in this model is explained by the motion of PeV electrons along the current sheet the orientation of which has rather random character. Therefore one can observe sudden increase of synchrotron emission with shifted synchrotron peak only when the particle beam is directed by chance towards the observer [241]. It should be noted that the condition $E > B$ implies a breakdown of MHD. Wile this "phase transition" of plasma can (or even should) take place at the acceleration of the pulsar wind [97], the formation of reconnection current sheets in the termination shock region seems a less obvious realization; this question needs further theoretical studies.

There are other ways to explain the extension of the gamma-ray spectra of flares beyond the "nominal" synchrotron cutoff. The simplest (although somewhat artificial) solution could be to assume that the acceleration and radiation regions are effectively separated, namely, the electrons are accelerated in a region of low magnetic field, then enter and radiate in a region with stronger magnetic field. A more realistic modification of this possibility is the acceleration and radiation of electrons in the termination shock region at the presence of stochastic magnetic field [112]. The strong magnetic field fluctuations, which can be facilitated by multiple forced magnetic field reconnections, is a key element in this scenario. While the acceleration and energy loss rates depend on $\langle B^2 \rangle$, the synchrotron radiation in the cutoff region is contributed by higher statistical moments of the stochastic field distribution. This might shift the position of the synchrotron cutoff to higher energies. For example, in the case of Gaussian type *probability distribution function* of magnetic field, the cutoff energy can be increased up to 300–400 MeV [112]. This will be accompanied by a dramatic, by an order of magnitude increase of the gamma-ray flux around 1–2 GeV as has been observed during the Crab flares. This model does not address, however, another important feature of flares, namely their extremely fast variability, $\Delta t \leq 6$ h, which implies that the small size of the emitting region should be as small as $l \leq \Delta t \times c \approx 10^{15}$ cm. On the other hand, the gamma-ray luminosity at the peak of the April 2011 flare was 4×10^{36} erg/s, close to 1 % of the spin-down luminosity of the pulsar. It is difficult to propose a realistic scenario of channelling of substation fraction of the available energy into a compact regions with a linear size less than 10^{-3} of the termination shock radius! In phenomenological terms, the simplest (and quite obvious) solution of this general problem is to propose that we deal with an apparent but not intrinsic luminosity, i.e. assume that the emission is strongly beamed towards the observer within a small angle [241], or the flares are a result of strongly Doppler boosted radiation from compact structures (see e.g. Ref. [231]). These could be fragments of the wind termination shock (see e.g. Refs. [93, 176, 187]) or some specific compact features in the inner nebula (see e.g. [185, 249]).

The hypothesis of Doppler boosted radiation is especially attractive because it allows reduction of energy requirements by orders of magnitude, relaxes the con-

strains on the emission size, and offers different ways of sudden increase of the Doppler factor of the magnetized non thermal plasma. Yet, there are key issues, e.g. the origin of the relativistic motion of the synchrotron source and its large (mG scale) magnetic field, which can be addressed only by detailed relativistic MHD simulations. Such studies already have started, and it is likely that soon they will tell us much more about the Crab flares. On the other hand, it is likely that the Crab flares are not a rare events, so it is expected that the detection of new flares by *Fermi* and *AGILE* will provide more information about the features of this unique phenomenon. The search for IC gamma-rays at TeV energies during GeV synchrotron flares seems an obvious motivation for ground-based observations, although the expectations concerning the detectable counterpart TeV gamma-ray flares should not be overestimated.

TeV PWNe

The strong magnetic field in the Crab Nebula initiated by termination of the powerful pulsar wind dramatically reduces the efficiency of production of gamma-rays through IC scattering. The energy density of the magnetic field exceeds by more than two orders of magnitude the radiation density. Thus less than 1 % of the energy of accelerated electrons is released in IC gamma-rays, the rest being emitted trough synchrotron radiation. In the case of the Crab Nebula, the low gamma-ray production efficiency is compensated by the huge energy budget of the source provided by rotation of the pulsar and converted very effectively to multi-TeV electrons.

In other PWNe with less powerful winds, the resulting nebular magnetic fields can be weaker by a factor of 10–100, which makes these objects as potentially (very) effective gamma-ray emitters. The main target photons for production of VHE gamma-rays in these objects are contributed by the 2.7 K CMBR, thus the average energy of IC gamma-rays produced in the Thomson limit is

$$\varepsilon_{\text{ic}} \simeq 5 \times 10^9 E_{\text{TeV}}^2 \text{ eV}. \tag{54}$$

Equations (53) and (54) give a direct relation between the characteristic energies of synchrotron and IC photons produced by the same electrons:

$$\varepsilon_{\text{s}} \simeq 7(\varepsilon_{\text{ic}}/1\text{TeV})B_{\mu\text{G}} \text{ eV}. \tag{55}$$

One should note that at the highest energies the Compton scattering proceeds in the Klein-Nishina regime, thus a non-negligible correction to the energy is needed. As the maximum energy of the upscatered photon is $\varepsilon_{\text{ic,max}} = (1 + 1/b)^{-1}E_{\text{e}}$, where $b \approx 0.005E_{\text{TeV}}$, the Klein-Nishina effect becomes significant at $E_{\text{e}} \geq 100 \text{ TeV}$.

For a given strength of magnetic field, the energy losses of electrons are shared between the synchrotron and IC radiation components as $L_\gamma/L_X = w_{\text{MBR}}/w_{\text{B}} \simeq 1 (B/3 \mu\text{G})^{-2}$. This implies that in a PWN with a nebular magnetic field of about 10 μG or less, the IC gamma-ray production efficiency could be as large as 10 %. Given that the rotational energy of pulsars is eventually released in relativistic elec-

trons accelerated at the termination shock, the detectors with sensitivities typical for the current IACT arrays should be able to reveal TeV gamma-ray emission from tens of PWNe of young pulsars with spin-down luminosities $L_0 \geq 10^{34} d_{\text{kpc}}^2$ erg/s [59]. The discovery of more than two dozens of extended TeV gamma-ray sources in the close vicinity of pulsars in the indicated rage of luminosities, supports this phenomenological prediction. Presently PWNe constitute the largest galactic TeV source population. Many previously dubbed "dark" TeV gamma-ray sources, including the first unidentified TeV gamma-ray source discovered by the HEGRA collaboration, TeV J2032+4130 [29], later have been identified with PWNe. Most of these identifications with PWNe are quite convincing, but still tentative, except for several HESS sources which are firmly identified with famous PWNe. VHE gamma-ray images of four PWNe, Vela X, MSH 15-52, the nebula associated with the pulsar PSR J1826-1334, and the Kookaburra complex are shown in Fig. 42.

A common feature for all TeV sources identified, either firmly or tentatively, with PWNe is that the gravity centres of the extended TeV images do not coincide with the positions of parent pulsars. Asymmetric, one-sided images of these PWNe have been found also in X-rays, but on significantly smaller scales. Although the mechanism which causes PWN offsets from the pulsar positions, is not yet firmly established, this effect could be linked to the propagation of the reverse shock created at the termination of the pulsar wind in highly inhomogeneous medium (for a discussion of this issue see [138]).

The second interesting feature of TeV PWNe is the significantly larger extension of TeV images compared to the X-ray images. Basically, this is the result of combination of several factors. (i) Generally, in PWNe with magnetic field of order of $10\,\mu$G or less, which seems to be the case of most of TeV PWNe, the electrons responsible for X-ray emission are more energetic than the electrons emitting TeV gamma-rays. Therefore the synchrotron burning of highest energy electrons results in a smaller size of the X-ray source. (ii) When electrons diffuse beyond the PWN boundary, they emit less synchrotron radiation (because of the reduced magnetic field), but they still effectively radiate gamma-rays via inverse Compton scattering on the universal CMBR. (iii) Finally, because of the high X-ray background, the sensitivities of X-ray detectors like Chandra and XMM-Newton are dramatically reduced beyond several angular minutes. This significantly limits the potential of these instruments for weak extended X-ray sources. In contrast, the sensitivity of IACT arrays remains almost unchanged approximately within $1°$ radius of FoV. All these factors make the IACT technique as the most powerful tool for studies of non thermal population of electrons in PWNe.

The similar morphologies of PWNe in X-ray and TeV gamma-ray bands supports the widely accepted view that both emission components are due to the radiation of the same population of multi-TeV electrons via synchrotron and inverse Compton channels, respectively. Generally, the inverse Compton models predict energy dependent morphology of PWNe in TeV gamma-rays caused by the energy losses and the propagation effects of multi-TeV electrons. Remarkably, a tendency of softening of the energy spectrum of gamma-rays with distance from the position of the pulsar has been found in the gamma-ray image of HESS J1825-137 [40]. This is indeed
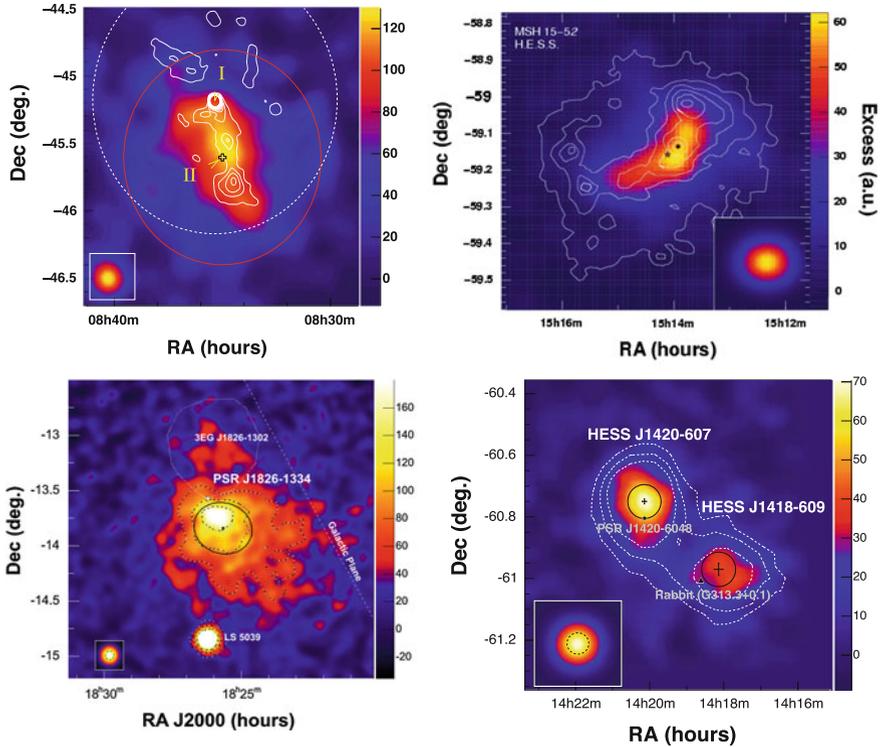
Fig. 42 TeV gamma-ray images of four PWNe: *top left* Vela X. The *white contours* correspond to the distribution of X-rays obtained with ROSAT. *Top right* The pulsar wind nebula MSH 15-52. The smoothed excess map is shown together with X-ray distribution (*white contours*) measured by ROSAT. The *black point* and the *black star* correspond to the pulsar position and the center of gravity of gamma-ray distribution, respectively. *Bottom left* The PWN HESS J1825-137. The position of the pulsar PSR J1826-1334 is marked by a *white triangle* (the bright point-source to the south is the microquasar LS 5039). *Bottom right* The map of TeV gamma-rays within the $1° × 1°$ FoV in the Kookaburra complex region. The positions of two separated sources are marked with error crosses. The position of the pulsar PSR J1420-6048 is marked with a *star*, the position of the source called Rabbit (G313.3+0.1) is marked with *black triangle*. In all figures the PSF of the instrument is shown in the *lower left* or *right* hand corners

a strong argument in favour of the inverse Compton origin of TeV emission of this source. On the other hand, the TeV gamma-ray luminosity of this source is about 10 % of the pulsar spin-down luminosity. While formally this is still within the available energy budget of the system, the IC cooling time for electrons responsible for $≤1$ TeV gamma-rays appears longer compared to the age of the source. This makes quite difficult the interpretation of gamma-ray emission by the current spin-down power of the pulsar. A natural solution to the problem of the energy budget crisis could be the contribution of "relic" electrons produced at the early epochs when the pulsar's spin-down luminosity was significantly higher [40].
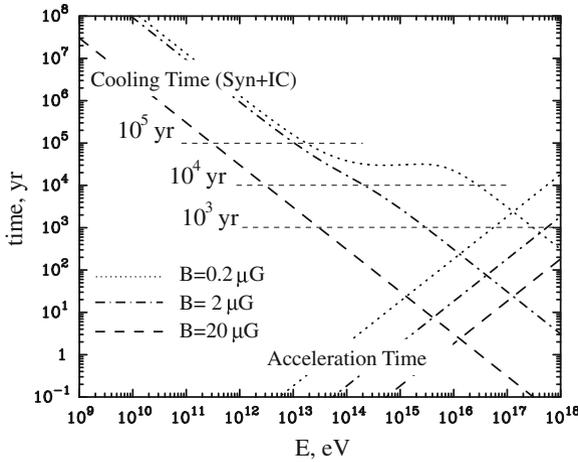
Fig. 43 Acceleration and radiative (synchrotron and IC) cooling time-scales of electrons in a PWN calculated for the strength of the magnetic field $B = 2 \times 10^{-7}$ G, $B = 2 \times 10^{-6}$ G and $B = 2 \times 10^{-5}$ G

Note that the significant enhancement of production rate of electrons at early epochs of pulsars has been independently argued in the context of observations of radio data [82, 109, 227]. However, unlike the low energy (radio) electrons, the "relic" TeV electrons from early epochs can survive only if the magnetic field does not significantly exceed 1 μG. This is demonstrated in Fig. 43, where the characteristic lifetimes of electrons are shown for 3 different values of the magnetic field—0.2, 2 and 20 μG. It is seen that while at the presence of magnetic field of 20 μG the radiative cooling time of electrons of energy 10 TeV is approximately 10^3 year, for the field of ≈1 μG their lifetime is increased almost by two orders of magnitude allowing accumulation of electrons from epochs as early as 10^5 year. This not only increases dramatically the number of potential TeV PWNe, but also can make very bright the TeV gamma-ray emission of some individual representatives of this source population. The key condition for realization of this scheme is the requirement of a weak nebular magnetic field, $B \ll 10$ μG.

The strength of the magnetic field varies over the nebula. In the cold ultra-relativistic wind zone, the toroidal component of the field (dominant at large scales) decreases linearly with the distance to the pulsar. In the shocked region, magnetic field may follow different trends depending on the magnetization parameter σ. At the termination shock with a characteristic radius

$$r_{\rm sh} \simeq 5 \cdot 10^{18} \left(\frac{L_{\rm SD}}{10^{37} \, {\rm erg/s}} \right)^{1/2} \left(\frac{p_{\rm ext}}{10^{-12} \, {\rm erg/cm^3}} \right)^{-1/2} \, {\rm cm}, \qquad (56)$$

the magnetic field is estimated
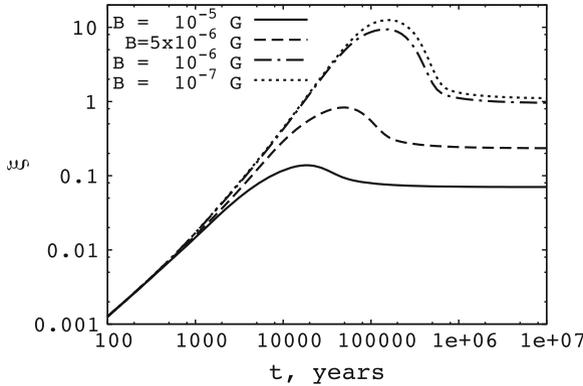
Fig. 44 The efficiency of radiation of IC gamma-rays in a PWN as a function of time, calculated for different strengths of magnetic field and for the history of the electron injection rate given by Eq. (58)

$$B \simeq 6 \left(\frac{\sigma}{3 \cdot 10^{-3}} \right)^{1/2} \left(\frac{p_{\text{ext}}}{10^{-12} \text{ erg/cm}^3} \right)^{1/2} \mu G, \tag{57}$$

where p_{ext} is external to the nebula pressure contributed by the matter, magnetic field, and cosmic rays outside the nebula. Interestingly, the B-field strength doesn't depend on the spin-down luminosity of the pulsar, but only on the external medium pressure. The estimated magnetic field strength is significantly smaller than interstellar magnetic field, $B_{\text{ism}} \approx 3\,\mu G$. This implies a rather specific cooling regime when electrons loose their energy mainly through IC scattering on the CMBR photons.

The higher spin-down luminosity of pulsars in the past has a dramatic effect on the radiation efficiency which usually is determined as the IC gamma-ray luminosity of the nebula to the current spin-down luminosity of the pulsar. In principle, the efficiency determined in this way can be larger than 1. The "trick" here is related to the relic electrons that have been accelerated in the past, but survived the energy losses, and today provide the dominant contribution to production of gamma-rays. It is seen from Fig. 44 where the radiation efficiency is shown for the time-dependent injection of electrons (assumed to be proportional to the spin-down luminosity) in the following form:

$$Q \propto \gamma^{-2} \left(1 + \frac{t}{t_0} \right)^{-2}, \tag{58}$$

with $t_0 = 5.3 \times 10^3$ year. This presentation corresponds to the case when the current spin-down luminosity of the pulsar of age 1.7×10^4 year is smaller by a factor of 10 compared to the spin-down luminosity at the birth of the pulsar (the so-called braking index was assumed to be equal 3). Note that for a strong magnetic field, the impact of the high initial luminosity is not significant. This is explained by the fact that multi-TeV electrons produced in the past do not survive
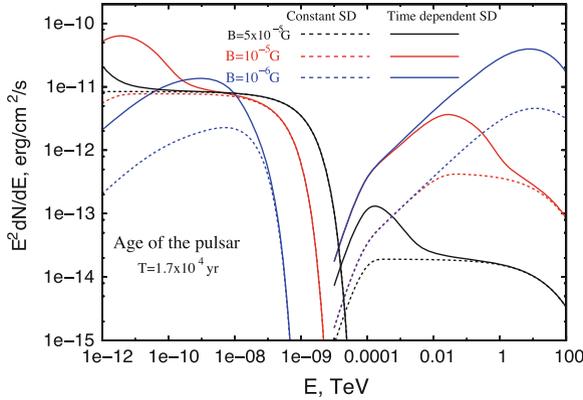
Fig. 45 The spectral energy distributions of the synchrotron and inverse Compton radiation of a PWN. The calculations are performed calculated for the steady injection (*dotted lines*) and time-dependent injection (*solid lines*) of electrons given by Eq. (58). The *curves* correspond to different strengths of the magnetic field. The shown fluxes correspond to the spin-down luminosity of the pulsar $L_0 = 10^{34} (d/1\,\mathrm{kpc})^2$ erg/s, where d is the distance to the source

because of fast synchrotron losses. On the other hand, for a very weak magnetic field, $B \leq 1\,\mu\mathrm{G}$, the gamma-ray efficiency can significantly exceed 100 %. Obviously, this does not imply a violation of conservation laws, but simply is a result of the chosen definition of the efficiency when the gamma-ray luminosity is compared with the *current* spin-down luminosity of the pulsar. In Fig. 45 the broad band spectral energy distributions are shown for three different values of the magnetic field, and assuming two different time histories of electron injection (following the evolution of the pulsar): (i) constant injection rate (dotted lines) and (ii) time-dependent electron injection given by Eq. (58) (solid lines).

For a very weak nebular magnetic field, $B \leq 1\,\mu\mathrm{G}$, a pronounced IC peak is formed around 10 TeV due to the radiative cooling break and the Klein-Nishina effect. With an increase of the magnetic field, the IC peak moves towards low energies and gradually disappears. However, in the case of higher injection rate of electrons in the past, the IC peak survives even for very large nebular magnetic field. Quite interestingly, a similar trend is observed also in the SED of synchrotron radiation; with the increase of the magnetic field, both the synchrotron and IC peaks are shifted towards low energies.

One of the nice features of VHE IC radiation on CMBR photons is its independence of any model parameter, except for the spatial and energy distributions of relativistic electrons. Thus, the high quality gamma-ray data can provide unambiguous information about the relativistic electrons at the present epoch. Coupled with the synchrotron radiation, IC gamma-rays give us precise information about the absolute strength and the spatial distribution of magnetic field over the nebula. They also allow us to explore, under certain model assumptions, the history of electron injection.

Fig. 46 The spectral energy distributions of IC gamma-rays arriving from different regions of the nebula (see the text for details)
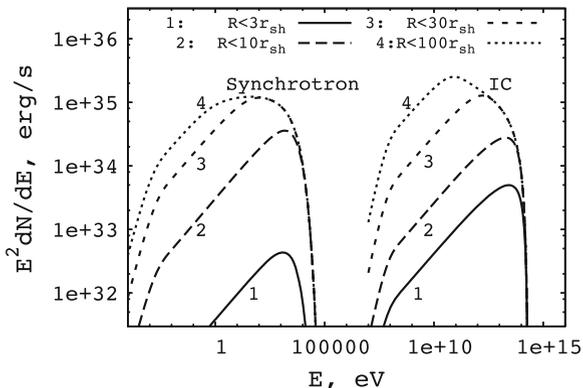
Because of the effects related to the propagation and radiative cooling of electrons, the IC radiation is characterized by strong variation of both the surface brightness and the spectral shape of gamma-rays from different regions of the nebula. It is demonstrated in Fig. 46 for SED of synchrotron and IC components of radiation integrated over four different regions within 3, 10, 30, and 100 termination shock radii. The pulsar's spin-down luminosity is fixed at the level $L_0 = 1.8 \times 10^{37}$ erg/s, and the magnetization parameter $\sigma = 10^{-3}$. The injection spectrum of electron has been assumed power-law, $dN/dE \propto E^{-2.2}$, with an exponential cutoff at 300 TeV. Note that while in the inner region of the nebula ($r \leq 3r_{sh}$) the IC component strongly dominates over the synchrotron radiation, beyond $r = 10r_{sh}$ the contributions of these components become comparable. This is explained by the hydrodynamical amplification of the magnetic field which is implemented in calculations following the formalism of Ref. [166]. While at the shock the field is expected quite week, less than 1 μG (see Eq. (57)), at larger distances it is increases up to a few μG.

Figure 46 shows that the SEDs of IC gamma-rays from the inner regions of the nebula sharply peak at energy $E \simeq 10$ TeV. The reason is that we see the radiation of uncooled electrons which quickly leave the compact region before suffering significant radiative losses even in the case of relatively large magnetic field. This seems to be the case of Vela X, a nearby PWN belongs to the powerful pulsar PSR J0835-4510 of age ≈11,000 year and spin-down luminosity $L_0 = 7 \times 10^{36}$ erg/s. Vela X has been established [41] as one of the strongest TeV gamma-ray sources in the Galaxy. The energy spectrum of the source is quite different from other galactic sources; it is very hard at low energies with photon index $\Gamma \approx 1.5$, and contains a high energy exponential cutoff resulting in a distinct maximum in the SED at 10 TeV (see Fig. 47).

Because of the nearby location of the source ($d \approx 300$ pc) we see, despite the large angular size of the gamma-ray image of order of $1°$ (see Fig. 42), only the central region with a linear size less than several pc. In this regard, Vela X is a perfect object for exploration of processes in the inner parts of the nebula close to the termination shock. The significantly improved sensitivity of the future Cherenkov Telescope Array and its superior angular resolution (1–2 arcmin at 10 TeV) should

Fig. 47 The spectral energy distribution of Vela X obtained with the HESS telescope array. The gamma-ray fluxes are shown per solid angle from two regions: (i) $\Theta < 0.8°$ (*full circles*) and (ii) $0.8° \leq \Theta \leq 1.2°$ (*open circles*)
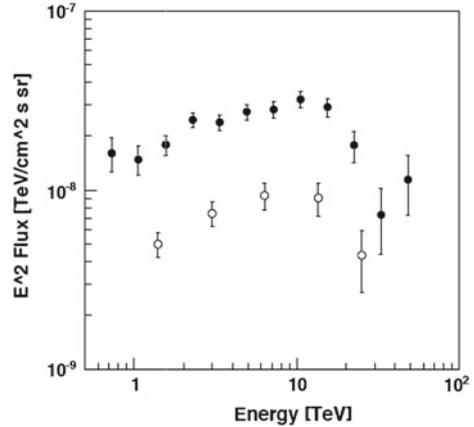
allow unique probes of relativistic electrons inside the region of the termination shock, i.e. just at the heart of the accelerator.

On the other hand, for exploration of distributions of relativistic electrons and magnetic fields beyond the termination shock, we need gamma-ray data up to at least several degrees from the pulsar. In general, to image Vela X on such large angular scales is a hard task for IACT arrays, especially given the expected decrease of the surface brightness of the source, as well as its location in a very complex region in the sky. Presently, the VHE gamma-ray studies of Vela X are limited by the region of angular size $1.2°$ (see Fig. 47). Hopefully CTA will extend observations beyond 2 or $3°$, but even this will be not sufficient for calorimetric studies of the entire source. Instead, the observations of other PWNe in environments of similar properties and powered by pulsars of comparable age and spin-down luminosity, but located at much larger distances, promise more information in this regard. In particular, the PWN MSH 52-15 seems a good counterpart of Vela X, given the spin-down luminosity $L_0 = 1.7 \times 10^{37}$ erg/s of the pulsar PSR B1509-58 that powers MSH 52-15. The TeV gamma-ray image of this source is about $0.2°$ [34] which implies, for the distance to the source of 5 kpc, a linear size of the gamma-ray nebula of about 15 pc.

Another important representative of middle-age PWNe is the TeV gamma-ray source HESS J1825-137, the nebula of the pulsar PSR J1826-1334. A remarkable feature of this source is its unusually large angular size (of order of $1°$; see Fig. 42) despite its far location ($d \approx 4$ kpc) and relatively modest spin-down luminosity of 2.8×10^{36} erg/s. This implies huge linear size of the gamma-ray source extending up to 80 pc which can be interpreted as the result of explosion of the supernova in an extremely low-density environment to allow such a quick expansion of the nebula. An alternative explanation could be that the external parts of the gamma-ray source represent a IC halo produced by the cloud of ultrarelativistic electrons which have left the nebula and propagate diffusively in the interstellar magnetic fields, a scenario discussed in Ref. [52]. High quality data from this source have been recently obtained
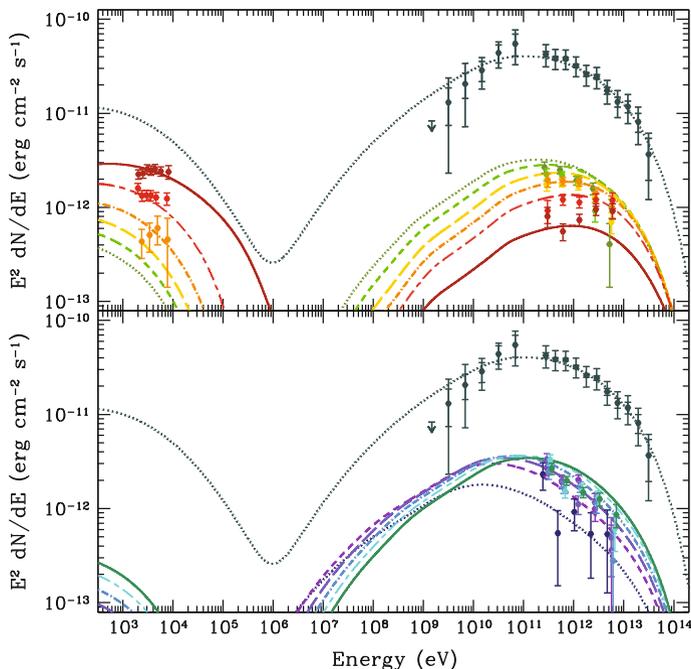
Fig. 48 SED of broad-band IC gamma-ray emission of the pulsar wind nebula HESS J1825-137 calculated for 12 zones with a constant 6 arcmin width of the zones: $0' - 6'$, $6' - 12'$, ... $66' - 72'$. The *theoretical curves* are shown together with observational points obtained with the *Suzaku* (the inner 6 zones), *Fermi LAT* (the entire nebula), and HESS (all 12 zones) telescopes [242]

also in X-ray and GeV gamma-ray bands by the *Suzaku* [238] and *Fermi* LAT [146] telescopes, respectively. Together with the reported spatial and spectral distributions of TeV gamma-rays [40], these data allow detailed multi-zone and time-dependent study of distributions of electrons and magnetic fields in the entire nebula. Such an attempt has been made recently in Ref. [242]; a careful three-dimensional treatment of the problem with inclusion of several important aspects like the magnetic field and velocity profiles, diffusion of electrons, etc., allowed the authors to make a quite strong conclusions [242]. While the energy-dependent profile of the gamma-ray morphology, as well as the significantly smaller size of the X-ray image compared to the gamma-ray image are naturally explained by the effects related to the radiative cooling of electrons, the time-dependent 3D modelling of this source shows for that for explanation of 10 TeV gamma-ray from outer parts of the nebula the diffusion of electrons should be very fast, in particular the time of diffusive escape adopted in the form $\tau_{esc} = 90(R/10\,\mathrm{pc})^2 (E_e/100\,\mathrm{TeV})^{-1}$, provides a very good fit of data shown in Fig. 48. Also, the results of this study demand an initial period of the pulsar close to 13 ms with the braking index $n = 1.9$. This implies an age of the source of about 40,000 year, twice older that thought before. The best-fit model requires

for the interior velocity a profile close to $v(r) \propto v^{-0.5}$, while for the magnetic field it demands $B(r) \propto r^{-0.7}$, with the field falling from a very large value of $400\,\mu G$ at the termination shock to $\approx 2\,\mu G$ at the periphery of the nebula. Also, the model requires quite strong dependence of the magnetic field from the spin-down luminosity, $B(t) \propto \sqrt{L(t)}$, i.e. it should be much higher at the early epochs. Actually, this is a quite significant deviation from the standard MHD picture [166] concerning both the strength of the field at the termination shock and its radial dependence. Obviously, these conclusions derived phenomenologically from the "bet-fit" requirement, need also a thorough theoretical justification. Also, it is possible that one should treat the TeV source as a combination of a more compact pulsar wind nebula and an extended IC gamma-ray halo produced by electrons which already have left the nebula, and diffusively expand in the interstellar medium. But in any case, the results of this work nicely demonstrate the potential of gamma-ray observations, especially in the TeV regime, for understanding the complex phenomena in PWNe.

In summary, the basic concept of formation of PWNe due to termination of ultrarelativistic electron–positron pulsar winds is strongly supported by X-ray and gamma-ray data both in GeV and TeV bands. In the case of most of TeV PWNe, the inverse Compton origin of radiation is the *only* real option to explain the reported fluxes, although in a few cases, e.g. for Vela X [158], the hadronic origin of gamma-ray emission formally cannot be fully excluded (see for a review Ref. [92]). However, the hadronic interpretation of TeV emission not only implies dramatic revision of the generally accepted paradigm of PWNe, but it is a redundancy and, more importantly, is not supported by observations.

8 Gamma-Ray Loud Binaries

In a binary system containing a relativistic object, a neutron start or a black hole, the gravitational energy of the compact object is transformed into thermal emission (typically in the X-ray band) radiated away by the hot accretion plasma. Therefore these objects traditionally are treated as sources of thermal radiation. However, they can operate also as particle accelerators. In fact, several different scenarios of effective particle acceleration can be realized in these systems linked to either the termination of the pulsar wind or to the internal shocks in the jet formed in the vicinity of the black hole. Moreover, thanks to the presence of dense target material provided by the companion star in the form of optical photons or gas, favourable conditions for gamma-ray production can be established as well.

The history of developments related to gamma-radiation of binary systems is rather controversial (see for a review Refs. [27, 247]). It is almost forgotten that in the 1980s these objects constituted the highlight source population of ground-based gamma-ray astronomy. The reports by several independent groups on detection of modulated TeV and PeV signals from Cyg X-3, Her X-1 and some other similar objects [247] initiated a significant interest to VHE gamma-ray astronomy. However, since the next generation of ground-based instruments with significantly

improved performance failed to confirm the early claims, these objects have not been anymore treated as important targets for gamma-ray observations. This stance has again changed after the discovery of galactic sources with relativistic jets, dubbed *microquasars* [193]. A clear message of this discovery was the important role of non-thermal processes, and hence possible effective gamma-ray production in these accretion-driven objects.

8.1 Microquasars: Not yet Proved TeV Emitters

During strong flares, the non-thermal power of synchrotron jets of microquasars (in the form of accelerated electrons and kinetic energy of the relativistic outflows) could be comparable or even exceed the thermal X-ray luminosity of the central black hole. The jets in microquasars are typically trans-relativistic, thus the acceleration of electrons can proceed, in principle, at a very high rate. Correspondingly the spectrum of synchrotron radiation of the jet can extend to hard X-rays [84, 191]. On the other hand, the high density photon fields provided by the accretion disk around the compact object, by the companion star, as well as produced by the jet itself, can create favourable conditions for effective production of high energy IC gamma-rays [84, 182, 215] provided that the ratio of energy densities of accelerated electrons and the magnetic field does not significantly deviate from the equipartition condition. Generally, this radiation is expected to have an episodic character associated with strong radio flares in objects like the classical microquasar GRS 1915+105.

The shocks in microquasars can accelerate protons as well. But the gamma-ray production through *pp* interactions may be effective only in high gas density regions (see for a review Ref. [107]). In particular, the old "*atmospheric target*" (bombardment of the normal star's atmosphere by a relativistic beam of particles accelerated at the compact object) or "*target crosses beam*" (interaction of a moving gas target with a beam of relativistic particles) scenarios can provide non-negligible production of high energy gamma-rays of hadronic origin. Although these models had been inspired by the reports of detection of VHE signals from Cyg X-3, Her X-1, which later were not confirmed (and eventually, discredited), the models themselves still present a certain conceptual interest as possible scenarios of efficient gamma-ray production.

Despite the several proposed viable scenarios of gamma-ray production in microquasars, and intensive searches for VHE gamma-ray signals from some famous representatives of this source population, we do not have yet a solid evidence in favour of particle acceleration and their radiation in the VHE regime. A possible exception is the marginal signal of VHE gamma-rays reported by the MAGIC collaboration from Cyg X-1 [74]. However the *astrophysical* origin of the claimed episodic event obviously needs further confirmation. Interestingly, a tentative (not yet confirmed) detection of GeV gamma-ray flare of Cyg X-1 has been claimed by the AGILE collaboration [216]. It is difficult to overestimate the astrophysical significance of these results in the case of their confirmation. AGILE and *Fermi* LAT telescopes are per-

fectly designed for such studies at GeV energies. In this regard, the detection by both instruments of several flaring episodes from another famous microquasar, Cyg X-3 [4, 233], is a remarkable observational achievement with important astrophysical implications (see e.g. Ref. [251]). On the other hand, the search for random/episodic VHE gamma-ray events from variable sources with small FoV Cherenkov telescopes is a quite hard task. Continuous monitoring of these source with future low-energy threshold particle arrays like HAWK, seems to be a more promising approach.

8.2 Binary Pulsars

The second type of binary systems predicted as potential VHE gamma-ray emitters are the systems which contain high spin-down luminosity pulsars [173, 232]. These sources work as a compact pulsar wind nebula located in an environment where the radiation and gas densities, as well as the magnetic field pressure are enhanced by more than ten orders compared to the interstellar medium. Correspondingly, all processes related to the termination of the shock, particle acceleration and radiation proceed in much shorter time-scales. The best candidate representing this class of sources for gamma-ray observations is PSR B1259-63/LS2883—a binary system consisting of a 48 ms pulsar in highly eccentric orbit around a massive companion star with a period 3.4 years. PSR B1259-63/LS 2883 was observed by the HESS array of telescopes around three periastron passages in 2004, 2007 and 2010 with detection at high statistical level on all three occasions [35, 47]. These observations revealed a variable character of gamma-radiation. Three observations have been conducted mainly either before or after the periastron passage. But fortunately some overlaps of same orbital phases was possible indicating on the general repetitive behaviour of the source. The cumulative lightcurves based on these observations periastron show a hint of two peaks around the periastron with a tendency of flux decrease towards the periastron. Unfortunately the source never has been observed exactly at the periastron passage (it was possible in 2004, but unfortunately, due to the moon and bad weather, such an opportunity was missed). The positions of the gamma-ray peaks coincide with the eclipse of the radio pulsed emission. This indicates that the origin of these peaks should have a relation to the position of the circumstellar disk. The inverse Compton scattering seems to be the most plausible gamma-ray production mechanism. However, while the measured spectrum is close to expectations [173], neither the flux minimum at the periastron, nor the enhanced fluxes before and after the periastron passage agree with the theoretical predictions [173] for the lightcurve. Apparently, for interpretation of these results we have to invoke more sophisticated scenarios than it was anticipated before.

Figures 49 and 50 show the variation of the VHE flux with the change of the true anomaly and the overall differential energy spectrum, respectively. Despite the significant changes of the absolute flux, no noticeable spectral variability was found, in particular on time-scales of months, for the data taken between April and August 2007 [47]. The geometry of the system has an impact on the lightcurve through the
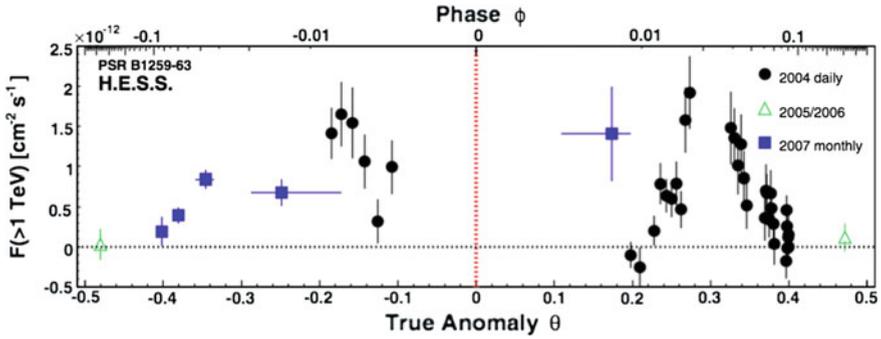
Fig. 49 Gamma-ray fluxes from PSR B1259-63 above 1 TeV as a function of the true anomaly. The data are from the HESS observations in 2004 (daily fluxes; *black points*) and 2007 (monthly fluxes; *blue squares*) [35, 47]. The *red vertical line* indicates the periastron passage. Also are shown the level of the overall flux in 2005 and 2006 (*green triangles*)
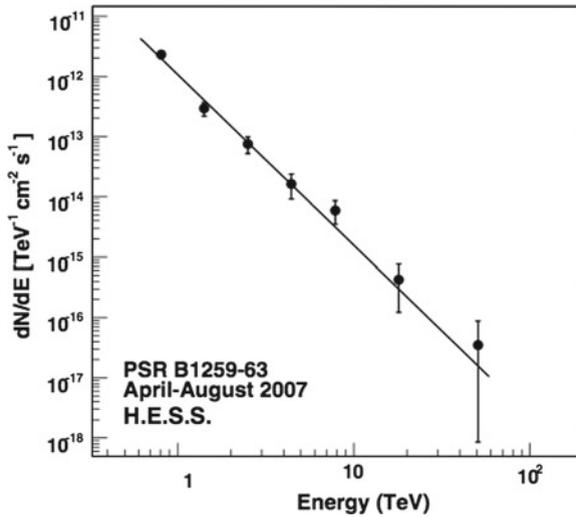
Fig. 50 The differential energy spectrum of PSR B1259-63 based on 52.5 h observations between April and August 2007. The spectrum is described by a power-law with flux normalisation $F_0 = (1.1 \pm 0.1_{stat} \pm 0.2_{sys}) \times 10^{-12}$ TeV^{-1} cm^{-2} s^{-1} and photon index $\Gamma = 2.8 \pm 0.2_{stat} \pm 0.2_{sys}$ [47]

anisotropic Compton scattering as the angle between the line of sight and the line connecting stars changes with the orbital phase. On the other hand, several factors can force the lightcurve of the IC signal to deviate from predictions based on the pure geometrical treatment of the problem. A strong impact on the radiation of the system has the variation of the magnetic field as well as non-radiative losses due to the expansion of the shocked region. The enhanced synchrotron losses due to the larger magnetic field close to the periastron could, in principle, explain the minimum

of the IC flux around the periastron passage. This should lead to an increase of the synchrotron X-radiation which, however, is not observed. Just opposite, the X-ray lightcurve contains a clear minimum at the periastron. It has been claimed [170] that the dominance of adiabatic losses could be a major reason causing the tendency of reduction of both X-ray and gamma-ray fluxes when the pulsar appears in the close vicinity of the companion star during the periastron passage. The analysis of the HESS data show that the TeV fluxes depend on the binary separation and hint at a peculiar non-radiative cooling profile around the periastron [167]. The adiabatic losses can be naturally linked to the interaction of the pulsar and stellar winds. The modelling of hydrodynamics of such interactions [103] confirms that adiabatic losses may indeed dominate over the radiative losses.The same interactions may lead to formation of relativistic outflows. The detailed hydrodynamic and MHD simulations revealed [103, 104] that already at distances comparable to the size of the binary system the bulk Lorentz factor of the shocked flow can be as large as $\Gamma \simeq 4$. This may result in significant anisotropy of radiation of accelerated electrons caused by Doppler boosting. Indeed, since the Doppler factor depends on the location of the pulsar along its orbit, it can modulate the IC gamma-ray flux [170], and thus lead to reduction of X-ray and TeV gamma-ray signals close to the periastron.

Finally, one should mention another possible reason for suppression of the X- and TeV gamma-ray emission; it can be caused by an "early" (sub-TeV) cutoff in the energy spectrum of electrons due to the enhanced Compton losses around the periastron [170]. An interesting feature of this interpretation is the expected anti correlation of the GeV and TeV gamma-ray fluxes. In particular it predicts an increase of the GeV flux in a narrow time-window around the periastron. The detection of the source by *Fermi* LAT with a modest energy flux (a few times 10^{-11} erg/cm^2 s) close to the periastron passage [10] agrees quite well with this prediction, but unfortunately the marginal gamma-ray signal as well as the inconsistency between spectra reported by two independent groups [10, 226] does not allow a certain conclusion in this regard. There is a hope that the TeV and GeV observations of the source during the next periastron passage of the source will significantly enhance our knowledge on temporal and spectral features of this source and correspondingly remove many uncertainties related to the conditions in the gamma-ray production region. What concerns the mechanism of production of gamma-rays, the arguments in favour of its inverse Compton origin seem quite solid. The detection of TeV gamma-rays approximately 50 days before the periastron passage in 2007 disfavours the stellar disk scenario [47], and therefore the hadronic origin of gamma-rays, at least within the current concept and knowledge regarding the stellar disk inclination, extension and the density profile.

Spectacular Flare of PSR B1259-63/LS2883

As discussed in Sect. 7.2, pulsars can produce potentially detectable gamma-ray emission also due to the bulk Comptonization of their cold ultrarelativistic winds. While this radiation component of isolated pulsars is generally weak, unless the wind

is accelerated (relatively) close to the pulsar [100], in binary systems the radiation is significantly enhanced because of the presence of target photons supplied by the optical companion [88, 169, 170]. On the other hand, while IC radiation of isolated pulsars can be readily separated from the extended synchrotron and IC emission of the surrounding nebula with distinct spectral and spatial features, in binary pulsars the separation of the radiation components produced before and after the termination of the pulsar wind is a more difficult task; it requires rather careful treatment of the spectral and temporal features of two radiation components. In this regard, the most promising object for exploration of processes of formation, acceleration and termination of pulsar winds in binary systems is PSR B1259-63/LS2883. While the observations with the HESS telescopes allowed a meaningful constraint of the wind's Lorentz factor, $\Gamma \leq 10^6$, the main hope for detection of IC radiation of the unshocked wind has been related to *Fermi* LAT [170].

The observations of PSR B1259-63/LS2883 with *Fermi* LAT around the periastron passage in 2010–2011 revealed a rather complex behaviour of the system in GeV gamma-rays [10, 226]. Close to the periastron passage, a rather weak signal has been detected which could be related to IC radiation of the wind—both before and after its termination [169]. Surprisingly, 30 days after the periastron passage, a spectacular flare has been recorded. It lasted approximately two weeks with a flux enhanced by more than an order of magnitude [10]. The flare is characterized by a sharp increase and a smoother decay over approximately 2 weeks. During the flare, gamma-ray luminosity has been increased to an extraordinary level, very close to the pulsar's spin-down luminosity $L = 8 \times 10^{35}$ erg/s. It is remarkable that such an extraordinary event was not accompanied by any noticeable change of flux at any other wavelength! Without any doubt, this flare represents a unique case in astrophysics when the available energy is fully converted to non thermal high energy radiation. On the other hand, in accordance with the current concept that the rotational energy of the pulsar is converted to kinetic energy of cold ultrarelativistic electron–positron wind, the binary pulsars can in principle operate as perfect ('100 %' efficient) gamma-ray emitters if the density of the surrounding target photon field would be high enough for realization of Comptonization of the pulsar wind in the saturation regime. Recently a model has been suggested which under certain assumption can address the most outstanding features of this unique event by radiation of the unshocked pulsar wind with a Lorentz factor $\Gamma_0 \simeq 10^4$. The combination of two effects both linked to the stellar disk, is a principal element in the proposed model. The first effect is related to the impact of the surrounding medium on the termination of the pulsar wind. Inside the disk, the wind can suffer an "early" termination resulting in suppression of its gamma-ray luminosity. When the pulsar escapes the disk, the conditions for termination of the wind undergo significant changes. This would lead to a dramatic increase of the pulsar wind zone, and thus to the proportional increase of the gamma-ray flux. On the other hand, if the parts of the stellar disk disturbed by the pulsar can supply infrared photons of density high enough for efficient (saturated) Comptonization of the wind, almost the entire kinetic energy of the pulsar wind could be converted to radiation, and thus the gamma-ray luminosity of the wind would achieve the level of the pulsar's spin-down luminosity.
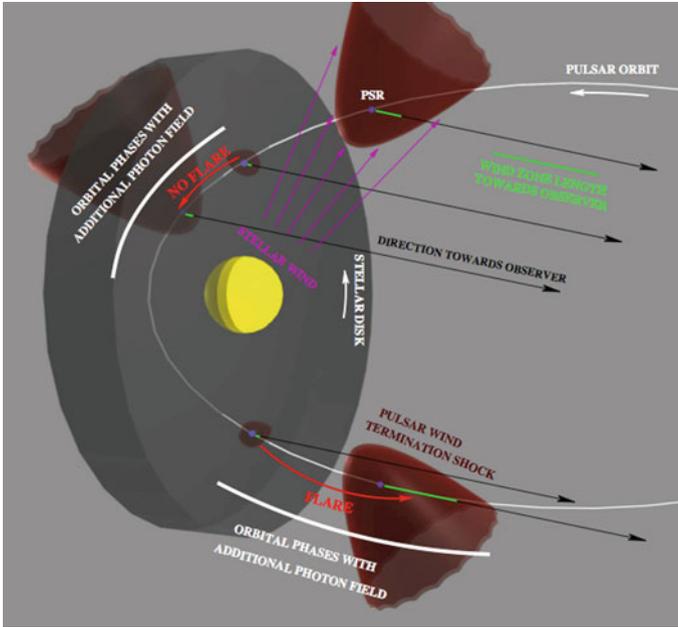
Fig. 51 The sketch of the scenario explaining the gamma-ray flare of PSR B1259-63/LS2883
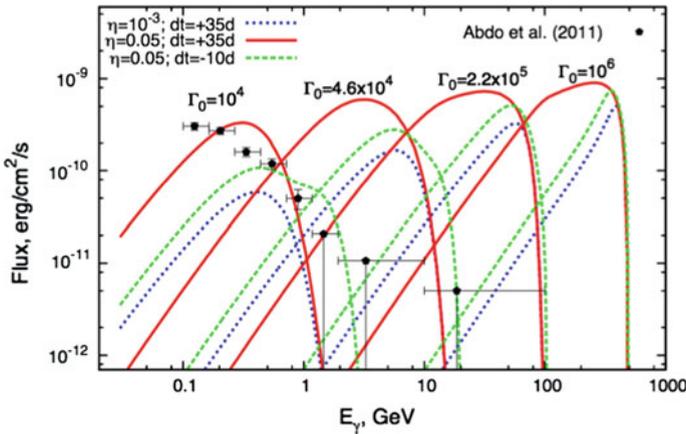
Fig. 52 SED of IC radiation of the unshocked pulsar wind at the epoch of 35 days after periastron passage. Calculations for different values of the ratio of ram pressures η of the interacting pulsar and stellar winds ($\eta = 0.001$—*dotted lines*; and $\eta = 0.05$—*solid lines*) are shown. The calculations are performed for two target photon fields: (i) radiation of the companion optical star and (ii) radiation of the stellar disk. Several initial pulsar wind bulk Lorentz factors indicated at the *curves* have been assumed. The spectral points reported by the *Fermi* LAT for the post-periastron flaring episode are shown with *pentagons*

The sketch of the proposed scenario is shown in Fig. 51. Calculations of spectral energy distributions of IC radiation of the unshocked wind performed within this scenario are presented in Fig. 52. The calculations for two key model parameters, the Lorentz factor of the wind $\Gamma_0 = 10^4$ and the ratio of ram pressures of the pulsar and stellar wind pressures $\eta = 0.05$ explain quite well the measurements of *Fermi* LAT.

If the proposed interpretation is correct, this would be the second case, after the Crab pulsar [62], of direct measurement of a pulsar wind's Lorentz factor. However, one should mention that there is another possible explanation of this extraordinary flare, namely by the Doppler boosted radiation of the post-shock flow (see e.g. Ref. [177]). Since the reported gamma-ray flux was close to pulsar's spin-down luminosity, strong boosting is a principal conditions of this interpretation Formally, this is not an extreme assumption as it can be look at the first glance. Relativistic flows can be formed at interactions of the pulsar and stellar winds [103], and therefore the broad-band radiation can be strongly affected by the Doppler boosting [128, 171]. The latter should amplify also the X- and TeV gamma ray fluxes. However, the lack of any noticeable activity during the flare at other wavelengths makes this interpretation less likely. Note that the apparent luminosity of the Doppler boosted radiation can easily exceed the limit set by the pulsars spin-down luminosity, but so far such flares have not been detected. On the other hand, the detection of such events in future observations would rule out the origin of gamma-radiation related to the unshocked pulsar wind.

8.3 Enigmatic TeV Binaries

Currently, the population of galactic binary systems detected in TeV gamma-rays consists of a few objects. In addition to PSR B1259-63/LS 2883, TeV gamma-rays have been reported from LS 5039 [36, 38] and LS I+61 303 [17, 18, 71]. Recently, one more object joint the club of "TeV binaries", HESS J0632+057. The source was discovered serendipitously in the survey of the galactic plane by HESS. Because of co-location with a variable radio and X-ray source at the position of the massive star MWC 148, this unidentified TeV source has been suggested [155] to resemble the TeV binary systems like LS I+61 303 or LS 5039. The subsequent gamma-ray observations with VERITAS and HESS [188] revealed variability in the TeV gamma-ray flux. Moreover, the X-ray observations with Swift [105] revealed a binary nature of the source with periodicity of $\approx$320 days. No GeV gamma-rays have been yet found from this source.

Except for PSR B1259-63/LS 2883, the nature of the compact stars in TeV binaries remains highly unknown. LS I+61 303 and LS 5039 often are treated as *microquasars*, i.e. objects with relativistic jets powered by accretion onto a black hole. However, unlike the classical representatives of this source population, e.g. GRS 1915+105 and Cyg X-1, these two objects are not univocally accepted as microquasars. They indeed show features quite different from classical microquasars. Therefore it has been advocated that they are analogues of the binary pulsar

PSR 1259-63 (see e.g. Ref. [126]). and therefore their non thermal activity is initiated by pulsar winds.

The intensive searches over the last years could not, however, reveal pulsed signals from these objects. While the radio emission could not show up, in principle, because of free-free absorption, the failure of detection of pulsed X-rays [211] can be explained by the beam pointing away from the line of sight of observer. However, even in this case it would be not so easy to hide the signatures of pulsars. For example, at interaction with the pulsar wind, the stellar wind will be heated and emit thermal X-ray emission. The lack of detection of thermal X-ray emission from LS 5039 implies a robust upper limit on the power of the pulsar wind [250] at the level close to the luminosity of GeV gamma-rays reported by *Fermi* LAT [3]. This would imply an extremely high efficiency of conversion of pulsar's spin-down luminosity to GeV gamma-rays. Alternatively one may assume that the gamma-radiation is strongly Doppler-boosted which, of course, cannot be *a priori* excluded.

The lack of thermal X-ray emission has been also invoked as an argument against the accretion origin of the power of compact objects in LS I+61 303 and LS 5039. However, an interesting recent study of wind accretion on a rotating black hole [90] in close binary systems shows that angular momentum of the accreted stellar wind might be not sufficient for formation of accretion disk, thus the X-ray emission would be suppressed. On the other hand, the direct wind accretion can activate the Blanford-Znajek process, and, as a result, a powerful jet can launch without accompanying thermal X-ray emission. It has been found that approximately 10 % of the accretion power could be channelled into the jet and initiate observable non thermal emission. Yet, in the case of LS 5039, one would need to invoke additional resources, in particular Doppler boosting or significant anisotropy of emission to explain the luminosity of reported GeV gamma-rays from this object.

The ambiguity related to the nature of the compact object in these binary systems limits, but does not prevent detailed phenomenological studies of particle acceleration and radiation processes in these complex systems. Over the last several years, the origin of TeV gamma-ray emission of LS I+61 303 and LS 5039 has been discussed in the context of their multi wavelength properties, under different assumptions concerning the radiation and absorption processes, the sites of particle acceleration and gamma-ray production, etc. Both sources have their specific feature, and should be treated quite differently. In this regard, one should mention that LS I+61 303 is a more "problematic" target for theoretical studies, despite the fact that it is one of the deepest observed VHE sources. Due to the low fluxes, inconvenient for observations orbital period and unpredictable behaviour at different phases, it is "much less well characterized" [157] than LS 5039.

LS 5039 is a binary system consisting of a massive star of huge optical luminosity ($\simeq 10^{39}$ erg/s) in a 3.9 day orbit around a compact object. It shows non thermal radio emission resolved as a bipolar-jet like structure with a speed $v \approx 0.2c$ [204]. TeV radiation from this source has been serendipitously discovered in 2005 during the galactic plane survey by HESS [36]. The subsequent deep observation of the source revealed a periodic nature of the signal—the source behaves as a perfect "TeV gamma-ray clock" [38]. Later, periodic gamma-ray signal has been found by *Fermi*
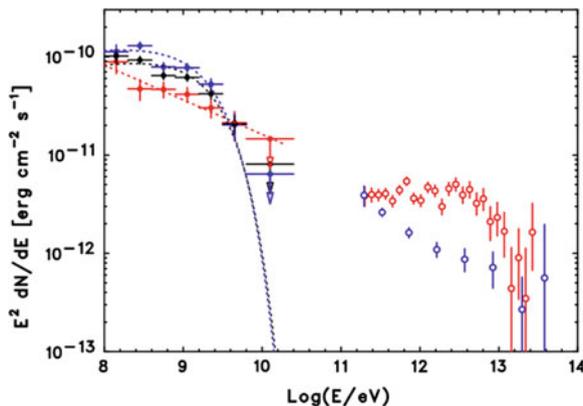
Fig. 53 High and VHE gamma-ray spectra of LS 5039 measured by HESS [38] and *Fermi* [3] for two broad orbital phase intervals defined as **INFC** $0.45 < \phi \leq 0.9$ (*red points*) and **SUPC** $\phi \leq 0.45$ & $\phi > 0.9$ (*blue points*). The *black points* and the *dotted line* represent the phase-averaged *Fermi* LAT spectrum

LAT also at GeV energies [3]. Both the GeV and TeV radiation components are modulated with 3.9 day period which coincides with the orbital period of the system derived from the Doppler-shifted optical lines. The major fraction of gamma-ray emission is contributed by the half of the orbit with a maximum close to the so-called *inferior conjunction* (the epoch when the compact object is lined-up along our line-of-sight in front of the optical star). Figure 53 shows the energy spectra of the source averaged in two broad phase intervals: (i) $0.45 \leq \phi \leq 0.9$ and (ii) $\phi \leq 0.45$ & $\phi \geq 0.9$. While the spectrum corresponding the phase interval (i) can be fitted by a hard power-law with a photon index ≈ 1.85 and exponential cutoff around 9 TeV, the interval (ii) is characterized by a quite different spectrum—a single power-law with a photon index ≈ 2.5.

The time variation of the absolute flux of TeV gamma-rays can be partly explained by the photon–photon absorption caused by the variation of the geometry and density of the stellar radiation field [108, 125]. However, the absorption effect alone hardly can explain the observed modulation of the energy spectrum of gamma-rays. Indeed, while for the given temperature of radiation of the companion star $kT = 3.5$ eV, the gamma-ray absorption should result in strongest modulation of the flux at low energies, $E \leq 1$ TeV, the HESS data show just opposite behaviour—a quite stable gamma-ray flux at 0.2 TeV, but strongly variable emission at energies above 1 TeV. Obviously, photon–photon pair production cannot be responsible for modulation of GeV gamma-rays either; the energy of these gamma-rays is well below the interaction threshold with optical photons. Thus, it is clear that additional processes should be invoked to explain the observed time variation of the energy spectrum and the absolute flux. In particular,

- variation of the maximum energy of accelerated electrons established in a radiation dominated environment may have a strong impact on the temporal behaviour of TeV gamma-ray spectra [38, 170]. While this effect may lead to a modulation of synchrotron X-rays produced, most likely, by the same population of electrons, its impact on the lightcurve of IC gamma-rays at GeV energies is quite small;
- the effect related to the cross-section of anisotropic Compton scattering, undoubtedly plays an important role in modulation of both GeV and TeV gamma-ray signals [38, 127, 168];
- the lightcurves of both synchrotron and IC components of radiation may be affected by hydrodynamic effects like adiabatic losses due to expansion of the production regions [225] and Doppler boosting of radiation due to relativistic outflows formed at the interaction of pulsar and stellar winds [128, 171].

The combination of all these effects results in rather complex correlations between different radiation components. Figure 54 demonstrates that the TeV and X-ray lightcurves can be described quite well by the same population of electrons under certain assumptions regarding the adiabatic and radiative losses of electrons. In Fig. 54 the phase dependence of adiabatic losses are derived phenomenologically from the Suzaku data. Although the adiabatic losses modulate the IC gamma-ray flux in a similar manner, the TeV gamma-rays unlike X-rays suffer significant distortion due to photon–photon absorption. It is likely that these two additional effects related to interactions of gamma-rays and electrons with the stellar radiation field are responsible for stronger changes of VHE gamma-ray flux compared to the X-ray flux.

Within any IC model of gamma-ray emission of LS 5039, the spectrum continues down to low energies with a peak around 10 GeV, unless one assumes a sharp *low-energy cutoff* in the energy spectrum of electrons. In the case of development of electromagnetic cascades triggered by photon–photon pair production, a maximum in SED appears close to 30 GeV, independent of the spectrum of primary electrons [26]. Therefore the report of the *Fermi* collaboration about detection of high energy gamma-rays was not a big surprise, especially given the detected change of the signal with orbital phase [3], which suggests a link to the VHE component of radiation. Nevertheless, the reported high flux and the energy spectrum with a cutoff at 2 GeV (see Fig. 53) challenge most of models suggested in the literature. The proposed in Ref. [3] interpretation that the presence of a cutoff in the spectrum may be indicative of magnetospheric emission, does not seems a realistic option since it cannot explain the origin of modulation of the signal. The GeV signal challenges also the non thermal energy budget of the source. In this regard, a possible Doppler boosting can help significantly to reduce the energetic requirements to the source.

The lack of information about the nature of the compact object in LS 5039 prevents us from comprehensive modelling of high energy processes in the source. However, a number of important conclusions can be derived from a general analysis of conditions concerning the acceleration of electrons and the radiation/absorption processes of gamma-rays. For example, the sheer fact of detection of gamma-rays with energy exceeding 10 TeV implies an extremely fast acceleration of electrons close to the theoretical limit; the parameter η in Eq. (28) which characterizes the acceleration
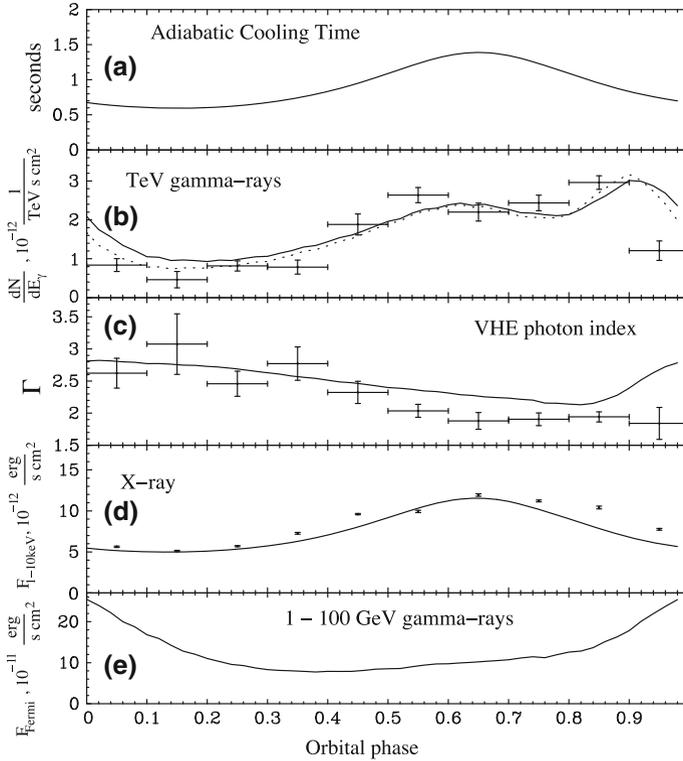
Fig. 54 *Light curves* characterizing LS 5039 [225]. **a** Phase dependence of adiabatic losses derived directly from the Suzaku data; **b** theoretical predictions for 1 TeV gamma-ray fluxes together with HESS data points; **c** calculations for photon indices around 1 TeV together with power-law indices reported by HESS; **d** calculated 1–10 keV synchrotron X-ray fluxes shown together with the Suzaku data; **e** predicted IC flux integrated over the interval 1–100 GeV. The total power injected in relativistic electrons with E^{-2} type spectrum is fixed at the level of 10^{37} erg/s. The cooled electron spectrum is formed due to radiative (IC and synchrotron) and non-radiative (adiabatic) losses

efficiency cannot significantly exceed 10 [168]. Even so, the accelerator should be located at periphery of the binary system, otherwise the severe energy losses of electrons in high density radiation field of the companion star will prevent acceleration to multi-TeV energies.

An independent conclusion regarding the location of the TeV gamma-ray production region can be derived from the analysis of the allowed level of gamma-ray absorption. If the gamma-ray emitter is located close to the compact object (i.e. deep inside the binary system), the magnetic field of the companion star should be less than a few Gauss. This is a necessary condition to allow pair cascades to develop and thus to reduce effectively the photon–photon absorption opacity. However, the magnetic field in O stars is expected to be much larger. For these (more realistic) fields the development of the cascade is suppressed; the major fraction of the energy

absorbed in the stellar photon field will be emitted by secondary pairs via synchrotron radiation in the X-ray band. In order to avoid overproduction of X-rays (compared to the detected fluxes), the TeV gamma-ray emitter should be located at distances $\geq 10^{12}$ cm from the compact object. This excludes the standard pulsar scenarios suggested for LS 5039 in which the emitter is located close to the pulsar (see e.g. Ref. [127]), but does not exclude the pulsar origin of the compact object. The supersonic outflow formed at interaction of pulsar and stellar winds, may effectively transport the available energy to the periphery of the binary system. Alternatively the energy of the compact object can be transported by a jet powered by radiatively inefficient stellar wind accretion onto a rotating black hole [90].

In summary, further multi-wavelength studies are needed to shed light on the nature of the compact object in LS 5039. The search for distinct pulsar or black-hole signatures seems to crucial in this regard. However, we perhaps should be prepared for discoveries of new observational features of this unusual object which could require a dramatic revision of standard concepts of current paradigms.

9 Summary

The remarkable success of observational gamma-ray astronomy, together with recent intensive theoretical and phenomenological studies of acceleration and radiation processes in astrophysical environments, supply a strong rationale for the further exploration of the sky at high and very-high energies.

In this chapter I tried to highlight the major results and advances in *galactic* VHE gamma-ray astronomy. The achievements in the *extragalactic* VHE gamma-ray astronomy are equally impressive with very important astrophysical and cosmological implications.[1]

The recent observational results have introduced non-negligible corrections and additions to the major scientific objections and motivations of gamma-ray astronomy. In particular, new interesting features have been revealed which require significant revisions of current theoretical models or even formulations of new concepts. It is expected that over the next decade the ongoing operation of *Fermi* LAT will be accompanied by observations with the next generation ground-based detectors, in particular CTA and HAWK. The data obtained in the enormous energy range from 100 MeV to 1 PeV will provide deep insight into a number of problems of high energy astrophysics and fundamental physics. In the area of galactic gamma-ray astronomy, I would put an emphasis on the following major objectives:

Identification of SNRs as the major source population responsible for the production of galactic cosmic rays. Since these sources should operate, by definition,

---

[1] A few selected topics linked to the origin of VHE gamma-ray emission of blazers and to the propagation of gamma-rays through the extragalactic radiation fields, have been partly covered in my Saas-Fee lectures. However, because of the chosen format and the imposed page limit, I decided not include them in this chapter.

as PeVatrons, the extension of gamma-ray observations up to 100 TeV and beyond seems to be the key condition for the success. Another breakthrough is expected from the improvement of the point spread function (PSF) of the atmospheric Cherenkov telescope arrays. An angular resolution of about 1 arcmin seems to be crucial, especially in the case of observations of young SNRs. Fortunately, such resolution in principle is achievable in the most important, $E \geq 10$ TeV energy band. This should allow purposeful searches for radiative signatures of PeV protons beyond the shells of young SNRs. Since the highest energy particles escape the shell typically after a few hundred years or even earlier, the radiative signatures of PeV particles are expected from the regions just beyond the shell rather than from the shell itself. The observations with an arcminute resolution should also allow localization of the gamma-ray production regions associated with dense core clumps in the shells, as well as with reverse shocks. Finally, the improvement of the sensitivity of the future IACT arrays by an order of magnitude should allow an extension of the distances to SNRs detectable in TeV gamma-rays up to 10 kpc. Thus, one may anticipate a dramatic increase of the number of TeV emitting young SNRs, many of which without previously detected counterparts at other wavelengths!

Search for other principal contributors to the galactic cosmic rays. Deep multiwavelength studies of the so-called "dark" (so far not identified) TeV gamma-ray sources may reveal a new population(s) of effective particle accelerators. We may anticipate also discovery of a large number of new "dark" TeV emitters in the surveys with the next generation ground-based detectors. The arrays of (relatively) large-field-of-view Cherenkov telescopes designed for detection of gamma-rays with angular resolution of order of 1 arcmin, primarily in the $E \geq 10$ TeV energy band, can be especially effective for serendipitous surveys of the galactic plane. Together with the ultrahigh energy gamma-rays, the hard X-ray imaging telescopes aboard the recently launched NuSTAR and the future ASTRO-H satellites, should be able to conduct effective searches of currently active PeVatrons through the synchrotron radiation of secondary electrons produced in hadronic interactions. A complementary information will be provided by *Fermi* LAT at lower energies. In particular, the gamma-ray observations from massive molecular clouds which operate as unique cosmic-ray barometers, can reveal the "smoking guns"—remnants of ancient (already faded away) PeVatrons, as well as provide an information about the environments which harbor these most effective cosmic ray accelerators in our Galaxy.

Physics of pulsar winds and pulsar wind nebulae. The large number of TeV emitting PWNe powered by pulsars of different age and spin-down luminosities provide an excellent observational material for comprehensive studies of the physics of formation and termination of ultrarelativistic pulsar winds. The unique feature of TeV gamma-ray emission is the dominance of the inverse Compton scattering of electrons on the photons of 2.7 K CMBR. This allows models-independent derivation of the energy spectra and spatial distributions of parent electrons. Combined with the X-ray data, accurate maps of magnetic fields can be extracted, at least for the central compact regions where X-ray data are typically available. In this regard, the nebula of the

Vela Pulsar is of special interest because of its proximity and high spin-down luminosity. The improvement of the flux sensitivity by an order of magnitude coupled with the angular resolution of the order of $\sim$1 arcmin achievable by the future atmospheric Cherenkov telescope arrays at energies above 10 TeV, should allow unique probes of relativistic electrons and magnetic fields inside the termination shock! Similar studies are possible also for the Crab Nebula, but invoking the temporal characteristics of the synchrotron gamma-ray emission during the recently established quasi-regular flares in the MeV–GeV band. The realization of these possibilities will provide a key insight into the magnetohydrodynamics and particle acceleration processes at the termination of pulsar winds resulting in formation of nonthermal synchrotron and IC nebulae. On the other hand, the detailed studies of energy-dependent morphologies of distant PWNe with large angular extensions (like HESS J1825-137), will greatly contribute to the understanding of the escape of relativistic electrons from the nebula and their diffusive propagation in the interstellar magnetic fields. What concerns the formation of the cold ultrarelativistic pulsar winds and exploration of their characteristics, in particular the location and maximum energy of the bulk motion acceleration, a unique and powerful method in this regard could be the combination of detailed temporal and spectroscopic studies of the characteristic pulsed gamma-ray emission of the Crab and perhaps some other high spin-down luminosity pulsars at energies from 10 GeV to a few TeV. The practical realization of such studies can be conducted by CTA with significantly improved (compared to the current telescope systems) flux sensitivities in the energy interval from tens of GeV to a few TeV.

Solving puzzles and revealing origin(s) of VHE binaries. So far, only several compact/variable VHE gamma-ray sources have be discovered in the Galaxy. The nature of these objects, except for PSR B1259-63/LS 2883, remains highly uncertain. PSR B1259-63/LS 2883 is firmly identified as a binary pulsar in which the processes of formation of the pulsar wind, and its termination with the subsequent particle acceleration proceed quite similarly to PWNe, but on much (orders of magnitude) shorter timescales. Therefore, this source offers a unique opportunity to study the complex MHD and particle acceleration processes in "on-line" regime. Although the gamma-ray fluxes from this object are weak both at GeV and TeV energies, for comprehensive temporal studies with *Fermi* LAT and HESS, the available data obtained during the epochs around the periastron of the system revealed significant deviations from the early predictions. This concerns the tendency of a minimum TeV flux at the periastron, the surprise GeV flare after several weeks of the periastron passage, etc. Despite the attempts to explain the strange behavior of the source, it is clear that a detailed theory can be developed only based on the higher quality gamma-ray data. The other TeV binaries are even more enigmatic. Although originally dubbed as microquasars, they show features quite different from the classical binary systems containing black holes. Therefore they have been advocated to be binary pulsars. On the other hand, no pulsed signals have been observed from these sources at any wavelength. A more exciting scenario for these objects could be a specific regime of accretion onto black hole—without formation of an accretion disk, but with a launch

of a powerful jet, e.g. due to the Blandford-Znajek process. One may anticipate that intensive theoretical studies will be conducted in both directions, but it is likely that a decisive answer on the origin of these objects in general, and on the nature of acceleration and radiation processes in particular, will come from new observations with the next generation of ground-based detectors. The performance of CTA concerning both the broader energy coverage and higher flux sensitivity, will be adequate (hopefully) to address all these issues.

References

1. Abdo, A.A., Ackermann, M., Ajello, M., Atwood, W.B., Axelsson, M., Baldini, L., Ballet, J., Barbiellini, G., Baring, M.G., Bastieri, D., et al.: The First Fermi Large Area Telescope Catalog of Gamma-ray Pulsars. ApJS **187**, 460–494 (2010)
2. Abdo, A.A., et al. (Fermi collaboration): A Population of Gamma-Ray Millisecond Pulsars Seen with the Fermi Large Area Telescope. Science **325**, 848– (2009)
3. Abdo, A.A., et al. (Fermi collaboration): Fermi/LAT observations of LS 5039. ApJ **706**, L56–L61 (2009)
4. Abdo, A.A., et al. (Fermi collaboration): Modulated High-Energy Gamma-Ray Emission from the Microquasar Cygnus X-3. Science **326**, 1512– (2009)
5. Abdo, A.A., et al. (Fermi collaboration): Fermi Large Area Telescope Observations of the Crab Pulsar And Nebula. ApJ **708**, 1254–1267 (2010)
6. Abdo, A.A., et al. (Fermi collaboration): Fermi Large Area Telescope Observations of the Supernova Remnant W28 (G6.4-0.1). ApJ **718**, 348–356 (2010)
7. Abdo, A.A., et al. (Fermi collaboration): Fermi-Lat Discovery of GeV Gamma-Ray Emission from the Young Supernova Remnant Cassiopeia A. ApJ **710**, L92–L97 (2010)
8. Abdo, A.A., et al. (Fermi collaboration): Gamma-Ray Emission from the Shell of Supernova Remnant W44 Revealed by the Fermi LAT. Science **327**, 1103– (2010)
9. Abdo, A.A., et al. (Fermi collaboration): Observation of Supernova Remnant IC 443 with the Fermi Large Area Telescope. ApJ **712**, 459–468 (2010)
10. Abdo, A.A., et al. (Fermi collaboration): Discovery of High-energy Gamma-ray Emission from the Binary System PSR B1259-63/LS 2883 around Periastron with Fermi. ApJ **736**, L11 (2011)
11. Abdo, A.A., et al. (Fermi collaboration): Gamma-Ray Flares from the Crab Nebula. Science **331**, 739– (2011)
12. Abdo, A.A., et al. (Fermi collaboration): Observations of the Young Supernova Remnant RX J1713.7-3946 with the Fermi Large Area Telescope. ApJ **734**, 28 (2011)
13. Abdo, A.A., et al. (Milagro Collaboration): TeV Gamma-Ray Sources from a Survey of the Galactic Plane with Milagro. ApJ **664**, L91–L94 (2007)
14. Abramowski, A., et al. (HESS collaboration): A new SNR with TeV shell-type morphology: HESS J1731–347. A&A **531**, A81 (2011)
15. Acciari, V.A., et al. (VERITAS collaboration): Observation of Extended Very High Energy Emission from the Supernova Remnant IC 443 with VERITAS. ApJ **698**, L133–L137 (2009)
16. Acciari, V.A., et al. (VERITAS collaboration): Discovery of TeV Gamma-ray Emission from Tycho's Supernova Remnant. ApJ **730**, L20 (2011)
17. Acciari, V.A., et al. (VERITAS collaboration): VERITAS Observations of the TeV Binary LSI + 61303 During 2008–2010. ApJ **738**, 3 (2011)
18. Acciari, V.A., et al. (VERTAS collaboration): Multiwavelength Observations of LSI + 61303 with Veritas, Swift, and RXTE. ApJ **700**, 1034–1041 (2009)
19. Acciari, V.A., et al. (VERTAS collaboration): Observations of the Shell-type Supernova Remnant Cassiopeia A at TeV Energies with VERITAS. ApJ **714**, 163–169 (2010)

20. Acero, F., et al. (HESS collaboration): First detection of VHE γ-rays from SN 1006 by HESS. A&A **516**, A62 (2010)
21. Acero, F., et al. (HESS Collaboration): Localizing the VHE γ-ray source at the Galactic Centre. MNRAS **402**, 1877–1882 (2010)
22. Achterberg, A., Gallant, Y.A., Kirk, J.G., Guthmann, A.W.: Particle acceleration by ultrarelativistic shocks: theory and simulations. MNRAS **328**, 393–408 (2001)
23. Actis, M., et al. (CTA consortium): Design concepts for the Cherenkov Telescope Array CTA: an advanced facility for ground-based high-energy gamma-ray astronomy. Experimental Astronomy **32**, 193–316 (2011)
24. Adriani, O., Barbarino, G.C., Bazilevskaya, G.A., et al. (Pamela collaboration): An anomalous positron abundance in cosmic rays with energies 1.5–100 GeV. Nature **458**, 607–609 (2009)
25. Adriani, O., Barbarino, G.C., Bazilevskaya, G.A., et al. (Pamela collaboration): PAMELA Measurements of Cosmic-Ray Proton and Helium Spectra. Science **332**, 69–71 (2011)
26. Aharonian, F., Anchordoqui, L., Khangulyan, D., Montaruli, T.: Microquasar LS 5039: a TeV gamma-ray emitter and a potential TeV neutrino source. Journal of Physics Conference Series **39**, 408–415 (2006)
27. Aharonian, F., Buckley, J., Kifune, T., Sinnis, G.: High energy astrophysics with ground-based gamma ray detectors. Reports on Progress in Physics **71**(9), 096,901 (2008)
28. Aharonian, F., et al. (HEGRA collaboration): Evidence for TeV gamma ray emission from Cassiopeia A. A&A **370**, 112–120 (2001)
29. Aharonian, F., et al. (HEGRA collaboration): An unidentified TeV source in the vicinity of Cygnus OB2. A&A **393**, L37–L40 (2002)
30. Aharonian, F., et al. (HEGRA Collaboration): Variations of the TeV energy spectrum at different flux levels of Mkn 421 observed with the HEGRA system of Cherenkov telescopes. A&A **393**, 89–99 (2002)
31. Aharonian, F., et al. (HEGRA Collaboration): The Crab Nebula and Pulsar between 500 GeV and 80 TeV: Observations with the HEGRA Stereoscopic Air Cerenkov Telescopes. ApJ **614**, 897–913 (2004)
32. Aharonian, F., et al. (HESS collaboration): Very high energy gamma rays from the direction of Sagittarius A*. A&A **425**, L13–L17 (2004)
33. Aharonian, F., et al. (HESS collaboration): Detection of TeV γ-ray emission from the shell-type supernova remnant RX J0852.0-4622 with HESS. A&A **437**, L7–L10 (2005)
34. Aharonian, F., et al. (HESS collaboration): Discovery of extended VHE gamma-ray emission from the asymmetric pulsar wind nebula in MSH 15–52 with HESS. A&A **435**, L17–L20 (2005)
35. Aharonian, F., et al. (HESS collaboration): Discovery of the binary pulsar PSR B1259-63 in very-high-energy gamma rays around periastron with HESS. A&A **442**, 1–10 (2005)
36. Aharonian, F., et al. (HESS collaboration): Discovery of Very High Energy Gamma Rays Associated with an X-ray Binary. Science **309**, 746–749 (2005)
37. Aharonian, F., et al. (HESS collaboration): Very high energy gamma rays from the composite SNR G0.9 + 0.1. A&A **432**, L25–L29 (2005)
38. Aharonian, F., et al. (HESS collaboration): 3.9 day orbital modulation in the TeV γ-ray flux and spectrum from the X-ray binary LS 5039. A&A **460**, 743–749 (2006)
39. Aharonian, F., et al. (HESS Collaboration): Discovery of very-high-energy γ-rays from the Galactic Centre ridge. Nature **439**, 695–698 (2006)
40. Aharonian, F., et al. (HESS collaboration): Energy dependent γ-ray morphology in the pulsar wind nebula HESS J1825-137. A&A **460**, 365–374 (2006)
41. Aharonian, F., et al. (HESS collaboration): First detection of a VHE gamma-ray spectral maximum from a cosmic source: HESS discovery of the Vela X nebula. A&A **448**, L43–L47 (2006)
42. Aharonian, F., et al. (HESS collaboration): Observations of the Crab nebula with HESS. A&A **457**, 899–915 (2006)
43. Aharonian, F., et al. (HESS collaboration): Primary particle acceleration above 100 TeV in the shell-type supernova remnant RX J1713.7-394 with deep HESS observations. A&A **464**, 235–243 (2007)

44. Aharonian, F., et al. (HESS collaboration): Discovery of very high energy gamma-ray emission coincident with molecular clouds in the W 28 (G6.4-0.1) field. A&A **481**, 401–410 (2008)

45. Aharonian, F., et al. (HESS collaboration): Discovery of Gamma-Ray Emission From the Shell-Type Supernova Remnant RCW 86 With Hess. ApJ **692**, 1500–1505 (2009)

46. Aharonian, F., et al. (HESS collaboration): Spectrum and variability of the Galactic center VHE γ-ray source HESS J1745-290. A&A **503**, 817–825 (2009)

47. Aharonian, F., et al. (HESS collaboration): Very high energy γ-ray observations of the binary PSR B1259-63/SS2883 around the 2007 Periastron. A&A **507**, 389–396 (2009)

48. Aharonian, F., Neronov, A.: High-Energy Gamma Rays from the Massive Black Hole in the Galactic Center. ApJ **619**, 306–313 (2005)

49. Aharonian, F., Neronov, A.: TeV Gamma Rays From the Galactic Center Direct and Indirect Links to the Massive Black Hole in Sgr A. Ap&SS **300**, 255–265 (2005)

50. Aharonian, F., Taylor, A.M.: Limitations on the photo-disintegration process as a source of VHE photons. Astroparticle Physics **34**, 258–266 (2010)

51. Aharonian, F.A.: TeV gamma rays from BL Lac objects due to synchrotron radiation of extremely high energy protons. New Astr. **5**, 377–395 (2000)

52. Aharonian, F.A.: Very high energy cosmic gamma radiation: a crucial window on the extreme Universe. River Edge, NJ: World Scientific Publishing (2004)

53. Aharonian, F.A., Akerlof, C.W.: Gamma-Ray Astronomy with Imaging Atmospheric Cerenkov Telescopes. Annual Review of Nuclear and Particle Science **47**, 273–314 (1997)

54. Aharonian, F.A., Atoyan, A.M.: Compton scattering of relativistic electrons in compact X-ray sources. Ap&SS **79**, 321–336 (1981)

55. Aharonian, F.A., Atoyan, A.M.: Cosmic gamma-rays associated with annihilation of relativistic e^+–e^- pairs. Physics Letters B **99**, 301–304 (1981)

56. Aharonian, F.A., Atoyan, A.M.: On the emissivity of $\pi\hat{0}$-decay gamma radiation in the vicinity of accelerators of galactic cosmic rays. A&A **309**, 917–928 (1996)

57. Aharonian, F.A., Atoyan, A.M.: On the origin of TeV radiation of SN 1006. A&A **351**, 330–340 (1999)

58. Aharonian, F.A., Atoyan, A.M.: Broad-band diffuse gamma ray emission of the galactic disk. A&A **362**, 937–952 (2000)

59. Aharonian, F.A., Atoyan, A.M., Kifune, T.: Inverse Compton gamma radiation of faint synchrotron X-ray nebulae around pulsars. MNRAS **291**, 162–176 (1997)

60. Aharonian, F.A., Atoyan, A.M., Nagapetyan, A.M.: Photoproduction of Electron–Positron Pairs in Compact X-Ray Sources. Astrophysics **19**, 187 (1983)

61. Aharonian, F.A., Belyanin, A.A., Derishev, E.V., Kocharovsky, V.V., Kocharovsky, V.V.: Constraints on the extremely high-energy cosmic ray accelerators from classical electrodynamics. Phys. Rev. D **66**(2), 023005 (2002)

62. Aharonian, F.A., Bogovalov, S.V., Khangulyan, D.: Abrupt acceleration of a 'cold' ultrarelativistic wind from the Crab pulsar. Nature **482**, 507–509 (2012)

63. Aharonian, F.A., Drury, L.O., Voelk, H.J.: GeV/TeV gamma-ray emission from dense molecular clouds overtaken by supernova shells. A&A **285**, 645–647 (1994)

64. Aharonian, F.A., et al. (HESS collaboration): High-energy particle acceleration in the shell of a supernova remnant. Nature **432**, 75–77 (2004)

65. Aharonian, F.A., Kelner, S.R., Prosekin, A.Y.: Angular, spectral, and time distributions of highest energy protons and associated secondary gamma rays and neutrinos propagating through extragalactic magnetic and radiation fields. Phys. Rev. D **82**(4), 043002 (2010)

66. Aharonian, F.A., Konopelko, A.K., Völk, H.J., Quintana, H.: 5@5—a 5 GeV energy threshold array of imaging atmospheric Cherenkov telescopes at 5 km altitude. Astroparticle Physics **15**, 335–356 (2001)

67. Aharonian, F.A., Kririllov-Ugriumov, V.G., Vardanian, V.V.: Formation of relativistic electron–photon showers in compact X-ray sources. Ap&SS **115**, 201–225 (1985)

68. Aharonian, F.A., Plyasheshnikov, A.V.: Similarities and differences between relativistic electron–photon cascades developed in matter, photon gas and magnetic field. Astroparticle Physics **19**, 525–548 (2003)

69. Aharonian, F.A., Sunyaev, R.A.: Gamma-ray line emission, nuclear destruction and neutron production in hot astrophysical plasmas—The deuterium boiler as a gamma-ray source. MNRAS **210**, 257–277 (1984)

70. Akhiezer, A.I., Merenkov, N.P., Rekalo, A.P.: On a kinetic theory of electromagnetic showers in strong magnetic fields. Journal of Physics G Nuclear Physics **20**, 1499–1514 (1994)

71. Albert, J., et al. (MAGIC collaboration): Variable Very-High-Energy Gamma-Ray Emission from the Microquasar LS I +61 303. Science **312**, 1771–1773 (2006)

72. Albert, J., et al. (MAGIC collaboration): Discovery of Very High Energy Gamma Radiation from IC 443 with the MAGIC Telescope. ApJ **664**, L87–L90 (2007)

73. Albert, J., et al. (MAGIC collaboration): Observation of VHE γ-rays from Cassiopeia A with the MAGIC telescope. A&A **474**, 937–940 (2007)

74. Albert, J., et al. (MAGIC collaboration): Very High Energy Gamma-Ray Radiation from the Stellar Mass Black Hole Binary Cygnus X-1. ApJ **665**, L51–L54 (2007)

75. Aleksic, J., et al. (MAGIC collaboration): Phase-resolved energy spectra of the Crab pulsar in the range of 50–400 GeV measured with the MAGIC telescopes. A&A **540**, A69 (2012)

76. Aliu, E., et al. (VERITAS collaboration): Detection of Pulsed Gamma Rays Above 100 GeV from the Crab Pulsar. Science **334**, 69– (2011)

77. Anchordoqui, L.A., Beacom, J.F., Goldberg, H., Palomares-Ruiz, S., Weiler, T.J.: TeV γ rays and neutrinos from photodisintegration of nuclei in Cygnus OB2. Phys. Rev. D **75**(6), 063001 (2007)

78. Anguelov, V., Vankov, H.: Electromagnetic showers in a strong magnetic field. Journal of Physics G Nuclear Physics **25**, 1755–1764 (1999)

79. Atoyan, A., Buckley, J., Krawczynski, H.: A Gamma-Ray Burst Remnant in Our Galaxy: HESS J1303-631. ApJ **642**, L153–L156 (2006)

80. Atoyan, A., Dermer, C.: Gamma Rays from the Tycho Supernova Remnant: Leptonic or Hadronic? ArXiv e-prints (2011)

81. Atoyan, A., Dermer, C.D.: TeV Emission from the Galactic Center Black Hole Plerion. ApJ **617**, L123–L126 (2004)

82. Atoyan, A.M.: Radio spectrum of the Crab nebula as an evidence for fast initial spin of its pulsar. A&A **346**, L49–L52 (1999)

83. Atoyan, A.M., Aharonian, F.A.: On the mechanisms of gamma radiation in the Crab Nebula. MNRAS **278**, 525–541 (1996)

84. Atoyan, A.M., Aharonian, F.A.: Modelling of the non-thermal flares in the Galactic microquasar GRS 1915 + 105. MNRAS **302**, 253–276 (1999)

85. Atoyan, A.M., Aharonian, F.A., Tuffs, R.J., Völk, H.J.: On the gamma-ray fluxes expected from Cassiopeia A. A&A **355**, 211–220 (2000)

86. Atwood, W.B., Abdo, A.A., Ackermann, M., Althouse, W., Anderson, B., Axelsson, M., Baldini, L., Ballet, J., Band, D.L., Barbiellini, G., et al.: The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission. ApJ **697**, 1071–1102 (2009)

87. Balbo, M., Walter, R., Ferrigno, C., Bordas, P.: Twelve-hour spikes from the Crab Pevatron. A&A **527**, L4 (2011)

88. Ball, L., Kirk, J.G.: Probing pulsar winds using inverse compton scattering. Astroparticle Physics **12**, 335–349 (2000)

89. Baring, M.G.: Synchrotron pair cascades in strong magnetic fields. A&A **225**, 260–276 (1989)

90. Barkov, M.V., Khangulyan, D.V.: Direct wind accretion and jet launch in binary systems. MNRAS **421**, 1351–1359 (2012). DOI LSI + 61303 or LS 5039

91. Bednarek, W.: Gamma-rays from synchrotron pair cascades in blazars? MNRAS **285**, 69–81 (1997)

92. Bednarek, W.: High energy processes in pulsar wind nebulae. Ap&SS **309**, 179–187 (2007)

93. Bednarek, W., Idec, W.: On the variability of the GeV and multi-TeV gamma-ray emission from the Crab nebula. MNRAS **414**, 2229–2234 (2011)

94. Berezhko, E.G., Ksenofontov, L.T., Völk, H.J.: Confirmation of strong magnetic field amplification and nuclear cosmic ray acceleration in SN 1006. A&A **412**, L11–L14 (2003)

95. Berezhko, E.G., Völk, H.J.: Hadronic versus leptonic origin of the gamma-ray emission from supernova remnant RX J1713.7-3946. A&A **492**, 695–701 (2008)

96. Berezinskii, V.S., Bulanov, S.V., Dogiel, V.A., Ptuskin, V.S.: Astrophysics of cosmic rays. Amsterdam: North-Holland (1990)

97. Beskin, V.S., Rafikov, R.R.: On the particle acceleration near the light surface of radio pulsars. MNRAS **313**, 433–444 (2000)

98. Blumenthal, G.R., Gould, R.J.: Bremsstrahlung, Synchrotron Radiation, and Compton Scattering of High-Energy Electrons Traversing Dilute Gases. Reviews of Modern Physics **42**, 237–271 (1970)

99. Boettcher, M., Schlickeiser, R.: The pair production spectrum from photon–photon annihilation. A&A **325**, 866–870 (1997)

100. Bogovalov, S.V., Aharonian, F.A.: Very-high-energy gamma radiation associated with the unshocked wind of the Crab pulsar. MNRAS **313**, 504–514 (2000)

101. Bogovalov, S.V., Chechetkin, V.M., Koldoba, A.V., Ustyugova, G.V.: Interaction of pulsar winds with interstellar medium: numerical simulation. MNRAS **358**, 705–715 (2005)

102. Bogovalov, S.V., Khangoulian, D.V.: On the origin of the torus and jet-like structures in the centre of the Crab Nebula. MNRAS **336**, L53–L55 (2002)

103. Bogovalov, S.V., Khangulyan, D., Koldoba, A.V., Ustyugova, G.V., Aharonian, F.A.: Modelling the interaction between relativistic and non-relativistic winds in the binary system PSR B1259-63/SS2883—II. Impact of the magnetization and anisotropy of the pulsar wind. MNRAS **419**, 3426–3432 (2012)

104. Bogovalov, S.V., Khangulyan, D.V., Koldoba, A.V., Ustyugova, G.V., Aharonian, F.A.: Modelling interaction of relativistic and non-relativistic winds in binary system PSR B1259-63/SS2883—I. Hydrodynamical limit. MNRAS **387**, 63–72 (2008)

105. Bongiorno, S.D., Falcone, A.D., Stroh, M., Holder, J., Skilton, J.L., Hinton, J.A., Gehrels, N., Grube, J.: A New TeV Binary: The Discovery of an Orbital Period in HESS J0632+057. ApJ **737**, L11 (2011)

106. Bonometto, S., Rees, M.J.: On possible observable effects of electron pair-production in QSOs. MNRAS **152**, 21 (1971)

107. Bosch-Ramon, V.: Theoretical overview on high-energy emission in microquasars. Ap&SS **309**, 321–331 (2007)

108. Böttcher, M., Dermer, C.D.: Photon–Photon Absorption of Very High Energy Gamma Rays from Microquasars: Application to LS 5039. ApJ **634**, L81–L84 (2005)

109. Bucciantini, N., Arons, J., Amato, E.: Modelling spectral evolution of pulsar wind nebulae inside supernova remnants. MNRAS **410**, 381–398 (2011)

110. Buehler, R., Scargle, J.D., Blandford, R.D., Baldini, L., Baring, M.G., Belfiore, A., Charles, E., Chiang, J., D'Ammando, F., Dermer, C.D., Funk, S., Grove, J.E., Harding, A.K., Hays, E., Kerr, M., Massaro, F., Mazziotta, M.N., Romani, R.W., Saz Parkinson, P.M., Tennant, A.F., Weisskopf, M.C.: Gamma-Ray Activity in the Crab Nebula: The Exceptional Flare of 2011 April. ApJ **749**, 26 (2012)

111. Bulik, T., Rudak, B., Dyks, J.: Spectral features in gamma-rays expected from millisecond pulsars. MNRAS **317**, 97–104 (2000)

112. Bykov, A.M., Pavlov, G.G., Artemyev, A.V., Uvarov, Y.A.: Twinkling pulsar wind nebulae in the synchrotron cut-off regime and the γ-ray flares in the Crab Nebula. MNRAS **421**, L67–L71 (2012)

113. Casanova, S., Aharonian, F.A., Fukui, Y., Gabici, S., Jones, D.I., Kawamura, A., Onishi, T., Rowell, G., Sano, H., Torii, K., Yamamoto, H.: Molecular Clouds as Cosmic-Ray Barometers. PASJ **62**, 769– (2010)

114. Casanova, S., Jones, D.I., Aharonian, F.A., Fukui, Y., Gabici, S., Kawamura, A., Onishi, T., Rowell, G., Sano, H., Torii, K., Yamamoto, H.: Modeling the Gamma-Ray Emission Produced by Runaway Cosmic Rays in the Environment of RX J1713.7-3946. PASJ **62**, 1127–1134 (2010)

115. Cheng, K.S., Chernyshov, D.O., Dogiel, V.A., Ko, C.M., Ip, W.H.: Origin of the Fermi Bubble. ApJ **731**, L17 (2011)

116. Cheng, K.S., Chernyshov, D.O., Dogiel, V.A., Ko, C.M., Ip, W.H., Wang, Y.: The Fermi Bubble as a Source of Cosmic Rays in the Energy Range $\geq 10^{15}$ eV. ApJ **746**, 116 (2012)
117. Chernyakova, M., Malyshev, D., Aharonian, F.A., Crocker, R.M., Jones, D.I.: The High-energy, Arcminute-scale Galactic Center Gamma-ray Source. ApJ **726**, 60 (2011)
118. Coppi, P.S., Blandford, R.D.: Reaction rates and energy distributions for elementary processes in relativistic pair plasmas. MNRAS **245**, 453–507 (1990)
119. Crocker, R.M., Aharonian, F.: Fermi Bubbles: Giant, Multibillion-Year-Old Reservoirs of Galactic Center Cosmic Rays. Physical Review Letters **106**(10), 101102 (2011)
120. Crusius, A., Schlickeiser, R.: Synchrotron radiation in random magnetic fields. A&A **164**, L16–L18 (1986)
121. de Jager, O.C., Harding, A.K.: The expected high-energy to ultra-high-energy gamma-ray spectrum of the Crab Nebula. ApJ **396**, 161–172 (1992)
122. Del Zanna, L., Amato, E., Bucciantini, N.: Axially symmetric relativistic MHD simulations of Pulsar Wind Nebulae in Supernova Remnants. On the origin of torus and jet-like features. A&A **421**, 1063–1073 (2004)
123. Derishev, E.V., Aharonian, F.A., Kocharovsky, V.V., Kocharovsky, V.V.: Particle acceleration through multiple conversions from a charged into a neutral state and back. Phys. Rev. D **68**(4), 043003 (2003)
124. Drury, L.O., Aharonian, F.A., Voelk, H.J.: The gamma-ray visibility of supernova remnants. A test of cosmic ray origin. A&A **287**, 959–971 (1994)
125. Dubus, G.: Gamma-ray absorption in massive X-ray binaries. A&A **451**, 9–18 (2006)
126. Dubus, G.: Gamma-ray binaries: pulsars in disguise? A&A **456**, 801–817 (2006)
127. Dubus, G., Cerutti, B., Henri, G.: The modulation of the gamma-ray emission from the binary LS 5039. A&A **477**, 691–700 (2008)
128. Dubus, G., Cerutti, B., Henri, G.: Relativistic Doppler-boosted emission in gamma-ray binaries. A&A **516**, A18 (2010)
129. Eichler, D., Wiita, P.J.: Neutron beams in active galactic nuclei. Nature **274**, 38 (1978)
130. Ellison, D.C., Bykov, A.M.: Gamma-ray Emission of Accelerated Particles Escaping a Supernova Remnant in a Molecular Cloud. ApJ **731**, 87 (2011)
131. Ellison, D.C., Patnaude, D.J., Slane, P., Raymond, J.: Efficient Cosmic Ray Acceleration, Hydrodynamics, and Self-Consistent Thermal X-Ray Emission Applied to Supernova Remnant RX J1713.7-3946. ApJ **712**, 287–293 (2010)
132. Enomoto, R., et al. (CANGAROO collaboration): The acceleration of cosmic-ray protons in the supernova remnant RX J1713.7-3946. Nature **416**, 823–826 (2002)
133. Finkbeiner, D.P.: Microwave Interstellar Medium Emission Observed by the Wilkinson Microwave Anisotropy Probe. ApJ **614**, 186–193 (2004)
134. Freire, P.C.C., et al. (Fermi collaboration): Fermi Detection of a Luminous γ-Ray Pulsar in a Globular Cluster. Science **334**, 1107– (2011)
135. Fritz, K.D.: Synchrotron emission spectra from shockwaves in active galactic nuclei—an energy and space dependent diffusion coefficient. A&A **214**, 14–28 (1989)
136. Fukui, Y., Sano, H., Sato, J., et al.: A Detailed Study of the Molecular and Atomic Gas Toward the gamma-ray SNR RX J1713.7-3946: Spatial TeV gamma-ray and ISM Gas Correspondence. ArXiv e-prints (2011)
137. Gabici, S., Aharonian, F.A., Casanova, S.: Broad-band non-thermal emission from molecular clouds illuminated by cosmic rays from nearby supernova remnants. MNRAS **396**, 1629–1639 (2009)
138. Gaensler, B.M., Slane, P.O.: The Evolution and Structure of Pulsar Wind Nebulae. ARA&A **44**, 17–47 (2006)
139. Gaisser, T.K.: Cosmic rays and particle, physics (1990)
140. Ginzburg, V.L., Syrovatskii, S.I.: The Origin of Cosmic Rays. New York: Macmillan (1964)
141. Giordano, F., Naumann-Godo, M., Ballet, J., Bechtol, K., Funk, S., Lande, J., Mazziotta, M.N., Rainò, S., Tanaka, T., Tibolla, O., Uchiyama, Y.: Fermi-LAT Detection of the Young SuperNova Remnant Tycho. ArXiv e-prints (2011)

142. Giuliani, A., et al. (AGILE collaboration): AGILE detection of GeV γ-ray emission from the SNR W28. A&A **516**, L11–L14 (2010)

143. Giuliani, A., et al. (AGILE collaboration): Neutral Pion Emission from Accelerated Protons in the Supernova Remnant W44. ApJ **742**, L30–L34 (2011)

144. Gold, T.: Rotating Neutron Stars as the Origin of the Pulsating Radio Sources. Nature **218**, 731–732 (1968)

145. Gould, R.J., Schréder, G.P.: Opacity of the Universe to High-Energy Photons. Physical Review **155**, 1408–1411 (1967)

146. Grondin, M.H., Funk, S., Lemoine-Goumard, M., Van Etten, A., Hinton, J.A., Camilo, F., Cognard, I., Espinoza, C.M., Freire, P.C.C., Grove, J.E., Guillemot, L., Johnston, S., Kramer, M., Lande, J., Michelson, P., Possenti, A., Romani, R.W., Skilton, J.L., Theureau, G., Weltevrede, P.: Detection of the Pulsar Wind Nebula HESS J1825-137 with the Fermi Large Area Telescope. ApJ **738**, 42 (2011)

147. Guilbert, P.W., Fabian, A.C., Rees, M.J.: Spectral and variability constraints on compact sources. MNRAS **205**, 593–603 (1983)

148. Harding, A.K., Muslimov, A.G.: Pulsar Pair Cascades in a Distorted Magnetic Dipole Field. ApJ **726**, L10 (2011)

149. Heitler, W.: Quantum theory of radiation (1954)

150. Hewish, A., Bell, S.J., Pilkington, J.D.H., Scott, P.F., Collins, R.A.: Observation of a Rapidly Pulsating Radio Source. Nature **217**, 709–713 (1968)

151. Higdon, J.C., Lingenfelter, R.E.: OB Associations, Supernova-generated Superbubbles, and the Source of Cosmic Rays. ApJ **628**, 738–749 (2005)

152. Hillas, A.M.: TOPICAL REVIEW: Can diffusive shock acceleration in supernova remnants account for high-energy galactic cosmic rays? Journal of Physics G Nuclear Physics **31**, 95–131 (2005)

153. Hinton, J.A., Aharonian, F.A.: Inverse Compton Scenarios for the TeV Gamma-Ray Emission of the Galactic Center. ApJ **657**, 302–307 (2007)

154. Hinton, J.A., Hofmann, W.: Teraelectronvolt Astronomy. ARA&A **47**, 523–565 (2009)

155. Hinton, J.A., Skilton, J.L., Funk, S., Brucker, J., Aharonian, F.A., Dubus, G., Fiasson, A., Gallant, Y., Hofmann, W., Marcowith, A., Reimer, O.: HESS J0632+057: A New Gamma-Ray Binary? ApJ **690**, L101–L104 (2009)

156. Hofmann, W.: Performance Limits for Cherenkov Instruments. ArXiv Astrophysics e-prints **astro-ph/0603076** (2006)

157. Holder, J.: TeV Gamma-ray Astronomy: A Summary. ArXiv e-prints (2012)

158. Horns, D., Aharonian, F., Santangelo, A., Hoffmann, A.I.D., Masterson, C.: Nucleonic gamma-ray production in Vela X. A&A **451**, L51–L54 (2006)

159. Inoue, T., Yamazaki, R., Inutsuka, S.i., Fukui, Y.: Toward Understanding the Cosmic-Ray Acceleration at Young Supernova Remnants Interacting with Interstellar Clouds: Possible Applications to RX J1713.7-3946. ApJ **744**, 71 (2012)

160. Iyudin, A.F., Diehl, R., Bloemen, H., Hermsen, W., Lichti, G.G., Morris, D., Ryan, J., Schoenfelder, V., Steinle, H., Varendorff, M., de Vries, C., Winkler, C.: COMPTEL observations of Ti-44 gamma-ray line emission from CAS A. A&A **284**, L1–L4 (1994)

161. Jones, F.C.: Calculated Spectrum of Inverse-Compton-Scattered Photons. Physical Review **167**, 1159–1169 (1968)

162. Kamae, T., Karlsson, N., Mizuno, T., Abe, T., Koi, T.: Parameterization of γ, $e^{+/-}$, and Neutrino Spectra Produced by p-p Interaction in Astronomical Environments. ApJ **647**, 692–708 (2006)

163. Katz, B., Waxman, E.: In which shell-type SNRs should we look for gamma-rays and neutrinos from P-P collisions? JCAP **1**, 18 (2008)

164. Kelner, S.R., Aharonian, F.A.: Energy spectra of gamma rays, electrons, and neutrinos produced at interactions of relativistic protons with low energy radiation. Phys. Rev. D **78**(3), 034013 (2008)

165. Kelner, S.R., Aharonian, F.A., Bugayov, V.V.: Energy spectra of gamma rays, electrons, and neutrinos produced at proton–proton interactions in the very high energy regime. Phys. Rev. D **74**(3), 034,018 (2006)

166. Kennel, C.F., Coroniti, F.V.: Magnetohydrodynamic model of Crab nebula radiation. ApJ **283**, 710–730 (1984)
167. Kerschhaggl, M.: TeV flux modulation in PSR B1259-63/LS2883. A&A **525**, A80 (2011)
168. Khangulyan, D., Aharonian, F., Bosch-Ramon, V.: On the formation of TeV radiation in LS 5039. MNRAS **383**, 467–478 (2008)
169. Khangulyan, D., Aharonian, F.A., Bogovalov, S.V., Ribo, M.: Post-Periastron Gamma Ray Flare from PSR B1259-63/LS 2883 as a Result of Comptonization of the Cold Pulsar Wind. ArXiv e-prints (2011)
170. Khangulyan, D., Hnatic, S., Aharonian, F., Bogovalov, S.: TeV light curve of PSR B1259-63/SS2883. MNRAS **380**, 320–330 (2007)
171. Khangulyan, D.V., Aharonian, F.A., Bogovalov, S.V., Koldoba, A.V., Ustyugova, G.V.: Hydrodynamics of Interaction of Pulsar and Stellar Winds and its Impact on the High Energy Radiation of Binary Pulsar Systems. International Journal of Modern Physics D **17**, 1909–1915 (2008)
172. Kirk, J.G.: Particle Acceleration in Relativistic Current Sheets. Physical Review Letters **92**(18), 181101 (2004)
173. Kirk, J.G., Ball, L., Skjaeraasen, O.: Inverse Compton emission of TeV gamma rays from PSR B1259-63. Astroparticle Physics **10**, 31–45 (1999)
174. Kirk, J.G., Lyubarsky, Y., Petri, J.: The Theory of Pulsar Winds and Nebulae. In: W. Becker (ed.) Astrophysics and Space Science Library, *Astrophysics and Space Science Library*, vol. 357, p. 421 (2009)
175. Komissarov, S.S., Lyubarsky, Y.E.: Synchrotron nebulae created by anisotropic magnetized pulsar winds. MNRAS **349**, 779–792 (2004)
176. Komissarov, S.S., Lyutikov, M.: On the origin of variable gamma-ray emission from the Crab nebula. MNRAS **414**, 2017–2028 (2011)
177. Kong, S.W., Cheng, K.S., Huang, Y.F.: Modelling the light curves of PSR B1259-63/LS 2883-II. The effects of anisotropic pulsar wind and Doppler-boosting. ArXiv e-prints (2012)
178. Kozlovsky, B., Murphy, R.J., Ramaty, R.: Nuclear Deexcitation Gamma-Ray Lines from Accelerated Particle Interactions. ApJS **141**, 523–541 (2002)
179. Kusunose, M., Takahara, F.: A Leptonic Model of Steady High-energy Gamma-Ray Emission from Sgr A*. ApJ **748**, 34 (2012)
180. Landau, L., Rumer, G.: The Cascade Theory of Electronic Showers. Royal Society of London Proceedings Series A **166**, 213–228 (1938)
181. Levinson, A.: Particle Acceleration and Curvature TeV Emission by Rotating, Supermassive Black Holes. Physical Review Letters **85**, 912–915 (2000)
182. Levinson, A., Blandford, R.: On the Jets Associated with Galactic Superluminal Sources. ApJ **456**, L29 (1996)
183. Linden, T., Lovegrove, E., Profumo, S.: The Morphology of Hadronic Emission Models for the Gamma-Ray Source at the Galactic Center. ArXiv e-prints (2012)
184. Liu, S., Melia, F., Petrosian, V., Fatuzzo, M.: Stochastic Acceleration in the Galactic Center HESS Source. ApJ **647**, 1099–1105 (2006)
185. Lobanov, A.P., Horns, D., Muxlow, T.W.B.: VLBI imaging of a flare in the Crab nebula: more than just a spot. A&A **533**, A10 (2011)
186. Lunardini, C., Razzaque, S.: High Energy Neutrinos from the Fermi Bubbles. ArXiv e-prints (2011)
187. Lyutikov, M., Balsara, D., Matthews, C.: Crab GeV flares from the corrugated termination shock. MNRAS p. 2769 (2012)
188. Maier, G., for the VERITAS Collaboration, Skilton, J., for the HESS Collaboration: VHE Observations of the Binary Candidate HESS J0632 + 057
189. Malkov, M.A., Diamond, P.H., Sagdeev, R.Z.: On the Gamma-Ray Spectra Radiated by Protons Accelerated in Supernova Remnant Shocks near Molecular Clouds: The case of Supernova Remnant RX J1713.7-3946. ApJ **624**, L37–L40 (2005)
190. Malkov, M.A., O'C Drury, L.: Nonlinear theory of diffusive acceleration of particles by shock waves. Reports on Progress in Physics **64**, 429–481 (2001)

191. Markoff, S., Falcke, H., Fender, R.: A jet model for the broadband spectrum of XTE J1118 + 480

192. Mertsch, P., Sarkar, S.: Fermi Gamma-Ray "Bubbles" from Stochastic Acceleration of Electrons. Physical Review Letters **107**(9), 091101 (2011)

193. Mirabel, I.F., Rodríguez, L.F.: A superluminal source in the Galaxy. Nature **371**, 46–48 (1994)

194. Montmerle, T.: On gamma-ray sources, supernova remnants, OB associations, and the origin of cosmic rays. ApJ **231**, 95–110 (1979)

195. Morlino, G., Amato, E., Blasi, P.: Gamma-ray emission from SNR RX J1713.7-3946 and the origin of galactic cosmic rays. MNRAS **392**, 240–250 (2009)

196. Morlino, G., Caprioli, D.: Strong evidences of hadron acceleration in Tycho's Supernova Remnant. ArXiv e-prints (2011)

197. Mücke, A., Engel, R., Rachen, J.P., Protheroe, R.J., Stanev, T.: Monte Carlo simulations of photohadronic processes in astrophysics. Computer Physics Communications **124**, 290–314 (2000)

198. Mücke, A., Protheroe, R.J., Engel, R., Rachen, J.P., Stanev, T.: BL Lac objects in the synchrotron proton blazar model. Astroparticle Physics **18**, 593–613 (2003)

199. Murase, K., Beacom, J.F.: Very-high-energy gamma-ray signal from nuclear photodisintegration as a probe of extragalactic sources of ultrahigh-energy nuclei. Phys. Rev. D **82**(4), 043008 (2010)

200. Nolan, P.L., Abdo, A.A., Ackermann, M., Ajello, M., Allafort, A., Antolini, E., Atwood, W.B., Axelsson, M., Baldini, L., Ballet, J., et al.: Fermi Large Area Telescope Second Source Catalog. ApJS **199**, 31 (2012)

201. Ochelkov, I.P., Usov, V.V.: Curvature radiation of relativistic particles in the magnetosphere of pulsars. I—Theory. Ap&SS **69**, 439–460 (1980)

202. Ostrowski, M., Bednarz, J.: Comment on the first-order Fermi acceleration at ultra-relativistic shocks. A&A **394**, 1141–1144 (2002)

203. Pacini, F.: Rotating Neutron Stars, Pulsars and Supernova Remnants. Nature **219**, 145–146 (1968)

204. Paredes, J.M., Martí, J., Ribó, M., Massi, M.: Discovery of a High-Energy Gamma-Ray-Emitting Persistent Microquasar. Science **288**, 2340–2342 (2000)

205. Parizot, E., Marcowith, A., van der Swaluw, E., Bykov, A.M., Tatischeff, V.: Superbubbles and energetic particles in the Galaxy. I. Collective effects of particle acceleration. A&A **424**, 747–760 (2004)

206. Piddington, J.H.: The Crab Nebula and the Origin of Interstellar Magnetic Fields. Australian Journal of Physics **10**, 530 (1957)

207. Pittori, C., et al. (AGILE Collaboration): First AGILE catalog of high-confidence gamma-ray sources. A&A **506**, 1563–1574 (2009)

208. Plyasheshnikov, A.V., Aharonian, F.A.: Characteristics of air showers produced by extremely high energy gamma-rays. Journal of Physics G Nuclear Physics **28**, 267–288 (2002)

209. Plyasheshnikov, A.V., Aharonian, F.A., Völk, H.J.: On the potential of the imaging atmospheric Cherenkov technique for study of cosmic gamma rays in the energy range 10–100 TeV. Journal of Physics G Nuclear Physics **26**, 183–201 (2000)

210. Prantzos, N., Boehm, C., Bykov, A.M., Diehl, R., Ferrière, K., Guessoum, N., Jean, P., Knoedlseder, J., Marcowith, A., Moskalenko, I.V., Strong, A., Weidenspointner, G.: The 511 keV emission from positron annihilation in the Galaxy. Reviews of Modern Physics **83**, 1001–1056 (2011)

211. Rea, N., Torres, D.F., Caliandro, G.A., Hadasch, D., van der Klis, M., Jonker, P.G., Méndez, M., Sierpowska-Bartosik, A.: Deep Chandra observations of TeV binaries—II. LS 5039. MNRAS 416, 1514–1521 (2011)

212. Rees, M.J., Gunn, J.E.: The origin of the magnetic field and relativistic particles in the Crab Nebula. MNRAS **167**, 1–12 (1974)

213. Renaud, M., Vink, J., Decourchelle, A., Lebrun, F., den Hartog, P.R., Terrier, R., Couvreur, C., Knödlseder, J., Martin, P., Prantzos, N., Bykov, A.M., Bloemen, H.: The Signature of ^{44}Ti in Cassiopeia A Revealed by IBIS/ISGRI on INTEGRAL. ApJ **647**, L41–L44 (2006)

214. Romani, R.W.: Gamma-Ray Pulsars: Radiation Processes in the Outer Magnetosphere. ApJ **470**, 469 (1996)
215. Romero, G.E., Kaufman Bernadó, M.M., Mirabel, I.F.: Recurrent microblazar activity in Cygnus X-1? A&A **393**, L61–L64 (2002)
216. Sabatini, S., et al. (AGILE collaboration): Episodic Transient Gamma-ray Emission from the Microquasar Cygnus X-1. ApJ **712**, L10–L15 (2010)
217. Sano, H., Sato, J., Horachi, H., et al.: Star-forming Dense Cloud Cores in the TeV Gamma-ray SNR RX J1713.7-3946. ApJ **724**, 59–68 (2010)
218. Snowden, S.L., Egger, R., Freyberg, M.J., McCammon, D., Plucinsky, P.P., Sanders, W.T., Schmitt, J.H.M.M., Truemper, J., Voges, W.: ROSAT Survey Diffuse X-Ray Background Maps. II. ApJ **485**, 125 (1997)
219. Stecker, F.W.: Photodisintegration of Ultrahigh-Energy Cosmic Rays by the Universal Radiation Field. Physical Review **180**, 1264–1266 (1969)
220. Stecker, F.W.: Cosmic gamma rays. NASA Special, Publication **249** (1971)
221. Striani, E., et al. (AGILE collaboration): The Crab Nebula Super-flare in 2011 April: Extremely Fast Particle Acceleration and Gamma-Ray Emission. ApJ **741**, L5 (2011)
222. Su, M., Slatyer, T.R., Finkbeiner, D.P.: Giant Gamma-ray Bubbles from Fermi-LAT: Active Galactic Nucleus Activity or Bipolar Galactic Wind? ApJ **724**, 1044–1082 (2010)
223. Svensson, R.: Steady mildly relativistic thermal plasmas—Processes and properties. MNRAS **209**, 175–208 (1984)
224. Svensson, R.: Non-thermal pair production in compact X-ray sources—First-order Compton cascades is soft radiation fields. MNRAS **227**, 403–451 (1987)
225. Takahashi, T., Kishishita, T., Uchiyama, Y., Tanaka, T., Yamaoka, K., Khangulyan, D., Aharonian, F.A., Bosch-Ramon, V., Hinton, J.A.: Study of the Spectral and Temporal Characteristics of X-Ray Emission of the Gamma-Ray Binary LS 5039 with Suzaku. ApJ **697**, 592–600 (2009)
226. Tam, P.H.T., Huang, R.H.H., Takata, J., Hui, C.Y., Kong, A.K.H., Cheng, K.S.: Discovery of GeV γ-ray Emission from PSR B1259-63/LS 2883. ApJ **736**, L10 (2011)
227. Tanaka, S.J., Takahara, F.: A Model of the Spectral Evolution of Pulsar Wind Nebulae. ApJ **715**, 1248–1257 (2010)
228. Tanaka, T., Allafort, A., Ballet, J., Funk, S., Giordano, F., Hewitt, J., Lemoine-Goumard, M., Tajima, H., Tibolla, O., Uchiyama, Y.: Gamma-Ray Observations of the Supernova Remnant RX J0852.0-4622 with the Fermi Large Area Telescope. ApJ **740**, L51 (2011)
229. Tanaka, T., Uchiyama, Y., Aharonian, F.A., Takahashi, T., et al.: Study of Nonthermal Emission from SNR RX J1713.7-3946 with Suzaku. ApJ **685**, 988–1004 (2008)
230. Tanimori, T., et al. (CANGAROO collaboration): Discovery of TeV Gamma Rays from SN 1006: Further Evidence for the Supernova Remnant Origin of Cosmic Rays. ApJ **497**, L25 (1998)
231. Tavani, M.: The variable Crab Nebula. ArXiv e-prints (2011)
232. Tavani, M., Arons, J.: Theory of High-Energy Emission from the Pulsar/Be Star System PSR 1259-63. I. Radiation Mechanisms and Interaction Geometry. ApJ **477**, 439 (1997)
233. Tavani, M., et al. (AGILE collaboration): Extreme particle acceleration in the microquasar CygnusX-3. Nature **462**, 620–623 (2009)
234. Tavani, M., et al. (AGILE collaboration): Direct Evidence for Hadronic Cosmic-Ray Acceleration in the Supernova Remnant IC 443. ApJ **710**, L151–L155 (2010)
235. Tavani, M., et al. (AGILE collaboration): Discovery of Powerful Gamma-Ray Flares from the Crab Nebula. Science **331**, 736– (2011)
236. Timokhin, A.N.: Time-dependent pair cascades in magnetospheres of neutron stars—I. Dynamics of the polar cap cascade with no particle supply from the neutron star surface. MNRAS **408**, 2092–2114 (2010)
237. Torres, D.F., Marrero, A.Y.R., de Cea Del Pozo, E.: The GeV to TeV connection in the environment of SNR IC 443. MNRAS **408**, 1257–1266 (2010)
238. Uchiyama, H., Matsumoto, H., Tsuru, T.G., Koyama, K., Bamba, A.: Suzaku Observation of HESS J1825-137: Discovery of Largely-Extended X-Rays from PSR J1826-1334. PASJ **61**, 189 (2009)

239. Uchiyama, Y., Funk, S., Katagiri, H., Katsuta, J., Lemoine-Goumard, M., Tajima, H., Tanaka, T., Torres, D.: Fermi-LAT Discovery of GeV Gamma-ray Emission from the Vicinity of SNR W44. ArXiv e-prints (2012)

240. Uchiyama, Y., on behalf of the Fermi LAT collaboration: GeV Gamma Rays from Supernova Remnants Interacting with Molecular Clouds. ArXiv e-prints (2011)

241. Uzdensky, D.A., Cerutti, B., Begelman, M.C.: Reconnection-powered Linear Accelerator and Gamma-Ray Flares in the Crab Nebula. ApJ **737**, L40 (2011)

242. Van Etten, A., Romani, R.W.: Multi-zone Modeling of the Pulsar Wind Nebula HESS J1825-137. ApJ **742**, 62 (2011)

243. Vissani, F., Aharonian, F., Sahakyan, N.: On the detectability of high-energy galactic neutrino sources. Astroparticle Physics **34**, 778–783 (2011)

244. Völk, H.J., Berezhko, E.G., Ksenofontov, L.T.: Internal dynamics and particle acceleration in Tycho's SNR. A&A **483**, 529–535 (2008)

245. Volpi, D., Del Zanna, L., Amato, E., Bucciantini, N.: Non-thermal emission from relativistic MHD simulations of pulsar wind nebulae: from synchrotron to inverse Compton. A&A **485**, 337–349 (2008)

246. Wang, Q.D., Lu, F.J., Gotthelf, E.V.: G359.95-0.04: an energetic pulsar candidate near Sgr A*. MNRAS **367**, 937–944 (2006)

247. Weekes, T.C.: TeV Radiation from Binary Stars. In: J. Sahade, G.E. McCluskey, Y. Kondo (eds.) The Realm of Interacting Binary Stars, *Astrophysics and Space Science Library*, vol. 177, p. 297 (1992)

248. Weekes, T.C.: Very high energy gamma-ray astronomy. The Institute of Physics Publishing (2003)

249. Yuan, Q., Yin, P.F., Wu, X.F., Bi, X.J., Liu, S., Zhang, B.: A Statistical Model for the γ-ray Variability of the Crab Nebula. ApJ **730**, L15 (2011)

250. Zabalza, V., Bosch-Ramon, V., Paredes, J.M.: Thermal X-Ray Emission from the Shocked Stellar Wind of Pulsar Gamma-Ray Binaries. ApJ **743**, 7 (2011)

251. Zdziarski, A.A., Sikora, M., Dubus, G., Yuan, F., Cerutti, B., Ogorzałek, A.: The gamma-ray emitting region of the jet in Cyg X-3. MNRAS **421**, 2956–2968 (2012)

252. Zenitani, S., Hoshino, M.: The Generation of Nonthermal Particles in the Relativistic Magnetic Reconnection of Pair Plasmas. ApJ **562**, L63–L66 (2001)

253. Zirakashvili, V.N., Aharonian, F.: Analytical solutions for energy spectra of electrons accelerated by nonrelativistic shock-waves in shell type supernova remnants. A&A **465**, 695–702 (2007)

254. Zirakashvili, V.N., Aharonian, F.A.: Nonthermal Radiation of Young Supernova Remnants: The Case of RX J1713.7-3946. ApJ **708**, 965–980 (2010)

255. Zirakashvili, V.N., Aharonian, F.A.: Radioactivity and electron acceleration in supernova remnants. Phys. Rev. D **84**(8), 083,010 (2011)

Multi-Messenger Astronomy and Dark Matter

Lars Bergström

1 Preamble

Astrophysics, and more specifically astroparticle physics, has been going through tremendous progress during the last two decades. Still, one of the main problems, that of the nature of the dark matter, remains unsolved. With the help of accelerator experiments (at CERN's Large Hadron Collider (LHC) in particular, which started operation in 2010 and which is currently gathering an impressive integrated luminosity) we could soon hope to get a first indication of the mass scale for the new physics that is associated with dark matter. However, to actually prove that a particle discovered at accelerators has the right properties to constitute the astrophysical dark matter, complementary methods are needed. The fact that a candidate for dark matter is electrically neutral (as not to emit nor absorb light—that is what we mean with the term "dark") can plausibly be determined at accelerators. However, the coupling of the dark matter particles to other matter needs to be weak, and the lifetime of the dark matter particle needs to be at least of the order of the age of the universe. This cannot be tested at accelerators—the dark matter particles would leave the detector in some 100 ns. There could be very useful information still gathered at the LHC, as possibly decays of more massive states in the "dark sector" would be observable, and the missing energy could be estimated.

Fortunately, through observations of various types of messengers—radio waves, microwaves, IR, optical and UV radiation, X-rays, γ-rays and neutrinos, there is great hope that we could get an independent indication of the mass scale of dark matter. This variety of possible methods of indirect detection methods is a part of multi-messenger astronomy, and it is the second way by which we approach the dark matter problem. In particular, for models where the dark matter particles are

L. Bergström
Department of Physics, The Oskar Klein Centre, AlbaNova, Stockholm University,
SE-106 91 Stockholm, Sweden
e-mail: lbe@fysik.su.se

involved in breaking the electroweak symmetry of the Standard Model, so-called WIMP models (for weakly interacting massive particles), prospects of detection in the near future look promising. We will look in some detail on the properties of WIMP candidates, where the fact that they are massive means that they move non-relativistically in galactic halos, and form so-called cold dark matter (CDM). One thought earlier that neutrinos could be the dark matter, but they would constitute hot dark matter (HDM), which is not favoured by observations. Due to free-streaming motion, they would only form very large structures first, which then fragment into smaller scales, like galaxies. This scenario does not agree with observations, as it gives too little power on small scales. Of course, one may also consider an in-between scenario, warm dark matter, usually consisting of having a sterile neutrino (i.e., with no direct Standard Model couplings) in the keV mass region. These may perhaps have some virtue of explaining possible anomalies in dark matter distribution on the very smallest scales, but reliable methods are so far lacking to probe the dark matter distribution, and its couplings to baryons, on these scales.

As a third approach, ingenious experiments for direct detection employing solid state devices, liquid noble gases, etc. can be used to tell us about other important properties of dark matter, like the spin-dependent or spin-independent cross section of dark matter particle scattering on nucleons. Once signals start to be found (and there are some, however not undisputed ones, already), an exciting puzzle will present itself, putting all these pieces of information together. For indirect detection, astrophysical backgrounds that could mask or mimic dark matter signatures will often be a great challenge to overcome. It should therefore be useful to the reader to study also the accompanying articles by Felix Aharonian and Chuck Dermer in this volume—not the least to understand the very interesting aspects of those processes in their own right.

In this set of lectures, I will treat all of the dark matter-related aspects in some detail, and also cover some other current problems of astroparticle physics and cosmology. The sections in these lectures correspond roughly to the lectures at the Saas-Fee Course in Les Diablerets in March, 2010, i.e.,

- The particle universe: introduction, cosmological parameters.
- Basic cross sections for neutrinos and γ-rays; IceCube.
- Density of relic particles from the early universe.
- Dark matter: direct and indirect detection methods; the galactic centre & other promising DM sources.
- Neutrinos and antimatter from dark matter, Sommerfeld enhancement.
- Supersymmetric dark matter, DarkSUSY.
- Particular dark matter candidates (WIMPS, Kaluza-Klein particles, sterile neutrinos ...).
- Diffuse extragalactic γ-rays, Primordial black holes, Hawking radiation.
- Gravitational waves.

The order has been slightly changed (cf. the Table of Contents), and in many cases I have updated the material since the time of the lectures, referring to important developments (actually, quite a number of them) that have appeared after the time

of the Course. This is of course mandatory in a field that evolves so rapidly. For the more basic parts of this review, I have relied heavily on the textbook by Ariel Goobar and myself [1]. Also material from various reviews I have written over the last few years [2–5] has come to use, but also a lot of new material. With these lecture notes, I hope to convey at least some of the excitement I feel for this topic, which relates to some of the outstanding questions still with us in particle physics and cosmology.

2 The Particle Universe: Introduction

2.1 Introduction

One of the most impressive achievements of science is the development of a quite detailed understanding of the physical properties of the universe, even at its earliest stages. Thanks to the fruitful interplay between theoretical analysis, astronomical observations and laboratory experiments we have today very successful 'Standard Models' of both particle physics and cosmology. The Standard Model of particle physics involves matter particles: quarks which always form bound states such as neutrons and protons, and leptons like the electron which is charged and therefore can make up neutral matter when bound to nuclei formed by neutrons and protons. There are also neutral leptons, neutrinos, which do not form bound states but which play a very important role in cosmology and particle astrophysics as we will see throughout these lecture notes. The other important ingredients in the Standard Model of particle physics are the particles which mediate the fundamental forces: the photon, the gluons and the W and Z bosons.

The Standard Model of cosmology is the hot model, which states that the universe is not infinitely old but rather came into existence some 13.7 billion years ago. There may have been a short period with extremely rapid expansion, inflation, which diluted all matter, radiation and other structures (like magnetic monopoles) that might have existed before inflation. When inflation ended, there was a rapid heating (or, thus, rather re-heating) which meant a re-start of expansion, now governed by the relativistic degrees of freedom of our universe, i.e., radiation. The expansion started out in a state which after this small fraction of a second was enormously compressed and very hot (the relation between the density and the temperature can be determined by near-equilibrium thermodynamics at this epoch, when the expansion was "slow" and adiabatic). No bound states could exist because of the intense heat which caused immediate dissociation even of protons and neutrons into quarks if they were formed in the quark-gluon plasma. Subsequently, the universe expanded and cooled, making possible the formation of a sequence of ever more complex objects: protons and neutrons, nuclei, atoms, molecules, clouds, stars, planets,. . . . As we will see, the observational support for the big bang model is overwhelming, but it contains new elements, of dark matter and dark energy, that were not entirely expected. The key observations are:

- The present expansion of the universe.
- The existence of the cosmic microwave background radiation (CMBR), i.e. the relic radiation from the hot stage of the early universe, and measurements of the temperature variations therein.
- The presence of related structure in the late-time distribution of galaxies, so-called "baryon acoustic oscillations" (BAO).
- Supernova cosmology that measures the expansion history, with the surprising result that the cosmic expansion is accelerating.
- The successful calculations of the relative abundance of light elements in the universe, which accurately agrees with what would be synthesized in an initially hot, expanding universe.
- The concept of cosmological inflation, which successfully predicted the geometric flatness of the universe, (thus that the average density is near the critical density, i.e., $\Omega_{tot} = 1$ to an excellent approximation) and gave an explanation of the form of the nearly scale invariant, gaussian temperature fluctuations.
- The discovery of dark matter, pioneered by Zwicky in the 1930s, has stood the test of time and is now an established piece of the cosmological standard model. Dark energy, in its simplest form just a constant vacuum energy, is the other part which explains $\Omega_{tot} = 1$ and the accelerated expansion of the universe.

Several of these observations have been awarded the Nobel Prize, the latest thus being the prize for the discovery of the accelerated expansion of the universe through supernova observations (S. Perlmutter, B. Schmidt and A. Riess, 2011).

As another piece of evidence in favour of the big bang scenario, can be taken the fact that the oldest objects found in the universe—globular clusters of stars and some radioactive isotopes—do not seem to exceed an age around 13 billion years. This gives strong evidence for a universe with a finite age, such as the big bang model predicts.

In some areas, there are new pieces of information to await. For instance, one of the main objectives of the Planck satellite, which will present cosmological data in early 2013, is to search for non-gaussian features, which could tell us more about the mechanism of inflation.

Although there are still many puzzles and interesting details to fill in, both in the Standard Model of particle physics and in the big bang model, they do remarkably well in describing a majority of all phenomena we can observe in nature. Combined, they allow us to follow the history of our universe back to only about 10^{-10} s after the big bang using established physical laws that have been checked in the laboratory. Extending the models, there are inflationary scenarios that describe the evolution back to 10^{-43} s after the big bang!

Behind this remarkable success are the theories of General Relativity and Quantum Field Theory, which we use in these lecture notes. However, many fundamental aspects of the laws of nature remain uncertain and are the subject of present-day research. The key problem is, as it has been for many decades, to find a valid description of quantized gravity, something which is needed to push our limit of knowledge even closer to (and maybe eventually explaining?) the big bang itself.

In this section we will review some of the most striking observational facts about our universe.

2.2 Basic Assumptions

A basic concept in modern cosmology is that of the "Copernican principle", i.e. the supposition that the universe on the average is homogeneous and isotropic. Although this is definitely not true on galactic scales and smaller, the distribution of matter seems to become more and more smooth on large scales, and on the largest scales we can observe, probed by the CMBR, isotropy and homogeneity seems to be fulfilled. The inhomogeneities seem to be 10^{-5} or smaller, apart from a dipole component in the CMBR, which however has a natural interpretation in terms of motion of our galaxy towards other massive galaxies. Given isotropy and homogeneity, the most general line element is the one found by Friedmann, Lemaître, Robertson and Walker (FLRW),

$$ds^2 = dt^2 - a^2(t) \left(\frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right). \tag{1}$$

Measurements on the cosmic microwave background gives (and inflationary theories predicted) $k = 0$, i.e., a geometrically flat universe on large scales, to good accuracy. (There have been suggestions that some of the features of the homogeneous and isotropic model can be alternatively explained if we live in an inhomogeneous universe with large "bubbles" of atypical density. Although a logical possibility, combined constraints from galaxy surveys, supernova data, and the CMBR mean that we would have to live at a fine-tuned location near the centre of such a bubble [6]. We will thus not consider these scenarios.)

The scale factor $a(t)$ follows equations first derived by Friedmann from Einstein's equations in general relativity:

$$H(t)^2 \equiv \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G_N}{3} \rho_{tot}. \tag{2}$$

Here G_N is Newton's gravitational constant, and ρ_{tot} is the total average energy density of the universe. The time-dependent Hubble parameter $H(t)$, has a value today which is known as the Hubble constant,

$$H(t_0) \equiv H_0 = h \cdot 100 \text{ km s}^{-1} \text{Mpc}^{-1}. \tag{3}$$

This defines the dimensionless quantity $h \sim 0.7$, which has to be given by measurement.

The equation which determines the acceleration of the scale factor is also derived from Einstein's equations:

$$\frac{2\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 = -8\pi G_N p, \tag{4}$$

with p being the total pressure.

2.3 Energy and Pressure

In general, there are several components contributing to the energy density, at least matter, radiation and dark energy, where the simplest possibility is a constant vacuum energy—the modern version of Einstein's cosmological constant:

$$\rho_{tot} = \rho_m + \rho_{rad} + \rho_\Lambda. \tag{5}$$

For an isotropic and homogeneous model, the relevant elements of the energy-momentum tensor are

$$T^{ij} = p\delta_{ij} \tag{6}$$

$$T^{i0} = 0 \tag{7}$$

$$T^{00} = \rho_{tot} \tag{8}$$

and there is for each component contributing to p and ρ_{tot} a relation

$$p_i = w_i \cdot \rho_i \tag{9}$$

called the equation of state, which enables one to make predictions for the time evolution of the expansion of the universe and for the relative weights of the different energy components. For non-relativistic matter, the pressure is proportional to $(v/c)^2$, and therefore negligible, $p = 0$, i.e. $w_M = 0$. For radiation on the other hand, $p = \rho/3$, so $w_R = 1/3$. What is the equation of state for vacuum energy? This is easy to motivate from symmetry reasons (as was done already by Lemaître in the 1930s). The energy momentum tensor has to be proportional to the only available rank-2 tensor in empty space-time, namely the Minkowski metric tensor in the cosmic rest frame:

$$T_\Lambda^{\mu\nu} = \rho_\Lambda \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \rho_\Lambda & 0 & 0 & 0 \\ 0 & -\rho_\Lambda & 0 & 0 \\ 0 & 0 & -\rho_\Lambda & 0 \\ 0 & 0 & 0 & -\rho_\Lambda \end{pmatrix}. \tag{10}$$

This form is thus dictated by the requirement of Lorentz invariance. Comparing with the general form of the energy-momentum tensor which has ρ and p in the diagonal,

we thus see that the equation of state is $p = -\rho$, i.e., $w_\Lambda = -1$. The vacuum energy thus acts as a fluid with negative pressure.

2.4 Contributions to Vacuum Energy

How do we describe the contents of the universe, including vacuum energy? Based on its success in particle physics, we try to do it by using quantum field theory, with its particles and fields. A field is a dynamical quantity which is defined in all points of space and at all times. Particles are the lowest excitations of the fields. A particle is characterized by the mass m, spin s, charge Q, and maybe other internal quantum numbers.

The lowest excitations of the field, carrying energy E and three-momentum p can be quantized as harmonic oscillators fulfilling, in the cosmic rest frame (the reference frame where the CMBR looks maximally isotropic), the mass shell condition

$$p_\mu p^\mu = m^2, \tag{11}$$

where the four momentum

$$p^\mu = (E, \mathbf{p}) \tag{12}$$

and

$$p_\mu = (E, -\mathbf{p}). \tag{13}$$

For each possible momentum mode, there will, as for the simple quantum mechanical harmonic oscillator, be a zero-point energy

$$E_i = \omega(p_i) \left(n + \frac{1}{2} \right)_{n=0} = \sqrt{p_i^2 + m^2} \left(n + \frac{1}{2} \right)_{n=0} = \frac{1}{2} \sqrt{p_i^2 + m^2}. \tag{14}$$

However, for a given field, these have to be summed for all modes, meaning that there will be a huge zero-energy density

$$\rho_\Lambda = \frac{1}{2} \frac{1}{(2\pi)^3} \int d^3 p \sqrt{p^2 + m^2}. \tag{15}$$

The highly divergent integral has to be cut-off at some large energy scale, and the first guess is the Planck mass, $m_{Pl} \sim 10^{19}$ GeV, thus

$$\rho_\Lambda = \frac{1}{2} \frac{1}{(2\pi)^3} \int^{m_{Pl}} d^3 p \sqrt{p^2 + m^2} \sim m_{Pl}^4. \tag{16}$$

Unfortunately, this is orders of magnitude too large, and is the same disastrous result one would get by using simple dimensional analysis. Namely, what is the natural

scale of ρ_Λ? We see here that it is governed by the cut-off mass scale when new physics appears, and dimensional analysis gives that in units where $c = 1$ so that length is proportional to an inverse mass, and thus energy per unit volume becomes $[\rho_\Lambda] = [M^4]$. The only mass scale in gravity is m_{Pl}, thus

$$\rho_\Lambda^{th} \sim m_{Pl}^4. \tag{17}$$

Unlike other guesses in physics based on dimensional analysis, this is a terrible prediction. The present-day vacuum energy density of the universe is given by measurements of supernovae and the CMBR and is (using $k = 0$)

$$\rho_\Lambda^{obs} \sim 10^{-122} m_{Pl}^4 \ll m_{Pl}^4 \sim \rho_\Lambda^{th}. \tag{18}$$

To go back to our field theory result, the zero-point energy is really a consequence of the quantum mechanical commutator between the field and its canonical momentum. However, for fermions, anticommutators are used, meaning the sign of the vacuum energy is changed. So, introducing the fermion number $F = 1$ for fermions, $F = 0$ for bosons, one gets

$$\rho_\Lambda = \sum (-1)^F \frac{1}{2} \frac{1}{(2\pi)^3} \int^{m_{Pl}} d^3 p \sqrt{p^2 + m^2}. \tag{19}$$

Remarkably, if there are as many fermionic degrees of freedom as bosonic, and they pairwise have the same mass, the vacuum energy would vanish. Examples of theories having this property are supersymmetric theories, with unbroken supersymmetry. However, since we do not see 0.511 MeV scalar electrons (selectrons), supersymmetry has to be broken. Therefore large effects of the zero-point energy remain, and $\rho_\Lambda \sim m_{SUSY}^4$ with m_{SUSY} (1000 GeV, say) the scale of SUSY breaking. Better, but still enormously much "too high".

In summary, we have encountered one of the most severe problems of cosmology and particle astrophysics: Why is the cosmological constant so small, but still not zero? (By the way, nobody found a good reason that it should be exactly zero, anyway...) Supersymmetry alleviates the problem somewhat, but since supersymmetry is broken there remains large mass terms still giving a value some 50–60 orders of magnitude different from the observed value.

In cosmology the cosmological constant has a dramatic effect. Since it is related to the energy density of the vacuum, and the vacuum is growing in size due to the expansion, it will eventually dominate completely. Matter is on the other hand more and more diluted and becomes less and less important, and radiation is also diluted plus red-shifted:

$$\rho_m \sim (1 + z)^3, \quad \rho_r \sim (1 + z)^4, \quad \rho_\Lambda \sim (1 + z)^0. \tag{20}$$

We see that in the early universe (large redshifts), vacuum energy was irrelevant. Today matter and vacuum energy are almost equal (why now?). In the future, the

expansion rate will grow exponentially, as we will see in the section on inflation, Sect. 3.2.

To explain the smallness of Λ some people resort to (almost) non-scientific reasoning: the anthropic principle, or the landscape of string theory vacua. There the argument goes roughly like this: There exist an amazingly large number of different vacua, i.e., ground states, of the theory, and maybe all of these are realized somewhere in nature. But of course, those with very large values of Λ would accelerate so quickly that structure would not form in the universe and therefore no life could appear. But since we exist, we have to do so in one of the very few universes where life did evolve. Of course, this sounds more like post-dicting the properties of our universe rather than predicting them, which perhaps just shows the desperation in front of the problem of the size of the cosmological constant.

Let us have another look at Planck-mass phenomena. Consider the scattering of a photon on an electron, Compton scattering (we will treat this in detail later, see Fig. 3). The relation between the incident and outgoing wavelength as a function of scattering angle is given by

$$\lambda' - \lambda = \frac{2\pi\hbar}{m_e c} \left(1 - \cos\theta\right) = \frac{2\pi}{m_e} \left(1 - \cos\theta\right) \equiv \lambda_c \left(1 - \cos\theta\right). \tag{21}$$

Here $\lambda_c \equiv 2\pi/m_e$ is called the Compton wavelength (or radius) of the particle (the electron in this case). Note that we use here and onwards units such that $c = \hbar = 1$. This implies that time and length have the same dimension which is inversely proportional to the dimension of mass.

We will see in Sect. 9 the expression for the Schwarzschild radius (the radius which marks the limit of where light can leave the black hole)

$$r_S = \frac{2G_N M}{c^2} = 2G_N M \tag{22}$$

Thus, the Compton radius decreases with mass, but the Schwarzschild radius increases with mass. When are the two equal, i.e., how big must the mass be for the Compton radius to be smaller than the Schwarzschild radius? This is when quantum gravity should be important. (All details of particle properties are smeared out by quantum fluctuations on the order of the Compton wavelength or less, so for $\lambda_c > r_S$ the black hole properties should be unnoticeable.) We see

$$\frac{\lambda_c}{r_S} = \frac{\pi}{G_N M^2} \sim \frac{m_{Pl}^2}{M^2}. \tag{23}$$

Thus, when the mass of an elementary particle is larger than the Planck mass, its Compton radius is smaller than its Schwarzschild radius, which implies that we need quantum gravity! None exists yet, but perhaps string theory is the best bet for such a fundamental theory at the Planck scale? For an electron, $\lambda_c/r_S \sim 10^{45}$, so quantum

gravity effects are completely negligible at the particle level. The same is true for all other Standard Model particles.

2.5 Summary of Observations

To end this section where the main theoretical lines for describing the universe have been laid out, we summarize what we know about the cosmological parameters of the universe from the impressive recent measurements. Analyses combining high-redshift supernova luminosity distances, microwave background fluctuations (from the satellite WMAP) and baryon acoustic oscillations (BAO) in the galaxy distribution give tight constraints [7] on the present mass density of matter in the universe. This is usually expressed in the ratio

$$\Omega_M = \rho_M/\rho_c, \tag{24}$$

normalized to the critical density,

$$\rho_c = 3H_0^2/(8\pi G_N) = h^2 \times 1.9 \cdot 10^{-29} \, \mathrm{g\,cm^{-3}}. \tag{25}$$

The value obtained for the 7-year WMAP data [7] for cold dark matter for the (unknown) particle X making up the dark matter is $\Omega_X h^2 = 0.113 \pm 0.004$, which is around 5 times higher than the value obtained for baryons, $\Omega_B h^2 = 0.0226 \pm 0.0005$. Here $h = 0.704 \pm 0.014$ is the derived [7] present value of the Hubble constant in units of 100 km s^{-1} Mpc^{-1}. In addition, the WMAP data is consistent within a percent with a flat universe ($\Omega_{tot} = 1$) and a value for the dark energy component, e.g. the cosmological constant Λ, of $\Omega_\Lambda = 0.73 \pm 0.02$.

One of the main problems for cosmology and particle physics is to explain the measured density of dark matter, and to give candidates for the identity of the dark matter particles. The fact that dark matter is definitely needed on the largest scales (probed by WMAP), on galaxy cluster scales (as pointed out by Zwicky already in 1933 [8], and verified by gravitational lensing and the temperature distribution of X-ray emitting gas) all the way down to the smallest dwarf galaxies, means that solutions based on changing the laws of gravity seem less natural. In particular, the direct empirical proof of the existence of dark matter given by the "Bullet Cluster" [9] is very difficult to circumvent, as the X-ray signal from the baryonic matter and the gravitational lensing signal from dark matter are clearly separated.

Although the existence of a non-zero cosmological constant (or some similar form of dark energy) in the present-day universe came as a big surprise to most cosmologists and particle physicists, the most successful models of evolution in the universe do make use of a similar effect in models of inflation, as we will see in Sect. 3.2.

3 Relic Density of Particles

There are several important examples of freeze-out in the early universe, for instance
at the synthesis of light elements one second to a few minutes after the big bang,
and the microwave photons from the "surface of last scattering" several hundred
thousand years later. Before we calculate freeze-out, it is convenient to introduce
a formalism which considers freeze-out in general: that is, what happens when a
particle species goes out of equilibrium. A rigorous treatment has to be based on
the Boltzmann transport equation in an expanding background, but here we give a
simplified treatment (see, for example [1] for a more complete discussion).

There are several different contributions to $\Omega = \rho/\rho_c$, like radiation Ω_R, matter
Ω_M and vacuum energy Ω_Λ.

The equations of motion for the matter in the universe are given by the vanishing
of the covariant divergence of the energy-momentum tensor

$$T^{\alpha\beta}_{;\beta} = 0 \tag{26}$$

This gives, for the FLRW metric,

$$\frac{d}{dt}(\rho a^3) = -p\frac{d}{dt}a^3 \tag{27}$$

which shows that the change of energy in a comoving volume element is equal to
minus the pressure times the change in volume. This can be rewritten as

$$a^3\frac{dp}{dt} = \frac{d}{dt}[a^3(\rho + p)] \tag{28}$$

which can be interpreted as a conservation law for the entropy in a volume $a^3(T)$.
For radiation, where $p = \rho/3$, (27) gives $\rho \sim a^{-4}$. Note that all particles fulfilling
$mc^2 \ll k_B T$ have the equation of state of radiation.

The Friedmann equation is

$$H^2(t) = \frac{8\pi G_N \rho}{3} \tag{29}$$

where as a good approximation only the relativistic species contribute appreciably
to ρ. Note that the Hubble parameter $H(t)$ has units of 1/(time). This means in our
units that it has dimensions of mass. The age of the universe at a given time t is simply
of the order of $H^{-1}(t)$, at least when the scale factor increases as a power of t.

We now treat schematically the thermodynamics of the expanding universe. We
assume, which is true if reactions between different species of particles are rapid
enough, that we can use the thermodynamical quantities, temperature T, pressure p,
entropy density s, and other quantities, at each time t to describe the state of the
universe. The constituents have number density n and typical relative velocities v,

and scattering or annihilation cross-section σ, meaning that the interaction rate per particle Γ is given by

$$\Gamma = n\sigma v. \tag{30}$$

The condition that the interactions maintain equilibrium is that the interaction rate is larger than the expansion rate of the universe:

$$\Gamma \gg H \tag{31}$$

Typically, the number density of particles decreases faster with temperature and therefore with time than the Hubble parameter does. This means that at certain epochs some of the particle species will leave thermodynamic equilibrium. Their number density will be "frozen" at some particular value which then only changes through the general dilution due to the expansion. This "freeze-out" of particles is an important mechanism which explains the particle content of the universe we observe today.

Using relativistic statistical mechanics in the cosmic rest frame, the distribution function $f_i(\mathbf{p})$ for particle species of type i is

$$f_i(\mathbf{p}) = \frac{1}{e^{\frac{(E_i - \mu_i)}{T}} \pm 1} \tag{32}$$

with $E_i = \sqrt{\mathbf{p}^2 + m_i^2}$ the energy, μ_i is the chemical potential and T the temperature (we put $k_B = 1$). The minus sign is for particles that obey Bose-Einstein statistics (bosons) and the plus sign is for particles obeying the exclusion principle and therefore Fermi-Dirac statistics (fermions). To a good approximation the chemical potentials can be neglected in the very early universe.

We denote by g_i the number of internal degrees of freedom of particle i. The photon has two polarization states and therefore $g_\gamma = 2$. The neutrinos only have one polarization state, giving $g_\nu = 1$, electrons and muons have $g_{e,\mu} = 2$ (and the same numbers for the antiparticles).

With these definitions, the number density for species i is

$$n_i = \frac{g_i}{(2\pi)^3} \int f_i(\mathbf{p}) d^3 p, \tag{33}$$

and its energy density is

$$\rho_i = \frac{g_i}{(2\pi)^3} \int E_i(\mathbf{p}) f_i(\mathbf{p}) d^3 p. \tag{34}$$

The expression for the pressure is

$$p_i = \frac{g_i}{(2\pi)^3} \int \frac{|\mathbf{p}|^2}{3E_i(\mathbf{p})} f_i(\mathbf{p}) d^3 p. \tag{35}$$

In the nonrelativistic limit $T/m \ll 1$ we can solve the integrals analytically, and the result both for Fermi-Dirac and Bose-Einstein particles is

$$n_{NR} = g_i \left(\frac{mT}{2\pi} \right)^{\frac{3}{2}} e^{-\frac{m}{T}}, \tag{36}$$

$$\rho_{NR} = m \cdot n_{NR}, \tag{37}$$

and

$$p_{NR} = T \cdot n_{NR} \ll \rho_{NR} \tag{38}$$

For nonrelativistic matter, $\langle E \rangle = m + 3T/2$.

In the ultrarelativistic approximation, $T/m \gg 1$, the integrals can also be performed with the results

$$\rho_R = \frac{g_i}{6\pi^2} \int_0^\infty \frac{E^3 dE}{e^{\frac{E}{T}} \pm 1} = \begin{cases} \frac{\pi^2}{30} g_i T^4, & \text{Bose-Einstein} \\ \frac{7}{8} \left(\frac{\pi^2}{30} g_i T^4 \right), & \text{Fermi-Dirac}, \end{cases} \tag{39}$$

$$n_R = \begin{cases} \frac{\zeta(3)}{\pi^2} g_i T^3, & \text{Bose-Einstein} \\ \frac{3}{4} \left(\frac{\zeta(3)}{\pi^2} g_i T^3 \right), & \text{Fermi-Dirac}, \end{cases} \tag{40}$$

with $\zeta(x)$ is the Riemann zeta function, $\zeta(3) = 1.20206\ldots$ The average energy ρ/n for a relativistic particle is

$$\langle E \rangle_{BE} \sim 2.7T \tag{41}$$

and

$$\langle E \rangle_{FD} \sim 3.15T \tag{42}$$

For photons, with the mass $m_\gamma = 0$, and $g_\gamma = 2$, the expression for $\rho_\gamma(T) \sim T^4$ is the famous Stefan Boltzmann law for electromagnetic black-body radiation.

The total contribution to the energy and number density of all kinds of particles in the early universe is to a good approximation (since the energy and number density of a nonrelativistic species is exponentially suppressed),

$$\rho_R(T) = \frac{\pi^2}{30} g_{\text{eff}}(T) T^4 \tag{43}$$

$$p_R(T) = \frac{1}{3} \rho_R(T) = \frac{\pi^2}{90} g_{\text{eff}}(T) T^4 \tag{44}$$

where $g_{\text{eff}}(T)$ counts the total number of internal degrees of freedom (such as spin, colour, etc.) of the particles whose mass fulfill $m \ll T$, and which are in thermodynamic equilibrium with the "primordial cosmic soup" of particles in the early universe. The expression for $g_{\text{eff}}(T)$ has the factor 7/8 for fermions.

As an example, we calculate $g_{\text{eff}}(T)$ for a temperature of, say, 1 TeV when all the particles of the Standard Model were relativistic and in thermal equilibrium. The total number of internal degrees of freedom of the fermions is 90 and for the gauge and Higgs bosons 28, so the total expression for g_{eff} is

$$g_{\text{eff}}(T \sim 1\,\text{TeV}) = 28 + \frac{7}{8} \cdot 90 = 106.75 \tag{45}$$

If we insert the expression for the energy density into the Friedmann equation (29) we get for the radiation-dominated epoch in the early universe

$$H^2 = \frac{8\pi G_N}{3}\rho_R = \frac{8\pi G_N}{3}\frac{\pi^2}{30}g_{\text{eff}}T^4 = 2.76\frac{g_{\text{eff}}T^4}{m_{Pl}^2} \tag{46}$$

or

$$H = 1.66\sqrt{g_{\text{eff}}}\frac{T^2}{m_{Pl}} \tag{47}$$

This is a very important formula governing the physics of the early universe.

For radiation domination, it can be shown that

$$a(t) \sim \sqrt{t} \tag{48}$$

deriving from the equation of state $p = \rho/3$. For matter domination, that is, for $p \sim 0$, one has

$$a(t) \sim t^{\frac{2}{3}}. \tag{49}$$

So for radiation domination,

$$H = \frac{\dot{a}}{a} = \frac{1}{2t} \tag{50}$$

and the time-temperature relation becomes

$$t = 0.30\frac{m_{Pl}}{\sqrt{g_{\text{eff}}}T^2} \sim \left(\frac{1\,\text{MeV}}{T}\right)^2\,\text{s} \tag{51}$$

We now have to determine which particles are in thermal equilibrium at a given temperature, so that we can calculate $g_{\text{eff}}(T)$. The entropy $S(V, T)$ is introduced through

$$dS(V, T) = \frac{1}{T}\left[d(\rho(T)V) + p(T)dV\right] \tag{52}$$

this gives (see [1])

$$S(V, T) = \frac{V}{T}(\rho(T) + p(T))$$ (53)

and from the conservation of the energy-momentum tensor follows

$$\frac{d}{dt}\left(\frac{a^3}{T}[\rho(T) + p(T)]\right) = 0.$$ (54)

Identifying the volume V with $a^3(t)$ and comparing with (53) we find the law of conservation of entropy in the volume $a^3(t)$. Sometimes it is more useful to work with the entropy density $s(T)$ rather than the total entropy $S(V, T)$ within the volume V. The definition is thus:

$$s(T) \equiv \frac{S(V, T)}{V} = \frac{\rho(T) + p(T)}{T}$$ (55)

In the early universe, both the energy density and the pressure were dominated by relativistic particles with the equation of state $p = \rho/3$. Using (55) and the relativistic expressions for the energy density and the pressure (Eqs. (43) and (44)), gives density s

$$s = \frac{2\pi^2}{45}g_{\text{eff}}^s T^3$$ (56)

where g_{eff}^s is defined in a similar way as g_{eff}.

Since s and n_γ both vary as T^3 there is a simple relationship between them. With

$$n_\gamma = \frac{2\zeta(3)}{\pi^2}T^3$$ (57)

from Eq. (40), we find

$$s = \frac{\pi^4}{45\zeta(3)}g_{\text{eff}}^s n_\gamma \sim 1.8 g_{\text{eff}}^s n_\gamma$$ (58)

Following [1] we now consider a case of great interest for the dark matter problem. Suppose that there exists some kind of unknown particle χ, with antiparticle $\bar{\chi}$, that can annihilate each other and be pair created through processes $\chi + \bar{\chi} \leftrightarrow X + \bar{X}$, where X stands for any type of particle to which the χs can annihilate.[1] We further assume that the X particles have zero chemical potential and that they are kept in thermal equilibrium with the photons and the other light particles in the early universe (the X particles can be quarks, leptons etc.)

---

[1] The supersymmetric neutralino is actually its own antiparticle (just as the photon is its own antiparticle). The formalism is very similar in this case. In particular, a neutralino can annihilate with another neutralino giving other, non-supersymmetric particles in the final state.

How will the number density n_χ evolve with time (and therefore with temperature)? It is clear that in exact thermal equilibrium the number of χ particles in a comoving volume $N_\chi = a^3 n_\chi$ will be given by the equilibrium value $n_\chi^{EQ}(T)$ (see (40)). (In exact thermal equilibrium the rate for the process $\chi + \bar\chi \leftrightarrow X + \bar{X}$ is the same in both directions.) If the actual number density $n_\chi(T)$ is larger than the equilibrium density the reaction will go faster to the right: that is, the χ particles will annihilate faster than they are created. The depletion rate of χ should be proportional to $\sigma_{\chi\bar\chi \to X\bar{X}}|\mathbf{v}|n_\chi^2$ (quadratic in the density, since it should be proportional to the product of n_χ and $n_{\bar\chi}$, and these are equal). However, χ particles are also created by the inverse process, with a rate proportional to $(n_\chi^{EQ})^2$. We have thus 'derived' the basic equation that governs the departure from equilibrium for the species χ:

$$\frac{dn_\chi}{dt} + 3Hn_\chi = -\langle \sigma_{\chi\bar\chi \to X\bar{X}}|\mathbf{v}|\rangle [n_\chi^2 - (n_\chi^{EQ})^2]. \tag{59}$$

The left-hand side comes from $\frac{1}{a^3}\frac{d}{dt}[n_\chi a^3]$; the term proportional to $3H$ just expresses the dilution that automatically comes from the Hubble expansion. The expression $\langle \sigma_{\chi\bar\chi \to X\bar{X}}|\mathbf{v}|\rangle$ stands for the thermally averaged cross section times velocity. This averaging is necessary, since the annihilating particles have random thermal velocities and directions. Summing over all possible annihilation channels gives

$$\frac{dn_\chi}{dt} + 3Hn_\chi = -\langle \sigma_A|\mathbf{v}|\rangle [n_\chi^2 - (n_\chi^{EQ})^2], \tag{60}$$

where σ_A is the total annihilation cross section.

Using the time-temperature relation of Eq. (51) (for radiation dominance)

$$t = 0.30 \frac{m_{Pl}}{T^2 \sqrt{g_{\text{eff}}}} \tag{61}$$

this can be converted to an evolution equation for n_χ as a function of temperature. Introducing the dimensionless variable $x \equiv m_\chi/T$, and normalizing n_χ to the entropy density:

$$Y_\chi = \frac{n_\chi}{s} \tag{62}$$

gives after some intermediate steps

$$\frac{dY_\chi}{dx} = -\frac{m_\chi m_{Pl} c_{\text{eff}}}{x^2}\sqrt{\frac{\pi}{45}}\langle \sigma_A|\mathbf{v}|\rangle (Y_\chi^2 - (Y_\chi^{EQ})^2) \tag{63}$$

where

$$c_{\text{eff}} = \frac{g_{\text{eff}}^s}{\sqrt{g_{\text{eff}}}} \tag{64}$$

or, after some rearrangement,

$$\frac{x}{Y_\chi^{EQ}} \frac{dY_\chi}{dx} = -\frac{\Gamma_A}{H}\left[\left(\frac{Y_\chi}{Y_\chi^{EQ}}\right)^2 - 1\right], \tag{65}$$

where $\Gamma_A = n_\chi^{EQ}\langle\sigma_A|\mathbf{v}|\rangle$. This equation can be solved numerically with the boundary condition that for small x, $Y_\chi \sim Y_\chi^{EQ}$ (since at high temperature the χ particles were in thermal equilibrium with the other particles). We see from (65) that the evolution is governed by the factor Γ_A/H, the interaction rate divided by the Hubble expansion rate.

The solutions to these equations have to be obtained numerically in the general case to find the temperature T_f and therefore the value of x_f of freeze-out and the asymptotic value $Y_\chi(\infty)$ of the relic abundance of the species χ. There are, however, some simple limiting cases. If the species χ is relativistic at freeze-out, then Y_χ^{EQ} is not changing with time during the period of freeze-out, and the resulting $Y_\chi(\infty)$ is just the equilibrium value at freeze-out,

$$Y_\chi(\infty) = Y_\chi^{EQ}(x_f) = \frac{45\zeta(3)}{2\pi^4}\frac{g_{eff}}{g_{\text{eff}}^s(x_f)} \tag{66}$$

where $g_{eff} = g$ for bosons and $3g/4$ for fermions. A particle that was relativistic at freeze-out is called a hot relic. A typical example is the neutrino. The present mass density of a hot relic with mass m is

$$\Omega_\chi h^2 = 7.8 \cdot 10^{-2}\frac{g_{eff}}{g_{\text{eff}}^s(x_f)}\left(\frac{m_\chi}{1\text{ eV}}\right) \tag{67}$$

Note that today the motion of a particle with mass greater than the small number $T_0 = 2.73\text{ K} = 2.4 \cdot 10^{-4}\text{ eV}$ is of course non-relativistic and therefore the contribution to the energy density is dominated by its rest mass energy. A Standard Model neutrino has $g_{eff} = 2 \cdot 3/4 = 1.5$ and decoupled at a few MeV when $g_{\text{eff}}^s = g_{eff} = 10.75$. We find

$$\Omega_{\nu\bar{\nu}}h^2 = \frac{\sum_i m_{\nu_i}}{(93\text{ eV})}. \tag{68}$$

As we will see, present estimates of the neutrinos masses, based on the observation of neutrino oscillations, give a sum much less than 1 eV, which means that neutrinos are not the main form of dark matter. On the other hand, we are now rather certain that they do contribute a small fraction of non-baryonic dark matter!

This analysis has been valid for hot relics, or hot dark matter. For cold relics (particles that were non-relativistic at freeze-out) the Eq. (65) has to be solved numerically. There one finds that for massive particles in the mass range between, say, 10 GeV and a few TeV, $x_f \sim 1/20$, and moreover to a good approximation the relic density only depends on the cross section times velocity,

$$\Omega_X h^2 \simeq 0.11 \times \frac{2.8 \cdot 10^{-26} \text{ cm}^3 \text{ s}^{-1}}{\langle \sigma_A |\mathbf{v}| \rangle}. \tag{69}$$

Another striking result is that, if one gives typical gauge couplings to the particle X, and a mass of typical weak interaction magnitude (100–300 GeV, say), then $\langle \sigma_A v \rangle$ is such that the resulting $\Omega_X h^2 \sim 0.11$. This is the essence of what is sometimes called the "WIMP miracle".

As can be understood, the value of x_f when Y_χ leaves the equilibrium curve is lower for a smaller cross section σ_A. This is because of the fact that in thermal equilibrium, massive particles will eventually be exponentially suppressed. That is, more weakly interacting particles decouple earlier, and since the equilibrium curve for a non relativistic species drops fast with increasing x, more weakly coupled particles will have a higher relic abundance.

Going through the numerical analysis one finds that a hypothetical neutrino with mass $m_\nu \sim 3$ GeV would also have about the right mass to close the universe. On the other hand, the range between 90 eV and 3 GeV is cosmologically disallowed for a stable neutrino. There are arguments from large-scale structure formation that favour cold relics over hot relics, so such a neutrino would be a good dark matter candidate. Data from the LEP accelerator at CERN have, however, excluded any ordinary neutrino with a mass in the GeV range.

3.1 Coannihilations

There are instances when the simple treatment discussed here has to be improved. One example is for instance the possibility that entropy may be generated by other particles than those of the Standard Model, before, at, or after decoupling. Another example, which for instance appears in some supersymmetric models, is that of coannihilations. This was first discussed in [10], here we follow the more detailed treatment in [11].

We will here outline the procedure developed in [2, 11] which is used in Dark-SUSY [144]. For more details, see [2, 11]. DarkSUSY is a free FORTRAN package that can be used to compute a variety of dark matter related quantities, such the relic density and the scattering and annihilation rates to many different channels. It was developed for computations in the Minimal Supersymmetric extension to the Standard Model (MSSM), but it is modular and can be adapted to most WIMP models.

We consider annihilation of N particles with mass m_i and internal degrees of freedom g_i. For convenience, we may order them such that $m_1 \leq m_2 \leq \cdots \leq m_{N-1} \leq m_N$. For the lightest particle (which is the dark matter candidate, if a symmetry is guaranteeing the stability, like what is called R-parity for supersymmetry, see Sect. 5), we use both the notation m_1 and m_χ.

All heavier particles will eventually decay to the lightest, stable, and therefore we add the number densities up,

$$n = \sum_{i=1}^{N} n_i.$$

The scattering rate of particles on particles in the thermal background "soup" is generally much faster than the annihilation rate, since the background particle densities of Standard Model particles, n_{SM} is much larger than each of the particle densities in the dark sector n_i. The important SM particles are, as we have seen, those that are relativistic and cold dark matter particles (WIMPs) are nonrelativistic, and thus suppressed by the Boltzmann factor. Thus, the n_i distributions remain in thermal equilibrium during their ("chemical") freeze-out.

We then get

$$\frac{dn}{dt} = -3Hn - \langle \sigma_{\text{eff}} v \rangle \left(n^2 - n_{\text{eq}}^2 \right) \tag{70}$$

where

$$\langle \sigma_{\text{eff}} v \rangle = \sum_{ij} \langle \sigma_{ij} v_{ij} \rangle \frac{n_i^{\text{eq}} \, n_j^{\text{eq}}}{n^{\text{eq}} \, n^{\text{eq}}}. \tag{71}$$

with

$$v_{ij} = \frac{\sqrt{(p_i \cdot p_j)^2 - m_i^2 m_j^2}}{E_i E_j}. \tag{72}$$

Using the Maxwell-Boltzmann approximation for the velocity distributions one can derive the following expression for the thermally averaged annihilation cross section [11]

$$\langle \sigma_{\text{eff}} v \rangle = \frac{\int_0^\infty dp_{\text{eff}} \, p_{\text{eff}}^2 \, W_{\text{eff}} K_1 \left(\frac{\sqrt{s}}{T} \right)}{m_1^4 T \left[\sum_i \frac{g_i}{g_1} \frac{m_i^2}{m_1^2} K_2 \left(\frac{m_i}{T} \right) \right]^2}. \tag{73}$$

where K_1 (K_2) is the modified Bessel function of the second kind of order 1 (2), T is the temperature, s is the usual Mandelstam variable and

$$W_{\text{eff}} = \sum_{ij} \frac{p_{ij}}{p_{\text{eff}}} \frac{g_i g_j}{g_1^2} W_{ij}$$

$$= \sum_{ij} \sqrt{\frac{[s - (m_i - m_j)^2][s - (m_i + m_j)^2]}{s(s - 4m_1^2)}} \frac{g_i g_j}{g_1^2} W_{ij}. \tag{74}$$

Here,

$$p_{ij} = \frac{\left[s - (m_i + m_j)^2\right]^{1/2} \left[s - (m_i - m_j)^2\right]^{1/2}}{2\sqrt{s}}, \tag{75}$$

and the invariant annihilation rate is

$$W_{ij} = 4p_{ij}\sqrt{s}\sigma_{ij} = 4\sigma_{ij}\sqrt{(p_i \cdot p_j)^2 - m_i^2 m_j^2} = 4E_i E_j \sigma_{ij} v_{ij} \qquad (76)$$

and, finally, the effective momentum

$$p_{\text{eff}} = p_{11} = \frac{1}{2}\sqrt{s - 4m_1^2}. \qquad (77)$$

Since $W_{ij}(s) = 0$ for $s \leq (m_i + m_j)^2$, the terms in (74) are non-negative. For a two-body final state, W_{ij} is given by the expression

$$W_{ij}^{2-\text{body}} = \frac{|\mathbf{k}|}{16\pi^2 g_i g_j S_f \sqrt{s}} \sum_{\text{internal d.o.f.}} \int |\mathcal{M}|^2 \, d\Omega, \qquad (78)$$

that after some manipulations leads to (63). Here $\mathbf{k}$ is the final center-of-mass momentum, S_f is a symmetry factor equal to 2 for identical final particles.

So, what could the dark matter be? It turns out that in particle physics, there are hypothetical particles, like supersymmetric partners of ordinary particles, that have the right interaction strength and mass range to be promising dark matter candidates. In particular, the neutralino has all the properties of a good dark matter candidate. Since it is electrically neutral it does not emit or absorb radiation which makes it 'dark' (invisible matter is thus a better term than dark matter). The couplings of neutralinos are generally of weak interaction strength, but the large number of possible annihilation channels, which depends on the unknown supersymmetry breaking parameters, makes an exact prediction of mass and relic abundance uncertain. Scans of parameter space show, however, that a neutralino in the mass range between 30 GeV and a few TeV could give a relic density close to the critical density. We will later in these notes have much more to say about this.

3.2 Inflation

An important ingredient in today's cosmology is, as mentioned, the concept of inflation, which was introduced by Alan Guth in the early 1980s. Here we use the near constancy of H to write $a_2/a_1 = \exp(\int_{t_1}^{t_2} H(t)dt)$: that is, Einstein's equations including a cosmological constant read

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}. \qquad (79)$$

Here we have put the Λ term on the right hand side, which shows that a cosmological term acts as a stress-energy tensor, albeit with the unusual equation of state $p_{vac} = -\rho_{vac}$. (We have already used that one may trivially include vacuum energy in

the term proportional to G, with $\rho_\Lambda = \Lambda/(8\pi G)$.) The equation of state implies that the entropy density according to (55) is $s \sim \rho + p = 0$. This means that, when vacuum energy dominates, the entropy vanishes. This can be understood from statistical mechanics. Entropy is related to the total number of degrees of freedom, and the vacuum (at least if it is unique) is just one state, that is only one degree of freedom. Of course, the entropy that was in a patch before inflation will still be there after inflation—but it will be diluted by an exponentially large factor due to the expansion.

In the situation when the constant vacuum energy dominates the expansion, the Friedmann equation (2) becomes very simple:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{\Lambda}{3} \tag{80}$$

or

$$H = \frac{\dot{a}}{a} = \sqrt{\frac{\Lambda}{3}} = \text{const} \tag{81}$$

with the (de Sitter) solution

$$a \sim e^{Ht}. \tag{82}$$

In inflation, the expansion rate is constant, which causes an exponential growth of the scale factor.

In many models of inflation, the phase transition involving a scalar field, the inflation field, took place at temperatures around the hypothetical Grand Unification scale $T_{GUT} \sim 10^{15}$ GeV, at the corresponding Hubble time $H^{-1} \sim 10^{-34}$ s. If the universe stayed in the inflationary state for a short time, say 10^{-32} s, many e-folds of inflation took place. When inflation stopped, the huge vacuum energy of the inflation field went into creation of ordinary particles, and a reheating of the universe took place. The reheating temperature is of the order of the temperature of the phase transition, $T_{RH} \sim 10^{15}$ GeV if the inflation is strongly enough coupled to ordinary matter, as it is in many successful models of inflation.

Let us see what happened to a small region with radius of for example 10^{-23} cm before inflation. The entropy within that volume was only around 10^{14}, but after inflation the volume of the region has increased by a factor given by the cube of the scale factor, $(e^{100})^3 = 10^{130}$. Thus, after the entropy generated by reheating, the total entropy within the inflated region had grown to around 10^{144}. Entropy was generated because the equation of state changed from $p = -\rho$ to $p = \rho/3$, meaning that the entropy density $s \sim p + \rho$ increased dramatically.

This huge entropy increase solves many problems of cosmology. The "horizon problem"—i.e., how came that regions of the universe that are too far from each other to be causally connected today, still have exactly the same microwave background temperature—is solved since our whole observable universe arose from a very small thermalized volume before inflation, and the smooth region after inflation had sufficient entropy to encompass our observable universe.

During inflation the energy density, and the negative pressure, of the universe were constant, whereas the scale factor a increased exponentially. This means that the total Ω after inflation was exponentially close to unity. (Like a balloon which would inflate to be as big as the Earth would locally look very flat.) Thus, the present value should also be equal to unity with an accuracy of many decimal places, perhaps the most important successful prediction of inflation.

Even if $\Omega = 1$ is predicted, there is nothing telling us the subdivision of Ω into contributions from radiation, matter and vacuum energy. As we have noted, however, the 'natural' contribution of Ω_Λ is either extremely small or extremely large. Only during very brief epochs can Ω_Λ be of similar magnitude as the matter contribution Ω_M. This is actually a still unsolved problem, why is it that the energy density in matter ρ_M is about the same as ρ_Λ today?

The period of inflation and reheating is strongly non-adiabatic, since there was an enormous generation of entropy at reheating. After the end of inflation, the universe 'restarted' in an adiabatic phase with the standard conservation of aT, and it is because the universe automatically restarts from very special initial conditions given by inflation that the horizon and flatness problems are solved.

It is instructive to see how inflation can be produced in field theory. A Lagrangian density of the form

$$\mathcal{L} = \frac{1}{2}\partial^\mu \phi \partial_\mu \phi - V(\phi) \tag{83}$$

can be shown to give a contribution to the energy-momentum tensor $T^{\mu\nu}$ of the form

$$T^{\mu\nu} = \partial^\mu \phi \partial^\nu \phi - \mathcal{L}g^{\mu\nu}. \tag{84}$$

For a homogeneous state, the spatial gradient terms vanish, meaning that $T^{\mu\nu}$ becomes that of the perfect fluid type. If one would keep the gradient terms, one sees that they are divided by $a(t)^2$, which means that after a short period of inflation they are exponentially suppressed. The resulting ρ and p are

$$\rho = \frac{\dot{\phi}^2}{2} + V(\phi) \tag{85}$$

and

$$p = \frac{\dot{\phi}^2}{2} - V(\phi), \tag{86}$$

and we see that the equation of state $\rho = -p$ will be fulfilled if we can neglect the kinetic term $\sim \dot{\phi}^2$ (this is called "slow-roll" inflation).

The exact equations of motion of ϕ can be derived from the condition of vanishing covariant divergence of the energy-momentum tensor, $T_\nu^{\mu\nu} = 0$, which gives

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0 \tag{87}$$

This is similar to the equation of motion of a ball in a potential well with friction $\sim 3H\dot{\phi}$, and can be solved by elementary methods. We assume that at very high temperatures, $\phi = 0$ gives the locations of the minimum of the potential. Temperature dependent terms in the effective potential then generate another minimum for $\phi = \phi_{vac} \neq 0$, an example of what is called spontaneous symmetry breakdown. To produce a long enough period of inflation and a rapid reheating after inflation, the potential $V(\phi)$ has as mentioned to be of the "slow-roll" type, with the field spending a long time on the nearly flat, horizontal part of the potential. In the beginning, on the almost horizontal slow 'roll' towards a deep potential well, $\ddot{\phi}$ can be neglected, and the slow-roll equation of motion is

$$3H\dot{\phi} + V'(\phi) = 0, \tag{88}$$

together with the Friedmann equation

$$H^2 = \frac{8\pi G_N}{3}\left[\frac{1}{2}\dot{\phi}^2 + V(\phi)\right], \tag{89}$$

which during slow roll, when $\dot{\phi}^2$ is small, can be approximated by

$$H^2 = \frac{8\pi G_N}{3}V(\phi). \tag{90}$$

One can from this get an expression for the number N_ϕ of e-folds of the scale factor,

$$N_\phi \equiv \log\left(\frac{a_2}{a_1}\right) = \int H dt \sim \int_{\phi_1}^{\phi_2} \frac{V(\phi)}{V'(\phi)}d\phi. \tag{91}$$

Thus, for a large growth of the scale factor, $V(\phi)$ has to be very flat ($V'(\phi) \sim 0$). This may be unnatural except perhaps in some supersymmetric theories where 'flat directions' can occur because of the pattern of supersymmetry breaking. In a situation of such a slow roll of the inflation field, the exact form of the potential does not matter so much, and the relevant physics can be expressed in terms of the so-called slow-roll parameters

$$\varepsilon = -\frac{\dot{H}}{H^2} = 4\pi G_N \frac{\dot{\phi}^2}{H^2} = \frac{1}{16\pi G_N}\left(\frac{V'}{V}\right)^2 \tag{92}$$

$$\eta = \frac{1}{8\pi G_N}\left(\frac{V''}{V}\right) = \frac{V''}{3H^2} \tag{93}$$

where the second equation in (92) comes from taking the derivative of (89) and inserting into (87). The variable ε is a measure of the change of the Hubble expansion during inflation; for inflation to happen at all, $\varepsilon < 1$ is needed.

In the picture of the rolling ball, reheating corresponds to oscillations in the potential well. Thus, for enough entropy to be generated the well has to be rather steep. The problem of constructing a suitable potential is to simultaneously have it flat near $\phi = 0$ and steep near $\phi = \phi_{min}$.

A way to avoid a phase transition, and in fact the simplest model of inflation is the chaotic inflation model of Andrei Linde [12]. It relies on the fact that the key ingredient for inflation to occur is that the field is rolling slowly, so that the energy density is nearly constant during a sufficient number of e-foldings of the scale factor. Since the rolling is damped by the presence of the term proportional to H in (87), and H according to the Friedmann equation is given by the height of the potential (if kinetic terms can be neglected), inflation will be possible for any positive, power-law potential $V(\phi)$, for example the simplest $V(\phi) = \frac{1}{2}m^2\phi^2$, as long as the field values start out large. As Linde has argued, this may not be unreasonable since these initial values may be given stochastically ("chaotically") at the Planck epoch, and those regions where the field values are large start to inflate rapidly dominating the volume of the universe. There are also constructions relying on the existence of more than one scalar field, keeping the same general features but with changes in the details.

Since the value of the total energy density $\Omega = 1$ is found observationally in current measurements of the CMBR anisotropy which yield $\Omega = 1.003 \pm 0.010$, the most natural explanation seems to be that the universe has indeed gone through a period of inflation. An important test of inflation may be produced by the upcoming measurements from the Planck satellite of the detailed pattern of temperature fluctuations in the CMBR. Inflation predicts a nearly but not perfect scale-invariant spectrum of fluctuations (which is when the index of scalar fluctuations $n_s = 1$), and present measurements from WMAP give $n_s \sim 0.96$, in excellent agreement. Inflation could also have generated gravitational (tensor) waves during the phase transitions which would give a particular pattern ("B-modes") in the microwave sky. However, the amplitude of tensor to scalar fluctuations depends rather strongly on the model. It will be interesting to see whether the Planck satellite, when cosmological data are released in early 2013, will detect such a B-mode pattern.

4 Basic Cross Sections for Neutrinos and γ-Rays

Among the various messengers from the Galaxy and structures further away, neutrinos and γ-rays have the advantage that they follow straight lines (or to be more exact, geodesics; the deviations from straight lines can in almost all cases be safely neglected—exceptions are given for particles originating or travelling very near black holes). This is in contrast to all other cosmic rays, electrons, protons, nuclei, and antimatter (positrons, antiprotons and some antinuclei like antideuterons). Neutrons would in principle also travel rectilinearly apart from effects of their magnetic moment. However, their finite lifetime (of the order of 10 min in the rest frame) means that for energies less than a few TeV which is the energy range we will

generally be concerned with, they cannot travel over astrophysical distances. They β-decay to a proton, an electron and an (anti-)neutrino.

Although neutrinos and γ-rays (high-energy photons) are both encompassed in the Standard Model of particle physics and therefore in principle should interact with similar strengths given by gauge couplings, this is in practice not so. The reason is the difference that the photon is a massless, spin-1 gauge particle, i.e., a mediator of a force (the electromagnetic force, i.e., it couples to electrons and protons, and all other particles with electric charge) while the neutrino is a spin-1/2 matter particle which in turn interacts through the weak force mediated by the heavy W and Z bosons. The large, important difference of masses between weak bosons and the photon is due to the hitherto experimentally unverified, but hopefully soon to be verified mechanism, the Higgs mechanism. This breaks the gauge group of the Standard Model, leaving only the electromagnetic $U(1)_{em}$ unbroken and therefore the photon massless. It means that for energies up to 1 TeV or so, neutrinos have very small cross section, which however rises with energy, until the centre-of-mass energy is of the order of the W and Z masses, at which point neutrinos start to react roughly as strongly as photons. Let us now discuss in some more detail how some simple particle cross sections are computed.

4.1 Estimates of Cross Sections

The calculation of collision and annihilation cross sections, and decay rates of particles, is an important task in particle physics. Here we will present only a brief outline of how this is done, and focus on 'quick-and-dirty' estimates which may be very useful in cosmology and astrophysics. For the local microphysics in the FLRW model, only three interactions—electromagnetic, weak and strong—between particles need to be considered. The gravitational force is completely negligible between individual elementary particles—for instance, the gravitational force between the proton and the electron in a hydrogen atom is around 10^{40} times weaker than the electromagnetic force. However, gravity, due to its coherence over long range, still needs to be taken into account through its influence on the metric. This means that the dilution of number densities due to the time dependence of the scale factor $a(t)$ has to be taken into account. In the next section we will see how this is done.

Let us begin with the interaction strengths. The strength of the electromagnetic interaction is governed by the electromagnetic coupling constant g_{em}, which is simply the electric charge. As usual, we take the proton charge e as the basic unit and can thus write

$$g_{em} = Qe \qquad (94)$$

where Q is the charge of the particle in units of the proton charge (for a u-quark, for example, $Q_u = +2/3$). In our system of units,

$$\frac{e^2}{4\pi} \equiv \alpha_{em} \tag{95}$$

where α_{em} is the so-called fine structure constant which has the value of around $1/137$ at low energies.[2] The weak coupling constant is of similar magnitude:

$$g_w = \frac{e}{\sin\theta_W} \tag{96}$$

with θ_W the weak interaction (or Weinberg) angle, which has the numerical value $\sin^2\theta_W \sim 0.23$. The fact that the weak and electromagnetic coupling constants are of the same order of magnitude is of course related to the fact that they are unified in the Standard Model to the 'electroweak' interaction.

The coupling constant of the strong interaction, g_s, is somewhat higher. Also, it runs faster (it decreases) with energy than the electromagnetic coupling. At energies of a few GeV,

$$\alpha_s \equiv \frac{g_s^2}{4\pi} \sim 0.3 \tag{97}$$

Let us look at the Feynman diagram for a simple process like $e^+e^- \rightarrow \mu^+\mu^-$ (Fig. 1). The amplitude will be proportional to the coupling constants at both vertices, which in this case are both equal to e. The cross section, being proportional to the square of the amplitude, is thus proportional to $e^4 \propto \alpha^2$.

The total energy of the e^+e^- pair in the centre of momentum frame is $E_{cm}(e^+) + E_{cm}(e^-)$, which is conventionally noted with the Mandelstam variable as $\sqrt{s}$. Since the total momentum in this frame is zero, the four-momentum $p^\mu = (\sqrt{s}, 0, 0, 0)$ is identical to that of a massive particle of mass $M = \sqrt{s}$ which is at rest. Energy and momentum conservation then tells us that the photon in the intermediate state has this four-momentum. However, a freely propagating photon is massless, which means that the intermediate photon is virtual by a large amount. In quantum field theory one can show that the appearance of an intermediate state of virtual mass $\sqrt{s}$ for a particle with real rest mass M_i is suppressed in amplitude by a factor (called the propagator factor)

Fig. 1 A Feynman diagram representing the annihilation of an electron and a positron to a muon pair
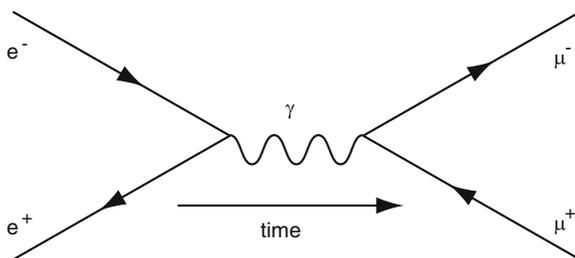
$$P(s) = 1/(s - m_i^2) \tag{98}$$

In this case ($m_\gamma = 0$), we have a factor of $1/s$. (If one does this rigorously, one should insert a small imaginary part in the denominator, which defines how the singularity on the mass shell is treated.) The outgoing particles (in this case the muons) have a large number of possible final states to enter (for example, all different scattering angles in the centre of momentum frame). This is accounted for by the so-called phase space factor ϕ, which generally grows as s for large energies. For the cross section σ

$$\sigma(e^+e^- \to \mu^+\mu^-) \propto \phi\left(\frac{\alpha^2}{s^2}\right) \tag{99}$$

with ϕ the phase space factor. If s is large compared to m_e^2 and m_μ^2, $\phi \propto s$, and

$$\sigma(e^+e^- \to \mu^+\mu^-) \sim \frac{\alpha^2}{s} \tag{100}$$

This is not an exact expression. A careful calculation (see next section) gives $4\pi\alpha^2/(3s)$, but it is surprisingly accurate and often accurate enough for the estimates we need in big bang cosmology.

Since the weak coupling strength is similar to the electromagnetic strength, the same formula is valid for, e.g., $\nu_e + e \to \nu_\mu + \mu$ which goes through W exchange (see Fig. 2). The only replacement we need is $1/s \to 1/(s - m_W^2)$ for the propagator, thus

$$\sigma(\bar{\nu}_e + e^- \to \bar{\nu}_\mu + \mu^-) \sim \frac{\alpha^2 s}{(s - m_W^2)^2} \tag{101}$$

When $s \ll m_W^2$, this gives $\sigma_w \sim \alpha^2 s/m_W^4$, which is a very small cross section, e.g., for MeV energies. One should notice, however, the fast rise with energy due to the factor s. This is the historical reason for the name 'weak interaction', which is really not appropriate at high energies (much larger than m_W), where the two types of cross sections become of similar size.
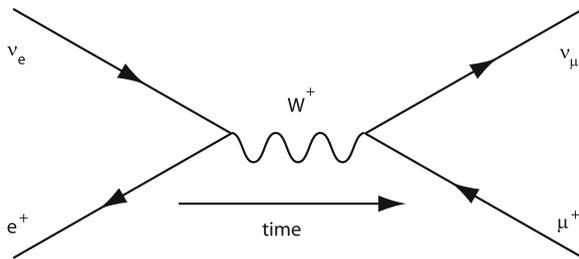
Fig. 2 A Feynman diagram representing the annihilation of an electron neutrino and a positron to a muon neutrino and a muon
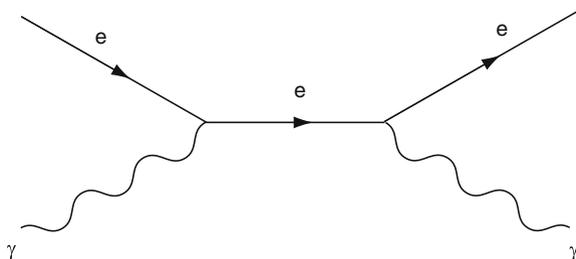
Fig. 3 A Feynman diagram representing the $\gamma e \to \gamma e$ process. In the classical limit, this is called Thomson scattering. The quantum version is called Compton scattering, and in the relativistic regime, the result is given by the Klein-Nishina formula

Note that once one remembers the factors of coupling constants and the propagators, the magnitude of cross sections can often be estimated by simple dimensional analysis. A cross section has the dimension of area, which in our units means $(\text{mass})^{-2}$. It is very useful to check that the expressions (100) and (101) have the correct dimensions.

A decay rate Γ can be estimated in a similar way. If a collection of identical unstable particles decay, their number density decreases as $e^{-\Gamma t}$ which means that Γ has the dimensions of mass.

A fermion has a propagator that behaves as $1/m$ (instead of $1/m^2$) at low energies. This means that the Thomson cross section $\sigma(\gamma e \to \gamma e)$ at low energies $E_\gamma \ll m_e$ can be estimated to be (see Fig. 3)

$$\sigma_T \equiv \sigma(\gamma e \to \gamma e) \sim \frac{\alpha^2}{m_e^2} \qquad (102)$$

4.2 Examples of Cross Section Calculations

The estimates we have given are in many cases sufficient for cosmological and astrophysical applications. However, there are cases when one would like to have a more accurate formula. We now provide only a couple of examples and summarize the general framework for calculation and the main results.

4.3 Definition of the Cross Section

The differential cross section $d\sigma/dt$ for $2 \to 2$ scattering $a + b \to c + d$ is given by the expression

$$\frac{d\sigma}{dt} = \frac{|\widetilde{T}|^2}{16\pi\lambda\left(s, m_a^2, m_b^2\right)} \tag{103}$$

where the Lorentz invariant Mandelstam variables are $s = (p_a + p_b)^2$, $t = (p_a - p_c)^2$ and $u = (p_a - p_d)^2$. $|\widetilde{T}|^2$ is the polarization-summed and squared quantum mechanical transition amplitude, while the function λ is defined as

$$\lambda(x, y, z) \equiv x^2 + y^2 + z^2 - 2xy - 2xz - 2yz. \tag{104}$$

For a $2 \rightarrow 2$ process, the kinematically allowed region in s is

$$s > (m_c + m_d)^2 \tag{105}$$

which can be understood from energy conservation: In the centre of momentum system (cms), where $\sqrt{s}$ corresponds to the total energy, at least the rest mass energy $m_c + m_d$ has to be provided.

The kinematical limits for t can be obtained from the condition $|\cos\theta_{ac}^{cms}| \leq 1$, with

$$\cos\theta_{ac}^{cms} = \frac{s(t - u) + (m_a^2 - m_b^2)(m_c^2 - m_d^2)}{\sqrt{\lambda(s, m_a^2, m_b^2)}\sqrt{\lambda(s, m_c^2, m_d^2)}}. \tag{106}$$

A typical calculation (following the treatment of [1]) involves computing the matrix element in terms of s and t and carrying out the t integration to obtain the total cross section.

In the one-photon exchange approximation, the cross section for the annihilation process $e^+e^- \rightarrow \mu^+\mu^-$ is

$$\sigma(e^+e^- \rightarrow \mu^+\mu^-) = \frac{2\pi\alpha^2}{s}\beta\left(1 - \frac{\beta^2}{3}\right) \tag{107}$$

where the only approximation made is to neglect m_e (this is allowed, since $m_e^2/m_\mu^2 \ll 1$). Here β is the velocity of one of the outgoing muons in the centre of momentum system, $\beta = \sqrt{1 - 4m_\mu^2/s}$. In the relativistic limit of $s \gg m_\mu^2$, ($\beta \rightarrow 1$), this becomes

$$\sigma\left(e^+e^- \rightarrow \mu^+\mu^-\right)_{\text{large } s} = \frac{4\pi\alpha^2}{3s} \tag{108}$$

in agreement with our simple estimate (100).
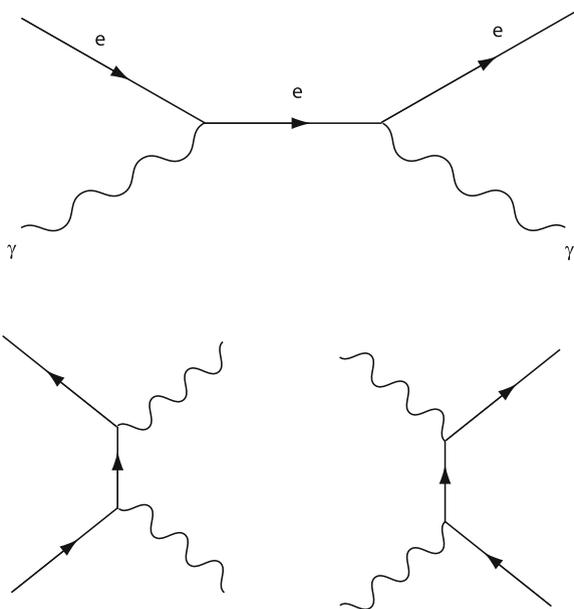
Fig. 4 The same Feynman diagram can, after rotation of the external legs, describe both $e\gamma \rightarrow e\gamma, e^+e^- \rightarrow \gamma\gamma$, and $\gamma\gamma \rightarrow e^+e^-$. Here time is as usual flowing from *left* to *right*

4.4 The $\gamma\gamma ee$ System

By different permutations of the incoming and outgoing particles, the basic $\gamma\gamma ee$ interaction (shown in Fig. 3) can describe all of the astrophysically important processes (see the contributions by F. Aharonian and C. Dermer in this volume) $\gamma\gamma \rightarrow e^+e^-, e^+e^- \rightarrow \gamma\gamma$, and $\gamma e^{\pm} \rightarrow \gamma e^{\pm}$, see Fig. 4.

For $\gamma\gamma \rightarrow e^+e^-$ the result is

$$\sigma\left(\gamma\gamma \rightarrow e^+e^-\right) = \frac{\pi\alpha^2}{2m_e^2}\left(1-\beta^2\right)\left[\left(3-\beta^4\right)\ln\left(\frac{1+\beta}{1-\beta}\right) + 2\beta\left(\beta^2-2\right)\right]$$

(109)

where β now is the velocity of one of the produced electrons in the centre-of-momentum frame, $\beta = \sqrt{1-4m_e^2/s}$. Near threshold, i.e. for small β, the expression in square brackets can be series expanded to $2\beta + \mathcal{O}(\beta^2)$, and thus

$$\sigma\left(\gamma\gamma \rightarrow e^+e^-\right)_{\text{small }\beta} \simeq \frac{\pi\alpha^2}{m_e^2}$$

(110)
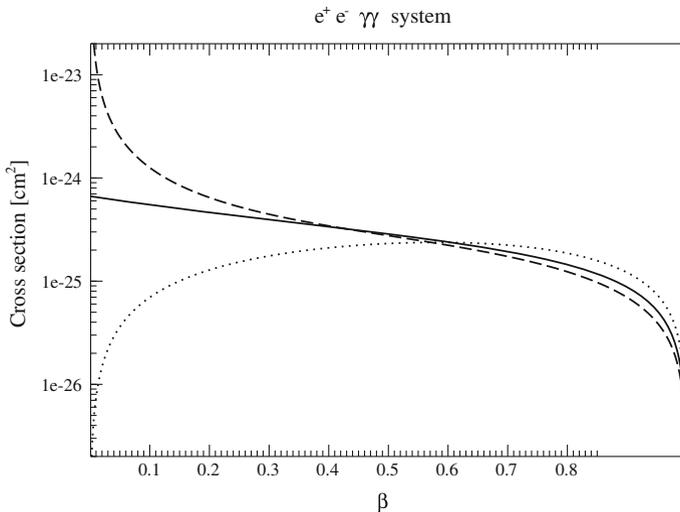
Fig. 5 The cross sections (in cm^2) for photon–photon annihilation (*dotted line*), $e^+e^- \rightarrow \gamma\gamma$ (*dashed line*) and Compton scattering (*solid line*) as a function of the cms velocity β of the electron

In the other extreme, $\beta \rightarrow 1$,

$$\sigma\left(\gamma\gamma \rightarrow e^+e^-\right)_{s \gg 4m_e^2} \simeq \frac{4\pi\alpha^2}{s}\left[\ln\left(\frac{\sqrt{s}}{m_e}\right) - 1\right] \tag{111}$$

One could in fact have guessed most of this to a fair amount of accuracy by the simple dimensional and vertex-counting rules. At low energy, the only available mass scale is m_e, so the factor α^2/m_e^2 could have been guessed for that reason. The factor β could also have been inferred with some more knowledge of non-relativistic partial wave amplitudes. At low energy, the $\ell = 0$ (S-wave) amplitude should dominate, and this contributes to the cross section proportionally to β. A partial wave ℓ contributes to the total cross section with a term proportional to $\beta^{2\ell+1}$. We see from Eq. (107) that in the case of $e^+e^- \rightarrow \mu^+\mu^-$ the S-wave dominates at low energy, but when $\beta \rightarrow 1$, the P-wave contribution is 1/3. At high energy, when m_e can be neglected, the dimensions have to be carried by s. Only the logarithmic correction factor in (111) could not have been easily guessed.

These formulas show that the $\gamma\gamma \rightarrow e^+e^-$ cross section rises from threshold to a maximum at intermediate energies and then drops roughly as $1/s$ at higher β, i.e., higher cms energy in the process (see Fig. 5).

The results for the reverse process $e^+e^- \rightarrow \gamma\gamma$ are of course extremely similar. Now, the process is automatically always above threshold. For $\beta \rightarrow 0$ (with β now the velocity of one of the incoming particles in the cm-system, still given by the formula $\beta = \sqrt{1 - 4m_e^2/s}$), the flux factor $\sim 1/\beta$ implicit in Eq. (103) diverges. Since the

outgoing photons move away with $\beta = c = 1$ there is no partial-wave suppression factor, and we can thus expect the cross section at low energy to behave as

$$\sigma \left(e^+ e^- \to \gamma\gamma\right)_{\text{low energy}} \sim \frac{\alpha^2}{\beta m_e^2} \tag{112}$$

and the high-energy behavior by the same formula, with m_e^2 replaced by s (and possibly a logarithmic factor). These expectations are borne out by the actual calculation, which gives

$$\sigma \left(e^+ e^- \to \gamma\gamma\right) = \frac{\pi\alpha^2 \left(1 - \beta^2\right)}{2\beta m_e^2} \left[\frac{3 - \beta^4}{2\beta} \ln\left(\frac{1 + \beta}{1 - \beta}\right) - 2 + \beta^2\right] \tag{113}$$

Note the similarity with Eq. (109). The $1/\beta$ behavior of the cross section (see the dashed curve in Fig. 5) was noted by Arnold Sommerfeld in the 1930s, and he showed how one can make an improved calculation valid at very small velocities by not only treating the annihilating particles as plane waves, but using wave functions appropriate for the attractive Coulomb interaction between the electron and positron. He thereby described a generic mechanism, the so-called Sommerfeld enhancement mechanism, which recently has played an important role for dark matter calculations, as we will see later in Sect. 6.3.

Compton and Inverse Compton Scattering

As the final example, we consider Compton scattering $\gamma + e^- \to \gamma + e^-$. Historically, this was first computed for an incoming beam of photons of energy ω which hit electrons at rest. Later on, the related process of a very high energy electron or positron colliding with a low-energy photon (such as coming from the cosmic microwave background, or from infrared or optical radiation created in stellar processes) and upscattering that photon to high, maybe GeV energy or higher, has been found to be very important in astrophysics. Despite being really one and the same process, the latter situation is often referred to as the inverse Compton or IC process. In fact, the inverse Compton process is one purely leptonic process of creating high-energy γ-rays, and could be important for the emission γ-rays in several cases, such as AGNs, GRBs and various supernova remnants. However, to completely prove such a scenario, it is important to search for, or give upper limits on, neutrino emission. In competing hadronic models of emission, where γ-rays mainly come from π^0 decays, one should also have essentially the same amount of charged pions which decay into a leptons and neutrinos. Also for some "leptophilic" models of dark matter, where electrons and muons are main annihilation products, inverse Compton processes may be quite important, e.g., near the galactic centre where radiation fields are large.

For scattering of a photon by an angle θ with respect to the incident photon direction, the outgoing photon energy ω' is given by energy-momentum conservation

$$\omega' = \frac{m_e \omega}{m_e + \omega \left(1 - \cos\theta\right)} \tag{114}$$

In this frame, the unpolarized differential cross section, the Klein-Nishina formula as it was first computed by Klein and Nishina shortly after Dirac had presented his equation describing relativistic electrons (and positrons), is

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{2m_e^2} \left(\frac{\omega'}{\omega}\right)^2 \left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right] \tag{115}$$

Integrated over all possible scattering angles this gives the total cross section

$$\sigma(\gamma + e \to \gamma + e) = \frac{\pi\alpha^2 \left(1 - \beta\right)}{m_e^2 \beta^3}$$

$$\times \left[\frac{4\beta}{1+\beta} + \left(\beta^2 + 2\beta - 2\right)\ln\left(\frac{1+\beta}{1-\beta}\right) - \frac{2\beta^3 \left(1 + 2\beta\right)}{(1+\beta)^2}\right] \tag{116}$$

where β is now the incoming electron velocity in the centre of momentum frame, $\beta = (s - m_e^2)/(s + m_e^2)$. If one expands this result around $\beta = 0$, one recovers the Thomson scattering result

$$\sigma_{\text{Thomson}} = \frac{8\pi\alpha^2}{3m_e^2} \sim 6.65 \cdot 10^{-25} \text{ cm}^2 \tag{117}$$

and the large-s, so-called Klein Nishina regime gives

$$\sigma_{\text{KN}} = \frac{2\pi\alpha^2}{s} \left[\ln\left(\frac{s}{m_e^2}\right) + \frac{1}{2}\right] \tag{118}$$

We see that for photon energies much larger than m_e—that is, in the Klein-Nishina regime—the Compton cross section falls quite rapidly.

In the classical Compton scattering situation, the outgoing photon energy is always less than the incoming one. Thus, energetic photons traveling through a gas of cold electrons will be 'cooled' by Compton scattering. In the IC case (for example for the cosmic microwave background radiation passing through a galaxy cluster with hot gas) energetic electrons may instead transfer energy to photons, thereby 'heating' them. For CMBR this is called the Sunyaev-Z'eldovich effect, and has a large range of applicability (for instance, it has recently been used to find galaxy clusters).

When computing actual numbers for the cross sections (which should have the dimensions of area) in our units, a useful conversion factor is

$$1 \text{ GeV}^{-2} = 0.389 \cdot 10^{-27} \text{ cm}^2 \tag{119}$$

In Fig. 5 the numerical results are summarized. The cross sections are shown (in cm^2) for $\gamma\gamma \to ee$, $ee \to \gamma\gamma$ and $\gamma e \to \gamma e$ as a function of the cms velocity β of the electron. We see in the figure the different behaviour at low cms velocity already discussed, but that they show a similar decrease at high energy.

Another process of great astrophysical importance is bremsstrahlung. By this is meant the emission of photons from charged particles which are accelerated or decelerated. If this acceleration is due to circular motion in a magnetic field, the term synchrotron radiation is used. Through these processes (unlike Compton scattering) the number of photons can change. This is needed, for instance in the early universe, if thermal equilibrium is to be maintained, since the number density of photons has to vary, as it depends strongly on temperature. Most of the produced photons have very low energy (long wavelength). If fast electrons pass through a region where synchrotron radiation and bremsstrahlung occur, these low-energy photons may be upscattered in energy through the inverse Compton process. This may for example explain the observations of very high-energy photons in active galactic nuclei.

For a detailed discussion of these and other quantum electrodynamic (QED) processes, see standard textbooks in quantum field theory, for example, [13], or a simplified treatment along the lines given here, in [1]. And, of course, for good examples of the use of these processes in astrophysics, see the accompanying lectures by F. Aharonian and C. Dermer in this volume.

4.5 Processes Involving Hadrons

Since protons and neutrons belong to the most common particles in the universe, it is of course of great interest to compute processes where these and other hadrons (such as pions) are involved. This is, however, not easy to do from first principles. The reason that in the previous section we could compute so accurately weak and electromagnetic processes is that we could use perturbation theory (as summarized, for example, in Feynman diagrams). The expansion parameter, the electroweak gauge coupling constant g or rather $\alpha_{ew} = g^2/(4\pi) \sim 10^{-2}$, is small enough that a lowest-order calculation is enough to obtain very accurate results.

In quantum chromodynamics (QCD), which also is a gauge theory just as QED, we also have a coupling constant α_s. Due to the fact that the gauge group of QCD is SU(3), which involves self-interactions of the 8 spin-1 gluons, there are important differences. We say that QCD is a non-abelian gauge theory whereas QED is based on the abelian group U(1) with only one spin-1 gauge field, the photon. One consequence of this difference is that QCD has what is called asymptotic freedom meaning that the coupling strength which is of order unity at a few hundred MeV, "runs" to smaller values for large energies. The energy scale is set, for example, by the energy or momentum transfer Q ($Q^2 \equiv -t$ with t the usual Mandelstam variable) in the process. Thus, for processes with large Q^2, we should be able to use low-order perturbative QCD, although with lower accuracy than for QED due to the possible importance of higher-order corrections. At low energies when the QCD coupling

becomes of the order unity perturbation theory breaks down. In the nonperturbative regime we have to rely on empirical methods, such as "QCD sum rules" [14] or large computer simulations, where one tries to solve QCD by formulating it as a field theory on a lattice. Although the problem is notoriously difficult, the agreement between the spectrum of hadrons, i.e., the masses and quantum numbers of the lowest-lying states with experimentally measured quantities, is quite satisfactory for the most recent numerical simulations [15].

For processes like proton-proton scattering at low energies, the picture of strong interactions being due to the exchange of gluons breaks down. Instead one may approximate the exchange force as being due to pions and other low-mass mesons with surprisingly good results (this is in fact what motivated Yukawa to predict the existence of pions). If one wants to make crude approximations of the strong interaction cross section in this regime, $\sigma_s \sim 1/m_\pi^2$ is a good estimate.

In the perturbative regime at high Q^2, the scattering, for example, of an electron off a proton ('deep inelastic scattering') can be treated by the successful parton model. Here, the momentum of a hadron at high energy is shared between its different constituents. Among the constituents are of course the quarks that make up the hadron (two u and one d quarks in the case of the proton), i.e., the valence quarks. In addition, there may be pairs of quarks and antiquarks produced through quantum fluctuations at any given "snapshot" of the hadron. The incoming exchange photon sent out from an electron in ep scattering may hit these "sea quarks", which will therefore contribute to the scattering process.

Since the partons interact with each other, they can share the momentum of the proton in many ways. Thus, there will be a probability distribution, $f_i(x, Q^2)$, for a parton of type i (where i denotes any quark, antiquark or gluon) to carry a fraction x of the proton momentum. These functions cannot be calculated from first principles. However, once determined (by guess or by experimental information from various scattering and annihilation processes) at a particular value of Q_0^2, the evolution of the structure functions with Q^2 can be predicted. This analysis, as first convincingly performed by Altarelli and Parisi [16], gives rise to a predicted variation of the deep inelastic scattering probability with Q^2 (so-called scaling violations) which has been successfully compared with experimental data. The success of the perturbative QCD program, including the running of α_s in agreement with the asymptotic freedom prediction, and the agreement of scaling violations in several processes with data, resulted in a Nobel Prize for Gross, Wilczek and Politzer in 2004.

With the successful QCD parton model, we can now compute many electromagnetic and weak processes, including those when hadrons are involved. For instance, the neutrino-proton scattering cross section is be given by the scattering of a neutrino on a single quark or antiquark. This calculation is easily done in a way similar to how we computed the $\bar{\nu}_e + e^- \rightarrow \bar{\nu}_\mu + \mu^-$ cross section. The only change is that the contributions from all partons have to be summed over, and an integral of x performed.

As an example, we give the expression for the electromagnetic cross section $p + p \rightarrow \mu^+ + \mu^-$, which is called the Drell-Yan process, in the QCD parton model. The fundamental process must involve charged partons, i.e., quarks (since

we assume that strong interactions dominate and thus neglect the weak contribution), $q + \bar{q} \rightarrow \gamma^* \rightarrow \mu^+ + \mu^-$, with the (valence) quark taken from one of the protons and the (sea) antiquark from the other. The momentum transfer in the process is $Q^2 = \hat{s}$, where $\hat{s} = (p_{\mu^+} + p_{\mu^-})^2$. We know from (108) that the parton level cross section is $4\pi\alpha e_q^2 / 3\hat{s}$ (where we have to take into account that the quark charge e_q is not the unit charge). Since the parton from proton 1 carries the fraction x_1 and that from proton 2 x_2 of the respective parent proton, $\hat{s} = x_1 x_2 s$, with $s = (p_1 + p_2)^2$. The total cross section for producing a muon pair of momentum transfer $\hat{s}$ is thus

$$\frac{d\sigma}{d\hat{s}} = \frac{4\pi\alpha^2}{3\hat{s}} k_c \sum_q e_i^2 \int_0^1 dx_1 \int_0^1 dx_2$$
$$\times \left[f_q(x_1, \hat{s}) f_{\bar{q}}(x_2, \hat{s}) + f_{\bar{q}}(x_1, \hat{s}) f_q(x_2, \hat{s}) \right] \delta(\hat{s} - x_1 x_2 s)$$

Here k_c is a colour factor, which takes into account that for a given quark of a given colour, the probability to find in the other proton an antiquark with a matching (anti-) colour is 1/3. Thus, in this case $k_c = 1/3$. In the reverse process, $\mu^+ + \mu^- \rightarrow q + \bar{q}$, all the quark colours in the final state have to be summed over (each contributes to the cross section), so in that case $k_c = 3$.

4.6 Neutrinos

Neutrinos are the neutral counterparts of the charged leptons: e, μ and τ. There are therefore three types of "active" neutrinos in the Standard Model of particle physics: ν_e, ν_μ and ν_τ. Neutrinos are fermions, i.e., spin-$\frac{1}{2}$ particles. Apart from their possible gravitational interactions, they interact with matter only through the exchange of the mediators of the weak force, the W and Z bosons. They are fundamental particles without constituents, as far as is known, have extremely small masses and lack electric charge. Among the most spectacular events in astrophysics are supernova explosions. In a few seconds, more energy is released in neutrinos from the forming neutron star than all the electromagnetic emission from an entire galaxy over a decade.

Neutrinos would provide an important contribution to the total energy density of the universe if they had a mass in the eV range. Present-day analyses of the microwave background and the matter distribution favour as we mentioned, cold dark matter over hot dark matter which eV-scale Standard Model neutrinos would constitute. However, neutrinos are plausibly part of dark matter, as one of the most important discoveries of the last 15 years has been the definitive demonstration that neutrinos are not massless. It has been shown that neutrinos species can transform into each other through quantum mechanical mixing, which means that in principle they could also decay into each other, e.g., by emitting a photon. The huge mass scale of the W and Z bosons means, however, that the lifetime even for the heaviest neutrino is much longer than the age of the universe, so they are effectively stable. Their production through thermal processes in the early universe would then mean

that their number density is of the same order of magnitude today as that of the microwave background photons, and since they are electrically neutral they qualify as dark matter. This was in fact one of the most studied candidates for dark matter in the 1970s, when one did not know how many different neutrino types there were (this was fixed in the 1990s to 3 standard species, thanks to the CERN LEP experiments), neither very much about their mass. Today's knowledge give estimates of far less than a percent for their contribution to Ω.

There is of course a possibility that there may exist neutrinos with weaker couplings than standard model ones, and if they have mass in the keV range they would qualify as "warm" dark matter, that would still be allowed from structure formation. However, this involves some fine-tuning to get the correct relic density and to avoid other experimental bounds [17].

Neutrinos are also very important information carriers from violent astrophysical processes, which is why neutrino astrophysics is a very active field of research at the present time. An attractive property of neutrinos is in fact their feeble interactions at low energies, which means that they may penetrate regions with dense matter, e.g., the solar interior, or near remnants of supernovae, without being absorbed. Where other particles become trapped or can only propagate through very slow diffusive processes (for instance, it takes on the order of a million years for a photon created near the centre of the sun to diffuse out to the solar surface), neutrinos are able to escape. Neutrinos can thus connect regions of matter that would otherwise be isolated from each other. Because they are almost massless, they move effectively with the speed of light, which makes energy transfer (i.e., radiative heat conduction) very efficient, e.g., from the interior of the sun. Unfortunately, the fact that neutrinos are so weakly interacting, also means that they are extremely difficult to detect. As of today, the only neutrinos of astrophysical interest that have been detected are those created in the fusion processes in the solar interior, and the exceptional supernova in 1987, where a handful of neutrinos was detected a few hours before it was spotted optically in the Southern sky.

Neutrino interactions with matter are divided into two kinds, neutral current (NC) interactions mediated by the neutral Z bosons, and charged current (CC) interactions involving the exchange of W^+ and W^- bosons. NC interactions are responsible for annihilation reactions involving neutrinos,

$$e^+ + e^- \rightarrow \nu_\mu + \bar{\nu}_\mu$$

or for example, and elastic scattering interactions such as

$$\nu_\mu + e^- \rightarrow \nu_\mu + e^-.$$

In CC interactions there is a change of fermion type, of "flavour". For example, an antineutrino can be absorbed by a proton, producing a neutron and a positron in the final state. This comes about because at the parton level a u-quark in the proton is changed into a d-quark, which means it is transformed to a neutron. In this process

charge is transferred, both for the leptons as the neutrino becomes a charged lepton, and for the hadrons as the positively charged proton becomes a neutron.

4.7 Neutrino Interactions

For the neutrino process (the flavour-changing charged current interaction) $\bar{\nu}_e e^- \rightarrow \bar{\nu}_\mu \mu^-$ the cross section at low energies (but still high enough to produce the heavier muon) is

$$\sigma\left(\bar{\nu}_e e^- \rightarrow \bar{\nu}_\mu \mu^-\right) \sim \frac{g_w^4 s}{96\pi m_W^4} \tag{120}$$

Before it was known that W bosons existed, Enrico Fermi[3] had written a phenomenological theory for weak interactions with a dimensionful constant (the Fermi constant) G_F. The relation between Fermi's constant and the gauge theory quantities is

$$\frac{G_F}{\sqrt{2}} = \frac{g_w^2}{8m_W^2} \simeq 1.166 \cdot 10^{-5} \text{ GeV}^{-2} \tag{121}$$

Using the Fermi constant, the cross section can now be written

$$\sigma\left(\bar{\nu}_e e^- \rightarrow \bar{\nu}_\mu \mu^-\right) = \frac{G_F^2 s}{3\pi}. \tag{122}$$

We note that the cross section rises with $s \simeq 2E_\nu m_e$ and thus linearly with neutrino energy. When s starts to approach m_W^2, the W propagator $1/(s - m_W^2)$ has to be treated more carefully. It can be improved by writing it in the so-called Breit-Wigner form

$$\frac{1}{s - m_W^2} \rightarrow \frac{1}{s - m_W^2 + i\Gamma m_W} \tag{123}$$

where Γ is the total decay width (around 2 GeV) of the W. We see from this that a substantial enhancement of the cross section is possible for $s \simeq m_W^2$. This is an example of a resonant enhancement in the s-channel. For a target electron at rest, this resonance occurs at around 6.3 PeV and is sometimes called the Glashow resonance. If astrophysical sources exist which produce electron antineutrinos with such high energies, the prospects of detecting them would be correspondingly enhanced. However, well above the resonance, the cross section will again start to decrease like $1/s$, just as in the electromagnetic case, $e^+ e^- \rightarrow \mu^+ \mu^-$.

It should be noted that the latter process, $e^+ e^- \rightarrow \mu^+ \mu^-$, also receives a contribution from an intermediate Z boson. At low energies this is negligible, but due

---

[3] Enrico Fermi is well-known today as the one who has been honoured by giving the γ-ray satellite, FERMI, its name. This has to do with his description of acceleration processes in astrophysics.

to the resonant enhancement it will dominate near $s \simeq m_Z^2$. This is the principle behind the Z studies performed at the LEP accelerator at CERN (where all other fermion-antifermion pairs of the Standard Model were also produced except for $t\bar{t}$, which was not kinematically allowed). In a full calculation, the two contributions have to be added coherently and may in fact interfere in interesting ways, producing for example, a backward-forward asymmetry between the outgoing muons.

A detailed calculation for neutrino energies above around 5 MeV shows that the total cross section for the reaction $\nu_X e^- \to \nu_X e^-$ is well approximated by [18]:

$$\sigma_{\nu e} = C_X \cdot 9.5 \cdot 10^{-45} \cdot \left(\frac{E_\nu}{1 \text{ MeV}} \right) \text{ cm}^2 \tag{124}$$

where the flavour-dependent constants C_X are

$$C_e = 1 \tag{125}$$

and

$$C_\mu = C_\tau = \frac{1}{6.2} \tag{126}$$

The cross section is larger for electron neutrinos as they can, unlike the other neutrino species, couple to the electrons in the target through both NC and CC interactions.

Laboratory experiments have, so far, not succeeded in directly measuring the mass of any neutrino. Instead, the negative results have been expressed in the form of upper limits, due to the finite resolution of the experiments. The best (lowest) upper limits on the mass of the electron neutrino come from the studies of the electron energy spectrum in tritium decay:

$$^3\text{H} \to\, ^3\text{He} + e^- + \bar{\nu}_e$$
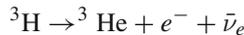
As the minimum amount of energy taken by the ν_e is its mass, the end-point energy of the emitted electron is a measurement of m_{ν_e}. According to these experiments the mass of the electron neutrino is lower than 3 eV at the 95 % confidence level [19]. With the KATRIN experiment being constructed in Karlsruhe, Germany, one hopes to decrease this upper limit (or find a non-zero value) by an order of magnitude [20].

The discovery of the tau neutrino was announced in 2000 by the DONUT collaboration at Fermilab, through appearance in charm meson decays in photographic emulsion.

Mixing of neutrino species is a very interesting quantum mechanical effect which may occur if the weak-interaction eigenstates ν_e, ν_μ and ν_τ are not the mass eigenstates that propagate in vacuum. We can then express a flavour or weak-interaction neutrino eigenstate, ν_f, as a linear superposition of orthogonal mass eigenstates, ν_m:

$$| \nu_f \rangle = \sum_m c_{fm} | \nu_m \rangle.$$

Of course, all three neutrinos may mix, but it is instructive to see what happens if just two of the neutrinos mix, e.g, $\nu_\mu \leftrightarrow \nu_e$ mixing with mixing angle θ. The time evolution of a muon neutrino wave function, produced e.g. in pion decays, with momentum p is then

$$|\nu_e(t)\rangle = -\sin\theta e^{-iE_1 t}|\nu_1\rangle + \cos\theta e^{-iE_2 t}|\nu_2\rangle \tag{127}$$

where E_1 and E_2 are the energies of the two mass eigenstates. Two energy levels arise if ν_1 and ν_2 have different masses, for the same momentum, p. Then, for small neutrino masses $m_i \ll E_i$,

$$E_i = p + \frac{m_i^2}{2p} \tag{128}$$

The probability $P(\nu_e \to \nu_e) = |\langle \nu_e \,|\, \nu_e \rangle|^2$, that an electron neutrino remains a ν_e after a time t then becomes

$$P(\nu_e \to \nu_e) = 1 - \sin^2(2\theta)\sin^2\left[\frac{1}{2}(E_2 - E_1)t\right] \tag{129}$$

For very small neutrino masses, using (128),

$$P(\nu_e \to \nu_e) = 1 - \sin^2(2\theta)\sin^2\left[\left(\frac{m_2^2 - m_1^2}{4E}\right)t\right] \tag{130}$$

where E is the energy of ν_e.

Thus the probability the electron neutrino transforms to a muon neutrino at a time t is

$$P(\nu_e \to \nu_\mu) = \sin^2(2\theta)\sin^2\left[\frac{\Delta m^2}{4E}t\right] \tag{131}$$

where $\Delta m^2 = |m_2^2 - m_1^2|$.

From (131) it is seen that the probability function for flavour change oscillates, with an amplitude given by $\sin^2(2\theta)$ and oscillation frequency $\sim \Delta m^2/E$. This is now the generally accepted reason for the deficit of solar electron neutrinos, as deduced by combining data from the Super-Kamiokande experiment in Japan (most recently [21]), which sees the deficit of electron neutrinos, with SNO in Canada, which has measured the neutral current cross section, which shows no deficit [22]. As the neutral current has the same strength for all three neutrinos, this is strong evidence that the total flux is unchanged, but the flux of electron neutrinos has decreased due to mixing.

Numerically, the oscillation length becomes

$$L_\nu = 1.27\left(\frac{E}{1\,\text{MeV}}\right)\left(\frac{1\,\text{eV}^2}{\Delta m^2}\right)\,\text{m.} \tag{132}$$

In fact, a direct proof that oscillations occur in the (anti-) neutrino sector is given by recent results from the KamLAND experiment [23], where reactor antineutrinos have been shown to oscillate over more than one period of oscillation in L/E.

4.8 Atmospheric Neutrinos

Neutrinos are copiously produced in the atmosphere by hadronic and muonic decays following the interaction of cosmic rays with atomic nuclei N,

$$
\begin{cases}
p/n + N \rightarrow \pi^+/K^+ + \cdots \\
\qquad\qquad\qquad \pi^+/K^+ \rightarrow \mu^+ + \nu_\mu \\
\qquad\qquad\qquad\qquad\qquad \mu^+ \rightarrow e^+ + \bar{\nu}_\mu + \nu_e, \\
p/n + N \rightarrow \pi^-/K^- + \cdots \\
\qquad\qquad\qquad \pi^-/K^- \rightarrow \mu^- + \bar{\nu}_\mu \\
\qquad\qquad\qquad\qquad\qquad \mu^- \rightarrow e^- + \nu_\mu + \bar{\nu}_e
\end{cases}
\tag{133}
$$

Studying the end result of these reactions one expects that there are about twice as many muon neutrinos than electron neutrinos produced in the atmosphere:

$$
\frac{\phi_{\nu_\mu} + \phi_{\bar{\nu}_\mu}}{\phi_{\nu_e} + \phi_{\bar{\nu}_e}} = 2
\tag{134}
$$

This expectation holds at low energies. At higher energies, additional effects have to be taken into account: for example, the competition between scattering and decay of the produced pions, and also time dilation. As the energy spectrum of the primary nuclei reaches out to $\sim 10^{20}$ eV, one expects neutrinos to be produced up to comparable energies.

Due to the complicated chain of reactions in the cascade, computer simulations are needed to find the differential spectrum of atmospheric neutrinos. One finds that there is a broad peak around 0.1 GeV (~ 1 cm^{-2} s^{-1}) and at very high energies, E_ν much larger than TeV, a falling flux $\sim E^{-3.7}$.

The cross section for neutrino-nucleon interactions in a target can be calculated by inserting the nucleon mass instead of m_e in our previous example. In the region of the maximum flux of atmospheric neutrinos the cross section is $\sigma_{\nu N} \sim 10^{-39}$ cm^2. The Super-K experiment showed that also atmospheric neutrinos oscillate, and that part of the muon neutrinos disappear due to $\nu_\mu \rightarrow \nu_\tau$ mixing taking place. (Due to the high $\tau^\pm$ lepton mass (1.8 GeV) the ν_τs generated by mixing will not have enough energy on average to make the charged current interaction $\nu_\tau + N \rightarrow \tau + X$ kinematically possible. Their contribution to the neutral current events is too small to be easily detected.) The Super-K data on the angular and energy dependence of the oscillation is in excellent agreement with the E/L ratio given by (132).

4.9 Neutrinos as Tracers of Particle Acceleration

A kiloton-size detector is necessary to observe neutrinos from sources as close as the Earth's atmosphere, or the Sun (this actually gave a Nobel Prize to Davies and Koshiba, in 2002). To be able to search other astrophysical objects, the required detector mass becomes very large, megaton or even gigaton (or volume of order km^3).

Consider the $\nu_\mu \to \mu$ charged current weak interaction in a medium,

$$\nu_\mu + N \to \mu + \cdots,$$

where N is a nucleon in the medium in or surrounding the detector. The muon range rises with energy, and around 1 TeV (10^{12} eV) it is more than 1 km. The detection area is therefore greatly enhanced at high energies. In water or ice, a good approximation of the muon range as a function of energy is given by

$$R_\mu \approx 2.5 \ln \left(2 \cdot \frac{E_\mu}{1 \text{ TeV}} + 1 \right) \text{ km} \tag{135}$$

The produced muon conserves, on average, the direction of the incoming neutrino. The average of the square of the ν_μ-μ angle is approximately (see [1])

$$\sqrt{\langle \theta^2 \rangle} \approx 2 \left(\frac{1 \text{ TeV}}{E_\nu} \right)^{\frac{1}{2}} \quad \text{deg.} \tag{136}$$

The cross section for neutrino interaction with a fixed target rises linearly with energy. Neutrino telescopes for very high energies become efficient at a few GeV, where the product of the neutrino-matter cross section and the muon range rises approximately as E_ν^2. Above 1 GeV, the induced flux of muons from atmospheric neutrinos, for example, is about 1 m^{-2} year^{-1}.

This detection scheme does not work as well for other types of neutrinos. Electrons (from $\nu_e + N \to e + \cdots$) have a very short range as they lose energy through radiative processes, due to their small mass. On the other hand τ leptons, the heaviest known charged leptons with $m_\tau = 1.78$ GeV, are produced in charged current interactions of ν_τ, but they are very short lived ($t_\tau \sim 10^{-13}$ s). Therefore they are not suitable for detection, except for the fraction of times where the τ decays into $\mu \bar{\nu}_\mu \nu_\tau$, which happens in roughly 20 % of the cases. However, in large neutrino detectors such as the IceCube, one may perhaps detect ultra-high-energy electron and τ neutrino events by the intense cascade of light that is produced by secondary electrons, positrons and photons. In the case of τ neutrinos, special relativity may help to produce a good signature. If sources of PeV (10^{15} eV) τ neutrinos exist, the produced charged τ lepton would have a relativistic γ factor as large as

$$\gamma \sim \frac{E_\nu}{m_\tau} \sim 10^6 \tag{137}$$

which means, thanks to time dilation, that in the detector reference frame the τ lepton will travel a distance $\gamma c t_\tau \sim 10$ m. The "double bang" created by the charged current interaction and the subsequent decay of the τ lepton, separated by 10 m, would be the signature of PeV τ neutrinos.

Since neutrinos oscillate, very energetic τ neutrinos could be produced by mixing with muon neutrinos created in high-energy pion decays in cosmic accelerators. This is, e.g., the basis for the experiments OPERA [24] at Gran Sasso and MINOS [25] at Fermilab, where a few events of produced τ leptons have in fact been reported.

In present detectors, only neutrino-induced muons moving upwards in the detectors (or downwards but near the horizon) are safe tracers of neutrino interactions. Most muons moving downwards have their origin in cosmic-ray nuclei interacting with the Earth's atmosphere and produce a very difficult background. At the surface of the Earth, the flux of downward-going muons produced in the atmosphere is about 10^6 times larger than the flux of neutrino-induced upward-moving muons.

By going underground, the material (rock, water, ice, etc.) above the detector attenuates the flux of atmospheric muons considerably. In addition, if it is experimentally possible to select events where a muon is moving upwards the Earth itself acts as a filter since only neutrino-induced muons can be produced upward-going close to the detector.

4.10 AMANDA, IceCube and Direct Detection of WIMPs

Neutrinos may give clues to the dark matter problem in another way than just being a small part of the dark matter due to their tiny mass. If the dark matter has a component that is massive and weakly coupled (electrically neutral) it will be non-relativistic at freeze-out, which is of course the WIMP paradigm of cold dark matter. A good template for a dark matter WIMP candidate is as we mentioned the lightest supersymmetric particle—plausibly the neutralino χ (see Sect. 5 for more details).

Neutralinos (or other WIMPs) have interactions with ordinary matter which are equally as small as those of neutrinos. However, since they move with non-relativistic velocity there is a chance that they become gravitationally trapped inside, for example, the Sun or the Earth. A neutralino scattering e.g., in the Sun will lose energy and fall further inside the solar volume, and successive scatterings in the solar medium will soon make it lose more and more energy. In the end, neutralinos will assemble near the centre. As they are their own antiparticles (they are Majorana fermions), they can annihilate with each other, resulting in ordinary particles (quarks, leptons, gauge particles).

As the annihilation rate is proportional to the scattering rate, and the interior of the Earth is almost entirely spin-0 nuclei, constraints on the spin-independent scattering rate from experiments described below in Sect. 8 mean that neutrinos from the center

of the Earth are not a very promising signal for canonical WIMPs. However, as the
Sun consists to some 75 % of single protons (i.e., hydrogen nuclei) with spin-1/2,
spin-dependent scattering is important and searching for neutrinos from the Sun
stands well in competition with other experiments. We will return also to this later.

Most of the annihilation products in the Sun create no measurable effect; they are
just stopped and contribute somewhat to the energy generation. However, neutrinos
have the unique property that they can penetrate the whole Sun without being much
absorbed, at least for WIMPs less massive than a few hundred GeV. An annihilating
neutralino pair of mass m_χ would thus give rise to high-energy neutrinos of energy
around $m_\chi/3$ or so (the reason that $E_\nu \neq m_\chi$ is that other particles created in the
annihilation process share the energy). The signal of high-energy neutrinos (tens to
hundreds of GeV—to be compared with the 'ordinary' MeV solar neutrinos) from
the centre of the Sun would be an unmistakable signature of WIMP annihilation.

The detection of muons in IceCube, for instance, relies on the Cherenkov effect.
This coherent emission of light follows a characteristic angle given by the Mach
relation

$$\cos\theta = \frac{1}{\beta n}$$

where β is the speed of the particle traversing the medium in units of the speed of
light and n is the index of refraction of the medium. The Cherenkov effect takes
place when

$$\beta > \frac{1}{n}.$$

Cherenkov radiation constitutes a very small fraction of the total energy loss of a
charged particle as it crosses a medium. The superluminal condition is fulfilled only
between the UV and near-infrared region of the electromagnetic spectrum. In water
or ice, for example, where the index of refraction for UV and optical wavelengths
averages around 1.3, the Cherenkov radiation cut-off in the UV region is around
70 nm. The differential energy loss into Cherenkov photons in water or ice is a few
percent of the total differential energy loss of a charged track moving with a speed
very close to c.

4.11 Water and Ice Cherenkov Telescopes

Neutrinos can thus be detected indirectly by the Cherenkov radiation from charged
leptons and hadrons produced in neutrino interactions with matter. The extremely
large detector volumes needed to detect neutrinos from distances beyond our Sun
makes the use of any other material than water or ice very difficult.

A typical detector consists of an array of light sensors (photomultipliers, PM)
with good time resolution (∼1 ns) distributed in the medium. The pattern of the hit
PMs, and relative arrival times, are then used to fit the direction of the particle that

generated the Cherenkov radiation. The correlation between the original direction of the neutrino and the produced charged lepton means that one may reconstruct the direction of the incoming neutrino.

Antares is a good prototype, for a larger detector being planned with the working name KM3NET, near Toulon in the Mediterranean. The AMANDA experiment at the South Pole was similarly an excellent working prototype, where the disadvantages related to the remote location of the telescope were compensated by the virtues of the glacier ice, found to be the clearest natural solid on Earth. The Cherenkov photons emitted along the path of a muon at some wavelengths can be selected hundreds of metres away from the muon track.

The AMANDA detector was a great success, but was too small and has recently been abandoned, replaced by a much larger detector, the IceCube, with 80 strings encompassing roughly a cubic kilometer of ice. Construction was finished in 2010, and at that time also a smaller and denser inset, the DeepCore detector, was completed. This allows a lower detection energy threshold which is particularly beneficial for the WIMP search. Unfortunately, despite the heroic effort to build the first large neutrino detector in this remote location, no astrophysical neutrino source including WIMPs has yet been detected, but it is only a year that data have been collected (for a recent review of dark matter detection in neutrino telescopes, see [26]).

5 Supersymmetric Dark Matter

As we have mentioned several times already, one of the prime candidates for the non-baryonic cold dark matter particle is provided by the lightest supersymmetric particle, most likely the lightest neutralino χ. Even if it would be that supersymmetry were not realized in nature, the neutralino is still important as a nice, calculable template for a generic WIMP.

In most versions of the low-energy theory which results from the largely unknown mechanism of supersymmetry breaking, there is a conserved multiplicative quantum number, R-parity:

$$R = (-1)^{3(B-L)+2S},\qquad(138)$$

where B is the baryon number, L the lepton number and S the spin of the particle. This implies that $R = +1$ for ordinary particles and $R = -1$ for supersymmetric particles. In fact, for phenomenological reasons, this symmetry is required, as its conservation protects the proton from decaying through the mediation of R-symmetry breaking interactions. The R-symmetry means that supersymmetric particles can only be created or annihilated in pairs in reactions of ordinary particles. It also means that a single supersymmetric particle can only decay into final states containing an odd number of supersymmetric particles. In particular, this makes the lightest supersymmetric particle stable, since there is no kinematically allowed state with negative R-parity which it can decay to. This is of course of utmost importance for the dark matter problem. Also other WIMP models of dark matter needs some

mechanism to prevent decay, and the simplest mechanism is a discrete symmetry like the double-valued (Z_2) R-symmetry. (Another reason for stability could be the quantum numbers of the particles in the theory. There are in fact models with high spin or isospin multiplets which also have a stable particle which could act as dark matter [27].)

Pair-produced neutralinos in the early universe which left thermal equilibrium as the universe kept expanding should, due to their stability, have a non-zero relic abundance today. If the scale of supersymmetry breaking is related to that of electroweak breaking, the neutralino will act as a WIMP and therefore a dark matter candidate with a relic density of the same order of magnitude as the value implied by the WMAP measurements. This is a very elegant and economical method to solve two of the most outstanding problems in fundamental science, dark matter and the unification of the basic forces, if they have a common element of solution—supersymmetry.

5.1 Supersymmetric Dark Matter Particles

If R-parity is conserved, the lightest supersymmetric particle should be stable. The most plausible candidate is the lightest neutralino χ. As we will see in Sect. 5.2 it is a mixture of the supersymmetric partners of the photon, the Z and the two neutral CP-even Higgs bosons present in the minimal extension of the supersymmetric standard model (MSSM). It is electrically neutral and thus neither absorbs nor emits light, and is stable, surviving since earliest epoch after the big bang. Its gauge couplings and mass means that for a large range of parameters in the supersymmetric sector a relic density is predicted in the required range to explain the observed $\Omega_\chi h^2 \sim 0.11$. Its electroweak couplings to ordinary matter also means that its existence as dark matter in our galaxy's halo may be experimentally tested.

Unfortunately, very little is known about how supersymmetry is broken (for a discussion, see [28]), and therefore any given supersymmetric model contains a large number of unknown parameters (of the order of 100). Such a large parameter space is virtually impossible to explore by present-day numerical methods, and therefore simplifying assumptions are needed. Fortunately, most of the unknown parameters such as CP violating phases influence the properties relevant for cosmology, and for detection, very little.

Usually, when scanning the large space of a priori unknown parameters in supersymmetry, one thus makes reasonable simplifying assumptions and accepts solutions as cosmologically appropriate if they give a neutralino relic density in the range

$$0.09 \lesssim \Omega_\chi h^2 \lesssim 0.12 \tag{139}$$

Recently, there has been a number of analyses where the relic density, and other parameters or experimental quantities known within some error bounds are allowed to vary. By using so-called Markov Chain Monte Carlo methods (MCMC), one can get a "global fit" of the best-fit models using statistical methods [29, 30]. Usually,

one employs what is called a Bayesian method which needs some assumption about the prior distribution of probabilities. In the case of mass parameters one may, for instance, choose linear or logarithmic scans. If experimental data are good enough, it can be shown that the choice of priors is not crucial. However, so far there has been a lack of experimental information, meaning that the predicted most likely regions in parameter space may depend quite sensitively on priors (see, e.g., [31]). Hopefully, the situation may soon change with new results from the LHC. A drawback of the method of global fits is that it is very computer intensive, meaning that only very simplified models of supersymmetry have been fully investigated so far.

Besides its interesting implications for cosmology, the motivation from particle physics for supersymmetric particles at the electroweak mass scale has become stronger due to the apparent need for 100 GeV–10 TeV scale supersymmetry to achieve unification of the gauge couplings in view of LEP results. (For an extensive review of the literature on supersymmetric dark matter up to mid-1995, see Ref. [32]. More recent reviews are [33] and [34]).

A great virtue of supersymmetry at the phenomenological level is that it gives an attractive solution to the so-called hierarchy problem, which is to understand why the electroweak scale at a few hundred GeV is so much smaller than the Planck scale $\sim 10^{19}$ GeV despite the fact that there is nothing in non-supersymmetric theories to cancel the severe quadratic divergences of loop-induced mass terms. In supersymmetric theories, the partners of differing spin would exactly cancel those divergencies (if supersymmetry were unbroken). Of course, supersymmetric models are not guaranteed to contain good dark matter candidates, but in the simplest models R-parity is conserved and the neutralino naturally appears as a good candidate.

The MSSM

The minimal supersymmetric extension of the standard model (MSSM) is defined by the particle content and gauge couplings required by supersymmetry and a gauge-invariant superpotential. Thus, to each particle degree of freedom in the non-supersymmetric Standard Model, there appears a supersymmetric partner with the same charge, colour etc, but with the spin differing by half a unit. The only addition to this doubling of the particle spectrum of the Standard Model concerns the Higgs sector. It turns out that the single scalar Higgs doublet is not enough to give masses to both the u- and d-like quarks and their superpartners (since supersymmetry forbids using both a complex Higgs field and its complex conjugate at the same time, which one does in the non-supersymmetric Standard Model). Thus, two complex Higgs doublets have to be introduced. After the usual Higgs mechanism, three of these states disappear as the longitudinal components of the weak gauge bosons leaving five physical states: two neutral scalar Higgs particles H_1 and H_2 (where by convention H_2 is the lighter state), one neutral pseudoscalar state A, and two charged scalars $H^{\pm}$.

The Z boson mass gets a contribution from the vacuum expectation values v_i of both of the doublets,

$$\left\langle H_1^1 \right\rangle = v_1, \qquad \left\langle H_2^2 \right\rangle = v_2, \tag{140}$$

with $g^2(v_1^2 + v_2^2) = 2m_W^2$. One assumes that vacuum expectation values (VEVs) of all other scalar fields (in particular, squark and sleptons) vanish, as this avoids color and/or charge breaking vacua.

The supersymmetric theory also contains the supersymmetric partners of the spin-0 Higgs doublets. In particular, two Majorana fermion states, higgsinos, appear as the supersymmetric partners of the electrically neutral parts of the H_1 and H_2 doublets. These can mix quantum mechanically with each other and with two other neutral Majorana states, the supersymmetric partners of the photon (the photino) and the Z (the zino). When diagonalizing the mass matrix of these four neutral Majorana spinor fields (neutralinos), the lightest physical state becomes an excellent candidate for cold dark matter.

The one-loop effective potential for the Higgs fields has to be used to obtain realistic Higgs mass estimates. The minimization conditions of the potential allow one to trade two of the Higgs potential parameters for the Z boson mass $m_Z^2 = \frac{1}{2}(g^2 + g'^2)(v_1^2 + v_2^2)$ (where $g = e/\sin\theta_W$, $g' = e/\cos\theta_W$) and the ratio of VEVs, $\tan\beta$. This ratio of VEVs

$$\tan\beta \equiv \frac{v_2}{v_1} \tag{141}$$

always enters as a free parameter in the MSSM, although it seems unlikely to be outside the range between around 1.1 and 60, with some preference for the higher values. The third parameter can further be re-expressed in terms of the mass of one of the physical Higgs bosons, for example m_A.

5.2 Higgs and Supersymmetry

At the ATLAS and CMS experiments at the CERN Large Hadron Collider (LHC), a discovery of the Higgs particle has not yet been claimed (by the end of 2011), as the statistical significance is still below the wanted $5\,\sigma$ (standard deviations). However, there are intriguing indications showing up at more than $3\,\sigma$ at a mass value around 125 GeV. If this would stand when more statistics is gathered in 2012, it could means that the Standard Model of particles and fields would be completed with a most wanted spin-0 boson, the Higgs particle. Moreover, a mass below 130 GeV is a firm prediction of supersymmetry, so it may also show the way to a whole new phenomenology, including a very interesting dark matter candidate—the lightest supersymmetric particle, generally thought to be the neutralino. As mentioned in Sect. 5.1, this is a quantum mechanical mixture of the supersymmetric partner of the photon, the neutral weak gauge boson Z and the neutral spin-1/2 partners of each of the two Higgs doublets which are needed by supersymmetry. In supersymmetric

theories, the most likely dark matter candidate is a quantum mechanical superposition, called the neutralino χ of electrically neutral supersymmetric fermions.

Of course, if the 125 GeV Higgs also signals the presence of supersymmetry, then a rich spectrum of particles, several of which may be in reach kinematically at the LHC, is expected. Even if supersymmetry is not realized in nature, it will continue to play a role as an important template for dark matter, as the neutralino is a very attractive, calculable candidate for a generic WIMP. We will return to this later.

5.3 The Neutralino Sector

The neutralinos $\tilde{\chi}_i^0$, of which the lightest is the dark matter candidate, are linear combination of the neutral gauge bosons $\tilde{B}$, $\tilde{W}_3$ (or equivalently $\tilde{\gamma}$, $\tilde{Z}$) and of the neutral higgsinos $\tilde{H}_1^0$, $\tilde{H}_2^0$. In this basis, their mass matrix

$$
\mathcal{M} = \begin{pmatrix}
M_1 & 0 & -\frac{g'v_1}{\sqrt{2}} & +\frac{g'v_2}{\sqrt{2}} \\
0 & M_2 & +\frac{gv_1}{\sqrt{2}} & -\frac{gv_2}{\sqrt{2}} \\
-\frac{g'v_1}{\sqrt{2}} & +\frac{gv_1}{\sqrt{2}} & 0 & -\mu \\
+\frac{g'v_2}{\sqrt{2}} & -\frac{gv_2}{\sqrt{2}} & -\mu & 0
\end{pmatrix}
\tag{142}
$$

can be diagonalized to give four neutral Majorana states,

$$
\tilde{\chi}_i^0 = a_{i1}\tilde{B} + a_{i2}\tilde{W}^3 + a_{i3}\tilde{H}_1^0 + a_{i4}\tilde{H}_2^0
\tag{143}
$$

($i = 1, 2, 3, 4$) the lightest of which, χ_1^0 or simply χ, is then the candidate for the particle making up the dark matter in the universe.

The coefficients in (143) are conveniently normalized such that for the neutralino

$$
\sum_{j=1}^{4} |a_{1j}|^2 = 1.
\tag{144}
$$

The properties of the neutralino are quite different depending on whether it consists mainly of gaugino ($j = 1, 2$) or higgsino ($j = 3, 4$) components. We therefore define a parameter, Z_g, which tells the size of the gaugino fraction:

$$
Z_g = \sum_{j=1}^{2} |a_{1j}|^2.
\tag{145}
$$

A neutralino is often said to be gaugino-like if $Z_g \gtrsim 0.99$, higgsino-like if $Z_g \lesssim 0.01$, and mixed otherwise.

To simplify, one often makes a diagonal ansatz for the soft supersymmetry-breaking parameters in the sfermion sector. This allows the squark mass matrices to be diagonalized analytically. Such an ansatz implies the absence of tree-level flavor changing neutral currents (FCNC) in all sectors of the model. In models inspired by low-energy supergravity with a universal scalar mass at the grand-unification (or Planck) scale the running of the scalar masses down to the electroweak scale generates off-diagonal terms and tree-level FCNC's in the squark sector.

In the estimates of detection rates here, we will adhere to a purely phenomenological approach, where the simplest unification and scalar sector constraints are assumed, and no CP violating phases outside those of the Standard Model, but no supergravity relations are used. This reduces the number of free parameters to be scanned over in numerical calculations to seven: $\tan\beta$, M_1, μ, m_A, and three parameters related to the sfermion sector (the exact values of the latter are usually not very important). In fact, on can reduce the number of parameters further by choosing, e.g., explicit supergravity models, but this only corresponds to a restriction to a subspace of our larger scan of parameter space. In fact, data from the LHC have already excluded large sectors of the simplified models.

The non-minimal character of the Higgs sector may well be the first experimental hint at accelerators of supersymmetry. At tree level, the H_2^0 mass is smaller than m_Z, but radiative (loop) corrections are important and shift this bound by a considerable amount. However, even after allowing for such radiative corrections it can hardly be larger than around 130 GeV. When there were some weak indications of a Higgs signature at 140 GeV in LHC data reported in mid-2011, this looked like bad news for the MSSM. However, with further data, the preferred mass is now around 125 GeV, which is easily accommodated.

5.4 Experimental Limits

The successful operation of the CERN accelerator LHC at centre of mass energies above 7 TeV without observing any supersymmetric particles, in particular squarks of gluinos, puts important constraints on the parameters of the MSSM. However, it may be that the mass scale of neutralinos is decoupled from the other supersymmetric particle masses (e.g, in "split susy" models [35]).

It has proven to be very difficult, however, to put very tight lower limits on the mass of the lightest neutralino, because of the multitude of couplings and decay modes of the next-to-lightest supersymmetric particle. The lightest neutralino can in general only be detected indirectly in accelerator experiments through the missing energy and momentum it would carry away from the interaction region.

The upper limit of dark matter neutralino masses in the MSSM is of the order of 7 TeV [11]. Above that mass, which is still far from the unitarity bound of 340 TeV [36], the relic density becomes larger than the allowed WMAP upper limit. To get

values for the lightest neutralino mass larger than a few hundred GeV, however, some degree of "finetuning" is necessary. (On the other hand, we have seen that for the other important unknown part of the energy density of the universe, the cosmological constant Λ, a "finetuning" of many orders of magnitude seems also to be necessary.)

By making additional well-motivated but not mandatory restrictions on the parameter space, such as in supergravity-inspired models or in simplified constrained MSSM models (CMSSM), one gets in general masses below 600 GeV [37, 38] for the lightest neutralino, but as mentioned these models are feeling some tension from early LHC data.

5.5 Supersymmetry Breaking

Supersymmetry is a mathematically beautiful theory, and would give rise to a very predictive scenario, if it were not broken in an unknown way which unfortunately introduces a large number of unknown parameters.

Breaking of supersymmetry has to be present since no supersymmetric particle has as yet been detected, and unbroken supersymmetry requires particles and sparticles to have the same mass. This breaking can be achieved in the MSSM by a soft potential which does not reintroduce large radiative mass shifts (and which indicates that the lightest supersymmetric particles should perhaps not be too much heavier than the 250 GeV electroweak breaking scale). The origin of the effective low-energy potential need not be specified, but it is natural to believe that it is induced through explicit breaking in a hidden sector of the theory at a high mass scale. The supersymmetry breaking terms could then be transmitted to the visible sector through gravitational interactions.

Another possibility is that supersymmetry breaking is achieved through gauge interactions at relatively low energy in the hidden sector. This is then transferred to the visible sector through some messenger fields which transform non-trivially under the Standard Model gauge group. However, we shall assume the "canonical" scenario in most of the following.

Since one of the virtues of supersymmetry is that it may establish grand unification of the gauge interactions at a common mass scale, a simplifying assumption is often used for the gaugino mass parameters,

$$M_1 = \frac{5}{3} \tan^2 \theta_W M_2 \simeq 0.5 M_2, \tag{146}$$

and

$$M_2 = \frac{\alpha_{\text{em}}}{\sin^2 \theta_W \alpha_s} M_3 \simeq 0.3 M_3, \tag{147}$$

where θ_W is the weak mixing angle, $\sin^2 \theta_W \approx 0.22$.

When using the minimal supersymmetric standard model in calculations of relic dark matter density, one should make sure that all accelerator constraints on supersymmetric particles and couplings are imposed. In addition to the significant restrictions on parameters given by LEP and LHC, the measurement of the $b \rightarrow s\gamma$ quark flavour changing process is providing important bounds, since supersymmetric virtual particles may contribute significantly to this loop-induced decay. There are also constraints arising if one wants to attribute the slightly abnormal value of $(g-2)$ for the muon [39] to supersymmetric contributions from virtual particles. The relic density calculation in the MSSM for a given set of parameters is nowadays accurate to a few percent or so [2].

5.6 Other Supersymmetric Candidates

Although the neutralino is considered by most workers in the field to be the preferred supersymmetric dark matter candidate, we mention briefly here also some other options.

If the axion, the spin-0 pseudoscalar field which solves the strong CP problem exists, and if the underlying theory is supersymmetric, there should also exist a spin-1/2 partner, the axino. If this is the lightest supersymmetric particle and is in the multi-GeV mass range, it could compose the cold dark matter of the universe (for a review, see [40]).

A completely different type of supersymmetric dark matter candidate is provided by so-called Q-balls [41], non-topological solitons predicted to be present in many versions of the theory. These are produced in a non-thermal way and may have a large lepton or baryon number. They could produce unusual ionization signals in neutrino telescopes, for example. However, the unknown properties of their precise formation mechanism means that their relic density may be far below the level of observability, and a value around the observationally favoured $\Omega_Q \sim 0.22$ may seem fortuitous (for a recent review of the physics of Q-balls, see [42]).

Of course, the possibility of dark matter being non-supersymmetric WIMPs still remains. However, the interaction cross sections should then be quite similar as for supersymmetric particles. Since, the rates in the MSSM are completely calculable once the supersymmetry parameters are fixed, these particles, in particular neutralinos, serve as important templates for reasonable dark matter candidates when it comes to designing experiments with the purpose of detecting dark matter WIMPs.

6 Detection Methods for Neutralino Dark Matter

The ideal situation would appear if supersymmetry were discovered at accelerators, so that direct measurements of the mass of the lightest supersymmetric particle, its couplings and other properties could be performed. This would give a way to

check from very basic principle if this particle is a good dark matter candidate—if it is electrically neutral and has the appropriate mass and couplings to give the required relic density to provide $\Omega_\chi h^2 \sim 0.11$. So far, no signal of supersymmetry has been found at either LEP, Fermilab, or LHC. An indirect piece of evidence for supersymmetry would be the discovery of a Higgs particle below around 130 GeV, since this is the maximal value of the lightest Higgs mass after radiative corrections, in the MSSM. In the non-supersymmetric Standard Model the Higgs could be much heavier. It is indeed encouraging that the first signs of the Higgs at LHC seems to correspond to a mass of 125 GeV.

If we assume a local neutralino halo density of $\rho_\chi = \rho_\odot \sim 0.4$ GeV/cm^3 [43], and a typical galactic velocity of neutralinos of $v/c \sim 10^{-3}$, the flux of particles of mass 100 GeV at the location of a detector at the Earth is roughly 10^9 m^{-2} s^{-1}. Although this may seem as a high flux, the interaction rate has to be quite small, since the correct magnitude of $\Omega_\chi h^2 \sim 0.11$ is only achieved if the annihilation cross section, and therefore by expected crossing symmetry also the scattering cross section, is of weak interaction strength.

The rate for direct detection of galactic neutralinos, integrated over deposited energy assuming no energy threshold, is

$$R = \sum_i N_i n_\chi \langle \sigma_{i\chi} v \rangle, \qquad (148)$$

where N_i is the number of nuclei of species i in the detector, n_χ is the local galactic neutralino number density, $\sigma_{i\chi}$ is the neutralino-nucleus elastic cross section, and the angular brackets denote an average over v, the neutralino speed relative to the detector.

The most important non-vanishing contributions for neutralino-nucleon scattering are the scalar-scalar coupling giving a spin-independent (SI) effective interaction, and the spin-dependent (SD) axial-axial interaction,

$$\mathcal{L}_{\text{eff}} = f_{SI} \left(\bar\chi \chi \right) \left(\bar{N} N \right) + f_{SD} \left(\bar\chi \gamma^\mu \gamma^5 \chi \right) \left(\bar{N} \gamma_\mu \gamma^5 N \right). \qquad (149)$$

Usually, it is the spin-independent interaction that gives the most important contribution in realistic target materials (such as Na, Cs, Ge, I, or Xe), due to the enhancement caused by the coherence of all nucleons in the target nucleus.

The neutralino-nucleus elastic cross section can be written as

$$\sigma_{i\chi} = \frac{1}{4\pi v^2} \int_0^{4m_{i\chi}^2 v^2} dq^2 G_{i\chi}^2(q^2), \qquad (150)$$

where $m_{i\chi}$ is the neutralino-nucleus reduced mass, q is the momentum transfer and $G_{i\chi}(q^2)$ is the effective neutralino-nucleus vertex. One may write

$$G_{i\chi}^2(q^2) = A_i^2 F_{SI}^2(q^2) G_{SI}^2 + 4\lambda_i^2 J(J+1) F_{SD}^2(q^2) G_{SD}^2, \qquad (151)$$

which shows the coherent enhancement factor A_i^2 for the spin-independent cross section. A reasonable approximation for the gaussian scalar and axial nuclear form factors is

$$F_{SI}(q^2) = F_{SD}(q^2) = \exp(-q^2 R_i^2 / 6\hbar^2), \tag{152}$$

$$R_i = (0.3 + 0.89 A_i^{1/3}) \text{ fm}, \tag{153}$$

which gives good approximation to the integrated detection rate [44] (but is less accurate for the differential rate [45]). Here λ_i is related to the average spin of the nucleons making up the nucleus. For the relation between G_{SI}, G_{SD} and f_{SI}, f_{SD} as well as a discussion of the several Feynman diagrams which contribute to these couplings, see e.g. [46–48]. One should be aware that these expressions are at best approximate. A more sophisticated treatment (see discussion and references in [32]) would, however, plausibly change the values by much less than the spread due to the unknown supersymmetric parameters.

For a target consisting of N_i nuclei the differential scattering rate per unit time and unit recoil energy E_R is given by

$$S_0(E_R) = \frac{dR}{dE_R} = N_i \frac{\rho_\chi}{m_\chi} \int d^3v \, f(\boldsymbol{v}) \, v \frac{d\sigma_{i\chi}}{dE_R}(v, E_R). \tag{154}$$

The nuclear recoil energy E_R is given by

$$E_R = \frac{m_{i\chi}^2 v^2 (1 - \cos\theta^*)}{m_i} \tag{155}$$

where θ^* is the scattering angle in the center of mass frame. The range and slope of the recoil energy spectrum is essentially given by non-relativistic kinematics. For a low-mass χ, the spectrum is steeply falling with E_R; interaction with a high-mass χ gives a flatter spectrum with higher cutoff in E_R.

The total predicted rate integrated over recoil energy above a given (generally detector-dependent) threshold can be compared with upper limits coming from various direct detection experiments. In this way, limits on the χ-nucleon cross section can be obtained as a function of the mass m_χ [49–52]. The cross section on neutrons is usually very similar to that on protons, so in general only the latter is displayed below. Major steps forward have been taken in recent years. For example, the CDMS-II experiment [53] and XENON100 [54] have been pushing the limits down by a large factor, reaching now 10^{-44} cm^2 for masses around 50 GeV. This together with a larger detector mass (for XENON, 1 tonne is presently being installed) and other improvements will enable a thorough search well beyond the present range of WIMP-nucleon cross sections. In Europe there are several other ambitious endeavours underway, such as DARWIN, a large liquid noble gas detector, and EURECA, a solid state detector.

The rate in (154) is strongly dependent on the velocity v of the neutralino with respect to the target nucleus. Therefore an annual modulation of the counting rate is in principle possible, due to the motion of the Earth around the Sun [55]. One can thus write

$$S(E_R, t) = S_0(E_R) + S_m(E_R) \cos\left[\omega(t - t_0)\right], \tag{156}$$

where $\omega = 2\pi/365 \, \text{days}^{-1}$. Starting to count time in days from January 1st, the phase is $t_0 = 153$ days since the maximal signal occurs when the direction of motion of the Earth around the Sun and the Sun around the galactic center coincide maximally, which happens on June 2nd every year [55]. Similarly, the counting rate is expected to be the lowest December 2nd every year. Here $S_0(E_R)$ is the average differential scattering rate in Eq. (154) and $S_m(E_R)$ is the modulation amplitude of the rate. The relative size of $S_m(E_R)$ and $S_0(E_R)$ depends on the target and neutralino mass as well as on E_R. Typically $S_m(E_R)$ is of the order of a few percent of $S_0(E_R)$, but may approach 10% for small m_χ (below, say, 50 GeV) and small E_R (below some 10 keV).

Since the basic couplings in the MSSM are between neutralinos and quarks, there are uncertainties related to the hadronic physics step which relates quarks and gluons with nucleons, as well the step from nucleons to nuclei. These uncertainties are substantial, and can plague all estimates of scattering rates by at least a factor of 2, maybe even by an order of magnitude [56, 57]. The largest rates, which as first shown in [46] could be already ruled out by contemporary experiments, are generally obtained for mixed neutralinos, i.e. with Z_g neither very near 0 nor very near 1, and for relatively light Higgs masses (since Higgs bosons mediate a scalar, spin-independent exchange interaction).

The experimental situation is becoming interesting as several direct detection experiments after many years of continuing sophistication are starting to probe interesting parts of the parameter space of the MSSM, given reasonable, central values of the astrophysical and nuclear physics parameters. Perhaps most striking is the 8 σ evidence for an annual modulation effect claimed to be seen in the NaI experiment DAMA/LIBRA [58] (see Sect. 8 where present data are summarized).

Many of the present day detectors are severely hampered by a large background of various types of ambient radioactivity or cosmic-ray induced activity (neutrons are a particularly severe problem since they may produce recoils which are very similar to the expected signal). A great improvement in sensitivity would be acquired if one could use directional information about the recoils. There are some very interesting developments also along this line, but a full-scale detector is yet to be built.

Direction-sensitive detectors would have an even bigger advantage over pure counting experiments if the dark matter velocity distribution is less trivial than the commonly assumed maxwellian.

6.1 Indirect Searches

Besides these possibilities of direct detection of supersymmetric dark matter (with even present indications of the existence of a signal [58]), one also has the possibility of indirect detection through neutralino annihilation in the galactic halo. This is becoming a promising method thanks to very powerful new detectors for cosmic gamma rays and neutrinos planned and under construction. Also, with time more has become known about the distribution of dark matter thanks to very ambitious N-body simulations [59–61], and a large amount of substructure has been found. This would enhance indirect detection, as it is proportional to the line-of-sight integral of the square of the number density.

There has been a balloon-borne detection experiment [62], with increased sensitivity to eventual positrons from neutralino annihilation, where an excess of positrons over that expected from ordinary sources was found. However, due to the rather poor quality of the data, it was not very conclusive.

In 2008, however, this changed completely when the data on the positron to electron ratio, rising with energy, from the satellite PAMELA was presented [63]. Somewhat later, FERMI-LAT reported a rise above the expectation from secondary production (by cosmic rays) also for the sum of positrons and electrons [64].

An unexpectedly high ratio of positrons over electrons was measured by PAMELA, in particular in the region between 10 and 100 GeV, where previously only weak indications of an excess had been seen [65–69]. This new precision measurement of the cosmic ray positron flux, which definitely disagrees with a standard background [70] has opened up a whole new field of speculations about the possible cause of this positron excess. Simultaneously, other data from PAMELA indicate that the antiproton flux is in agreement with standard expectations [71].

There are a variety of astrophysical models proposed for the needed extra primary component of positrons, mainly based on having nearby pulsars as a source [72, 73]. Although pulsars with the required properties like distance, age, and energy output are known to exist, it turns out not to be trivial to fit both FERMI and PAMELA data with these models (see, for example, [74, 75]). For this and other reasons, the dark matter interpretation, which already had been applied to the much more uncertain HEAT data [62] was one of the leading hypotheses.

It was clear from the outset that to fit the PAMELA positron data and FERMI's sum of positrons and electrons with a dark matter model a high mass is needed (on the order of 600 GeV to several TeV). However, since the local average dark matter density is well-known to be around 0.4 GeV/cm^3 [43], the number density decreases as $1/M_X$ and therefore the annihilation rate goes as $1/M_X^2$ with M_X the mass of the annihilating particle. This means that with $\langle \sigma v \rangle = 3 \cdot 10^{-26}$ cm^3/s, which is the standard value of the annihilation rate in the halo for thermally produced WIMPs (see Eq. (69)), the rate of positrons, even for a contrived model which annihilates to e^+e^- with unit branching ratio is much too small to explain the measured result.

To a good approximation, the local electron plus positron flux for such a model is given by, assuming an energy loss of $10^{-16}E^2$ GeVs^{-1} (with E in GeV) from

inverse Compton and synchrotron radiation,

$$E^3 \frac{d\phi}{dE} = 6 \cdot 10^{-4} E \left(\frac{1\,\text{TeV}}{M_X}\right)^2 \theta(M_X - E) B_{\text{tot}} \ \text{m}^{-2}\,\text{s}^{-1}\,\text{sr}^{-1}\,\text{GeV}^2, \quad (157)$$

which means that the boost factor B_{tot} must be of about 200 for a 600 GeV particle, that may otherwise explain the positron excess. Similar boost factors seem to be generic, also for supersymmetric models giving e^+e^- through internal bremsstrahlung [76].

Such a boost factor can in principle be given by a large inhomogeneity in the DM distribution which has to be very local, since positrons and electrons of several hundred GeV do not diffuse very far before losing essentially all their energy. Although not excluded, this would seem to be extremely unlikely in most structure formation scenarios. Therefore, most models rely on the Sommerfeld enhancement factor (see Sect. 6.3). This means a non-negligible amount of fine-tuning of the mass spectrum, in particular also for the degeneracy between the lightest and next-to-lightest particle in the new sector. For a detailed discussion of the required model-building, see [77]. Similar fine-tuning is needed for the decaying dark matter scenario, where the decay rate has to be precisely tuned to give the measured flux. Since the antiproton ratio seems to be normal according to the PAMELA measurements [71], the final states should be mainly leptons (with perhaps intermediate light new particles decaying into leptons). For an interesting such model, which may in fact contain an almost standard axion, see [78].

It seems that at present it is possible to construct models of the Sommerfeld enhanced type [79] which do marginally not contradict present data. However, constraints are getting severe and the dark matter solution to the positron excess is currently not as fashionable as a couple of years ago. It will be interesting, however, to see the first results from the AMS-02 experiment [80] on the International Space Station, which should appear in the summer of 2012.

A very rare process in proton-proton collisions, antideuteron production, may be less rare in neutralino annihilation [81]. However, the fluxes are so small that the possibility of detection seems marginal even in the AMS-02 experiment, and probably a dedicated space probe has to be employed [82].

6.2 Indirect Detection by γ-Rays from the Halo

With the problem of a lack of clear signature of positrons and antiprotons, one would expect that the situation of gamma rays and neutrinos is similar, if they only arise from secondary decays in the annihilation process. For instance, the gamma ray spectrum arising from the fragmentation of fermion and gauge boson final states is quite featureless and gives the bulk of the gamma rays at low energy where the cosmic gamma ray background is severe. However, an advantage is the directional

information that photons carry in contrast to charged particles which random walk through the magnetic fields of the Galaxy [83].

For annihilation into quark-antiquark pairs, or W and Z bosons, the continuous energy spectrum one gets after fragmentation into SM particles can rather well and conveniently be parametrized as

$$dN_{\text{cont}}(E_\gamma)/dE_\gamma = (0.42/m_\chi)e^{-8x}/(x^{1.5} + 0.00014), \qquad (158)$$

where m_χ is the WIMP mass and $x = E_\gamma/m_\chi$. For more detailed spectra, one may for instance use standard particle physics codes like PYTHIA [84] (as is done in [2]). One should note that for τ leptons in the final state (158) is not a good approximation, as this gives a harder spectrum.

Gamma-Ray Lines

An early idea was to look for a spectral feature, a line, in the radiative annihilation process to a charm-anticharm bound state $\chi\chi \to (\bar{c}c)_{\text{bound}} + \gamma$ [85]. However, as the experimental lower bound on the lightest neutralino became higher it was shown that form factor suppression rapidly makes this process unfeasible [86]. The surprising discovery was made that the loop-induced annihilations $\chi\chi \to \gamma\gamma$ [86–88] and $\chi\chi \to Z\gamma$ [89–91] do not suffer from any form factor suppression.

The rates of these processes are difficult to estimate because of uncertainties in the supersymmetric parameters, cross sections and halo density profile. However, in contrast to the other proposed detection methods they have the virtue of giving very distinct, "smoking gun" signals of monoenergetic photons with energy $E_\gamma = m_\chi$ (for $\chi\chi \to \gamma\gamma$) or $E_\gamma = m_\chi(1 - m_Z^2/4m_\chi^2)$ (for $\chi\chi \to Z\gamma$) emanating from annihilations in the halo.

The detection probability of a gamma-ray signal, either continuous or line, will of course depend sensitively on the density profile of the dark matter halo. To illustrate this point, let us consider the characteristic angular dependence of the γ-ray line intensity from neutralino annihilation $\chi\chi \to \gamma\gamma$ in the galactic halo. Annihilation of neutralinos in an isothermal halo with core radius a leads to a γ-ray flux along the line-of-sight (l.o.s.) direction $\hat{n}$ of

$$\frac{d\mathcal{F}}{d\Omega}(\hat{n}) \simeq (0.94 \times 10^{-13} \text{cm}^{-2}\text{s}^{-1}\text{sr}^{-1}) \left(\frac{\sigma_{\gamma\gamma} v}{10^{-29}\text{cm}^{-3}\text{s}^{-1}}\right) \left(\frac{\rho_\chi}{0.3\text{GeV cm}^{-3}}\right)^2$$
$$\times \left(\frac{100\text{GeV}}{m_\chi}\right)^2 \left(\frac{R}{8.5\text{kpc}}\right) J(\hat{n}) \qquad (159)$$

where $\sigma_{\gamma\gamma} v$ is the annihilation rate, ρ_χ is the local neutralino halo density and R is the distance to the galactic center. The integral $J(\hat{n})$ is given by

$$J(\hat{n}) = \frac{1}{R\rho_\chi^2} \int_{\text{l.o.s.}} \rho^2(\ell) d\ell(\hat{n}), \tag{160}$$

and is evidently very sensitive to local density variations along the line-of-sight path of integration. In the case of a smooth halo, its value ranges from a few at high galactic latitudes to several thousand for a small angle average towards the galactic center in the Navarro, Frenck and White (NFW) [105] model [92].

Since the neutralino velocities in the halo are of the order of 10^{-3} of the velocity of light, the annihilation can be considered to be at rest. The resulting gamma ray spectrum is a line at $E_\gamma = m_\chi$ of relative linewidth 10^{-3} (coming from the Doppler effect caused by the motion of the WIMP) which in favourable cases will stand out against background.

Detection of a γ-rate line signal would need a detector with very good energy resolution, like 1 % or better. This may be achieved by FERMI (although only upper limits have been presented so far [93]). However, the Russian satellite GAMMA-400 [94] seems to have very promising characteristics for this type of dark matter search, when it is launched by the end of this decade. This could be a very interesting new instrument in the search for γ-ray lines from annihilation (or decay) of dark matter.

The calculation of the $\chi\chi \to \gamma\gamma$ cross section is technically quite involved with a large number of loop diagrams contributing. A full calculation in the MSSM was performed in [95, 96]. Since the different contributions all have to be added coherently, there may be cancellations or enhancements, depending on the supersymmetric parameters. The process $\chi\chi \to Z\gamma$ is treated analogously and has a similar rate [89–91].

An important contribution, especially for neutralinos that contain a fair fraction of a higgsino component, is from virtual W^+W^- intermediate states. This is true both for the $\gamma\gamma$ and $Z\gamma$ final state for very massive neutralinos [89–91]. In fact, thanks to the effects of coannihilations [11], neutralinos as heavy as several TeV are allowed without giving a too large Ω. These extremely heavy dark matter candidates (which, however, would require quite a degree of finetuning in most supersymmetric models) are predominantly higgsinos and have a remarkably large branching ratio into the loop-induced $\gamma\gamma$ and $Z\gamma$ final states (the sum of these can be as large as 30 %). If there would exist such heavy, stable neutralinos, the gamma ray line annihilation process may be the only one which could reveal their existence in the foreseeable future (since not even LHC would be sensitive to supersymmetry if the lightest supersymmetric particle weighs several TeV). In fact the high branching ratio for higgsino annihilation to 2γ was the reason that Hisano et al. [97] took a closer look at the process and discovered the effect of Sommerfeld enhancement.

Internal Bremsstrahlung

The $\gamma\gamma$ process appears in a closed loop meaning that it is suppressed by powers of the electromagnetic coupling constant. An amusing effect appears, however, for Majorana fermions at even lower order. It was early realized that there could be important spectral features [98], and recently it has been shown that internal bremsstrahlung

(IB) from produced charged particles in the annihilations could yield a detectable "bump" near the highest energy for heavy gauginos or Higgsinos annihilating into W boson pairs, such as expected in split supersymmetry models [99]. In [100], it was furthermore pointed out that IB often can be estimated by simple, universal formulas and often gives rise to a very prominent step in the spectrum at photon energies of $E_\gamma = m_\chi$ (such as in lightest Kaluza-Klein particle (LKP) models [101] (see Sect. 7.8)). The IB process was thoroughly treated in [102], and here we summarize the main results.

In [98] it was shown that the radiative process $\chi^0\chi^0 \to f\bar{f}\gamma$ may circumvent the chiral suppression, i.e., the annihilation rate being proportional to m_f^2. This is normally what one would get for annihilation into a fermion pair from an S-wave initial state [103], as is the case in lowest order for non-relativistic dark matter Majorana particles in the Galactic halo (see also [104]). Since this enhancement mechanism is most prominent in cases where the neutralino is close to degenerate with charged sleptons, it is of special importance in the so-called stau coannihilation region.

A fermion final state containing an additional photon, $f\bar{f}\gamma$, is thus not subject to a helicity suppression. The full analytical expressions are lengthy, but simplify in the limit of $m_f \to 0$. Then one finds

$$\frac{\mathrm{d}N_{f^+f^-}}{\mathrm{d}x} = \lambda \times \left\{ \frac{4x}{\mu(\mu - 2x)} - \frac{2x}{(\mu - x)^2} - \frac{\mu(\mu - 2x)}{(\mu - x)^3} \log \frac{\mu}{\mu - 2x} \right\}, \quad (161)$$

with

$$\lambda = (1 - x)\alpha_{\mathrm{em}} Q_f^2 \frac{|\tilde{g}_R|^4 + |\tilde{g}_L|^4}{64\pi^2} \left(m_\chi^2 \langle\sigma v\rangle_{\chi\chi\to f\bar{f}} \right)^{-1}.$$

where $\mu \equiv m_{\tilde{f}_R}^2/m_\chi^2 + 1 = m_{\tilde{f}_L}^2/m_\chi^2 + 1$ and $\tilde{g}_R P_L$ ($\tilde{g}_L P_R$) is the coupling between neutralino, fermion and right-handed (left-handed) sfermion. This confirmed the result found in [98] for photino annihilation. Note the large enhancement factor m_χ^2/m_f^2 due to the lifted helicity suppression (from $\langle\sigma v\rangle_{\chi\chi\to f\bar{f}} \propto m_f^2 m_\chi^{-4}$), and another large enhancement that appears at high photon energies for sfermions degenerate with the neutralino.

Internal bremsstrahlung from the various possible final states of neutralino annihilations is included in DarkSUSY [2]. The total γ-ray spectrum is given by

$$\frac{dN^{\gamma,\mathrm{tot}}}{dx} = \sum_f B_f \left(\frac{dN_f^{\gamma,\mathrm{sec}}}{dx} + \frac{dN_f^{\gamma,\mathrm{IB}}}{dx} + \frac{dN_f^{\gamma,\mathrm{line}}}{dx} \right), \quad (162)$$

where B_f denotes the branching ratio into the annihilation channel f. The last term in the above equation gives the contribution from the direct annihilation into photons, $\gamma\gamma$ or $Z\gamma$, which result in a sharp line feature [89–91, 95, 96]. The first term is the contribution from secondary photons from the fragmentation of the fermion pair.
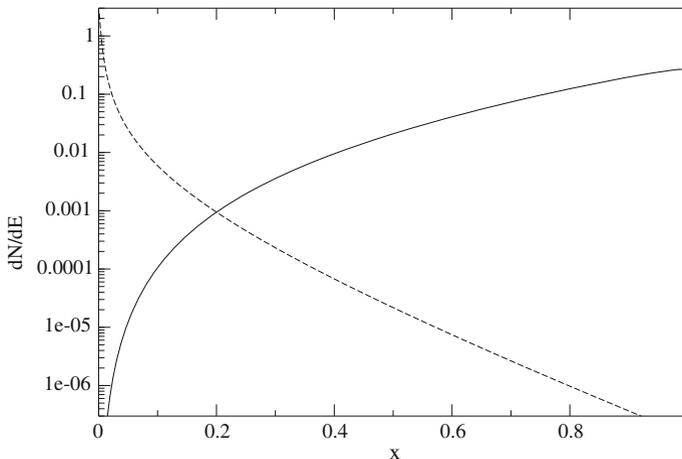
Fig. 6 The distribution of γ-rays from the internal bremsstrahlung process $\chi^0\chi^0 \to f\bar{f}\gamma$ is shown as the *solid line*, and compared to the standard case, (158) (*dashed line*). As can be seen, the internal bremsstrahlung process gives a very hard spectrum, which may counteract the fact that radiation of a photon always is suppressed by a factor of $\sim\alpha_{em}/\pi$

This "standard" part of the total γ-ray yield from dark matter annihilations shows a feature-less spectrum with a rather soft cutoff at $E_\gamma = m_\chi$.

In Fig. 6 an example of the energy distribution of photons given by (161) is shown.

Density Profile and γ-Ray Detection

To compute $J(\hat{n})$ in (160), a model of the dark matter halo has to be chosen. The universal halo profile found in simulations by Navarro, Frenk and White (NFW) [105] has a rather significant enhancement $\propto 1/r$ near the halo centre,

$$\rho_{\text{NFW}} = \frac{c}{r(a+r)}, \tag{163}$$

where c is a concentration parameter and a a typical length scale for the halo. In fact, more detailed later simulations have given a slightly different shape, the so-called Einasto profile,

$$\rho_{\text{Einasto}} = \rho_s e^{-\frac{2}{\alpha}\left[\left(\frac{r}{\alpha}\right)^\alpha - 1\right]}, \tag{164}$$

with $\alpha \sim 0.17$ for the Milky Way. Except near $r = 0$, this profile is actually quite similar to the NFW profile, and it has slightly higher density outside the very center. The local dark matter density near the solar system can be quite well determined [43] and is $\rho_0 \simeq 0.4$ GeV/cm^3. If these forms of the density can be applied to the Milky Way, this would lead to a much enhanced annihilation rate towards the galactic centre,
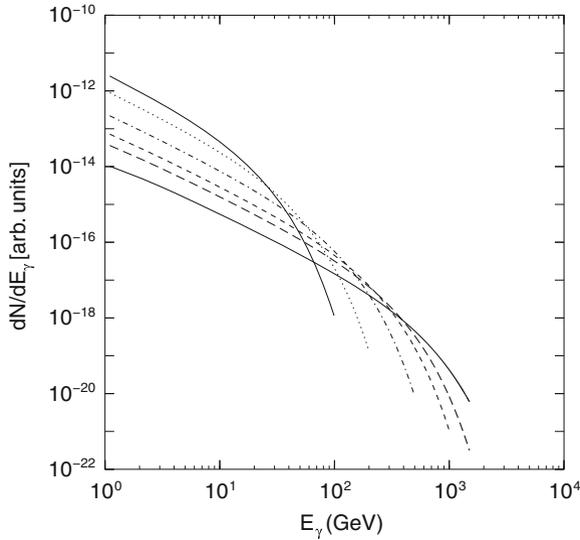
Fig. 7 The energy distribution of γ-rays from WIMP dark matter annihilation into a $b\bar{b}$ pair, for a dark matter particle mass of 100, 200, 500, 1000, 1500, and 3000 GeV, respectively. One can see that the bulk of the signal is at low energies. (Here the line signals from $\gamma\gamma$ and $Z\gamma$ have not been included)

and also to a very characteristic angular dependence of the line signal. This would be very beneficial when discriminating against the extragalactic γ ray background, and Imaging Air Cherenkov Telescope Arrays (IACTAs) are be eminently suited to look for these signals since they have an angular acceptance which is well matched to the angular size of the Galactic central region where a cusp is likely to be. Both H.E.S.S. [106], MAGIC [107] and Whipple [108] have searched for a signal at the galactic center or in other dark matter concentrations, but are still a couple of orders of magnitude above the flux which would correspond to the canonical WIMP flux, Eq. (69). Maybe with the planned CTA project [109] one may get to the interesting region of parameter space for supersymmetric or other WIMPs.

Also the energy threshold of present-day IACTAs is too high (of the order of 50 GeV or higher) to be very useful for WIMPs of 100 GeV or lighter. There have been discussions about a high-altitude detector with lower threshold, perhaps as low as 5 GeV [110], which would be very beneficial for dark matter detection, see Fig. 7.

Space-borne gamma-ray detectors, like the FERMI satellite have a much smaller area (on the order of 1 m^2 instead of $10^4 - 10^5$ m^2 for IACTAs), but a correspondingly larger angular acceptance so that the integrated sensitivity is in fact similar. This is at least true if the Galactic center does not have a very large dark matter density enhancement which would favour IACTAs. The total rate expected in FERMI can be computed with much less uncertainty because of the angular integration [111]. Directional information is obtained and can be used to discriminate against the diffuse

extragalactic background. A line signal can be searched for with high precision, since the energy resolution of FERMI is at the few percent level.

Indirect Detection Through Neutrinos

The density of neutralinos in the halo is not large enough to give a measurable flux of secondary neutrinos, unless the dark matter halo is very clumpy [112–114]. In particular, the central Galactic black hole may have interacted with the dissipationless dark matter of the halo so that a spike of very high dark matter density may exist right at the Galactic centre [115]. However, the existence of these different forms of density enhancements are very uncertain and depend extremely sensitively on presently completely unknown aspects of the formation history of the Milky Way.

More model-independent predictions (where essentially only the relatively well-determined local halo dark matter density is of importance) can be made for neutrinos from the centre of the Sun or Earth, where neutralinos may have been gravitationally trapped and therefore their density enhanced. As they annihilate, many of the possible final states (in particular, $\tau^+\tau^-$ lepton pairs, heavy quark–antiquark pairs and, if kinematically allowed, $W^\pm H^\mp$, $Z^0 H_i^0$, W^+W^- or Z^0Z^0 pairs) give after decays and perhaps hadronization energetic neutrinos which will propagate out from the interior of the Sun or Earth. (For neutrinos from the Sun, energy loss of the hadrons in the solar medium and the energy loss of neutrinos have to be considered [116, 117]). In particular, the muon neutrinos are useful for indirect detection of neutralino annihilation processes, since muons have a quite long range in a suitable detector medium like ice or water. Therefore they can be detected through their Cherenkov radiation after having been produced at or near the detector, through the action of a charged current weak interaction $\nu_\mu + A \to \mu + X$.

Detection of neutralino annihilation into neutrinos is one of the most promising indirect detection methods, and will be subject to extensive experimental investigations in view of the new neutrino telescopes (IceCube, ANTARES, KM3NET) planned or under construction [118]. The advantage shared with gamma rays is that neutrinos keep their original direction. A high-energy neutrino signal in the direction of the centre of the Sun or Earth is therefore an excellent experimental signature which may stand up against the background of neutrinos generated by cosmic-ray interactions in the Earth's atmosphere.

The differential neutrino flux from neutralino annihilation is

$$\frac{dN_\nu}{dE_\nu} = \frac{\Gamma_A}{4\pi D^2} \sum_f B_\chi^f \frac{dN_\nu^f}{dE_\nu} \tag{165}$$

where Γ_A is the annihilation rate, D is the distance of the detector from the source (the central region of the Earth or the Sun), f is the neutralino pair annihilation final states, and B_χ^f are the branching ratios into the final state f. dN_ν^f/dE_ν are the energy distributions of neutrinos generated by the final state f. Detailed calculations

of these spectra can be made using Monte Carlo methods [117, 119–121]. Effects of neutrino oscillations have also been included [122].

The neutrino-induced muon flux may be detected in a neutrino telescope by measuring the muons that come from the direction of the centre of the Sun or Earth. For a shallow detector, this usually has to be done in the case of the Sun by looking (as always the case for the Earth) at upward-going muons, since there is a huge background of downward-going muons created by cosmic-ray interactions in the atmosphere. The flux of muons at the detector is given by

$$
\frac{dN_\mu}{dE_\mu} = N_A \int_{E_\mu^{\mathrm{th}}}^{\infty} dE_\nu \int_0^{\infty} d\lambda \int_{E_\mu}^{E_\nu} dE'_\mu \; P(E_\mu, E'_\mu; \lambda) \; \frac{d\sigma_\nu(E_\nu, E'_\mu)}{dE'_\mu} \; \frac{dN_\nu}{dE_\nu},
$$
(166)

where λ is the muon range in the medium (ice or water for the large detectors in the ocean or at the South Pole, or rock which surrounds the smaller underground detectors), $d\sigma_\nu(E_\nu, E'_\mu)/dE'_\mu$ is the weak interaction cross section for production of a muon of energy E'_μ from a parent neutrino of energy E_ν, and $P(E_\mu, E'_\mu; \lambda)$ is the probability for a muon of initial energy E'_μ to have a final energy E_μ after passing a path-length λ inside the detector medium. E_μ^{th} is the detector threshold energy, which for "small" neutrino telescopes like Baksan, MACRO and Super-Kamiokande is around 1 GeV. Large area neutrino telescopes in the ocean or in Antarctic ice typically have thresholds of the order of tens of GeV, which makes them sensitive mainly to heavy neutralinos (above 100 GeV) [123]. Convenient approximation formulas relating the observable muon flux to the neutrino flux at a given energy exist [124].

The integrand in (166) is weighted towards high neutrino energies, both because the cross section σ_ν rises approximately linearly with energy and because the average muon energy, and therefore the range λ, also grow approximately linearly with E_ν. Therefore, final states which give a hard neutrino spectrum (such as heavy quarks, τ leptons and W or Z bosons) are usually more important than the soft spectrum arising from light quarks and gluons.

The rate of change of the number of neutralinos N_χ in the Sun or Earth is governed by the equation

$$
\dot{N}_\chi = C_C - C_A N_\chi^2
$$
(167)

where C_C is the capture rate and C_A is related to the annihilation rate Γ_A, $\Gamma_A = C_A N_\chi^2$. This has the solution

$$
\Gamma_A = \frac{C_C}{2} \tanh^2\left(\frac{t}{\tau}\right),
$$
(168)

where the equilibration time scale $\tau = 1/\sqrt{C_C C_A}$. In most cases for the Sun, and in the cases of observable fluxes for the Earth, τ is much smaller than a few billion years, and therefore equilibrium is often a good approximation ($\dot{N}_\chi = 0$ in (167)).

This means that it is the capture rate which is the important quantity that determines the neutrino flux.

The capture rate induced by scalar (spin-independent) interactions between the neutralinos and the nuclei in the interior of the Earth or Sun is the most difficult one to compute, since it depends sensitively on Higgs mass, form factors, and other poorly known quantities. However, this spin-independent capture rate calculation is the same as for direct detection. Therefore, there is a strong correlation between the neutrino flux expected from the Earth (which is mainly composed of spin-less nuclei) and the signal predicted in direct detection experiments [123, 125]. It seems that even the large (kilometer-scale) neutrino telescopes planned will not be competitive with the next generation of direct detection experiments when it comes to detecting neutralino dark matter, searching for annihilations from the Earth. However, the situation concerning the Sun is more favourable. Due to the low counting rates for the spin-dependent interactions in terrestrial detectors, high-energy neutrinos from the Sun constitute a competitive and complementary neutralino dark matter search. Of course, even if a neutralino is found through direct detection, it will be extremely important to confirm its identity and investigate its properties through indirect detection. In particular, the mass can be determined with reasonable accuracy by looking at the angular distribution of the detected muons [126, 127].

For the Sun, dominated by hydrogen, the axial (spin-dependent) cross section is important and relatively easy to compute. A good approximation is given by [32]

$$\frac{C_\odot^{\text{sd}}}{(1.3 \cdot 10^{23}\,\text{s}^{-1})} = \left(\frac{\rho_\chi}{0.3\,\text{GeV}\,\text{cm}^{-3}}\right)\left(\frac{100\,\text{GeV}}{m_\chi}\right)\left(\frac{\sigma_{p\chi}^{\text{sd}}}{10^{-40}\,\text{cm}^2}\right)\left(\frac{270\,\text{km/s}}{\bar{v}}\right),$$
(169)

where $\sigma_{p\chi}^{\text{sd}}$ is the cross section for neutralino-proton elastic scattering via the axial-vector interaction, $\bar{v}$ is the dark-matter velocity dispersion, and ρ_χ is the local dark matter density. The capture rate in the Earth is dominated by scalar interactions, where there may be kinematic and other enhancements, in particular if the mass of the neutralino almost matches one of the heavy elements in the Earth. For this case, a more detailed analysis is called for, but convenient approximations are available [32]. In fact, also for the Sun the spin-dependent contribution can be important, in particular iron may contribute non-negligibly.

A neutrino telescope of area around 1 km^2, which is roughly the size of IceCube, has discovery potential for a range of supersymmetric models, which cannot easily be probed using other methods, see [123].

6.3 Antimatter Detection of Dark Matter

Antimatter does not seem to be present in large quantities in the universe, as can be inferred from the absence of γ-ray radiation that would have been created in large amounts if astrophysical anti-objects would annihilate on their matter counterparts

(this would also cause deviations from the pure black-body form of the cosmic microwave background, something which is very severely limited by WMAP data and will be further probed by the Planck satellite). In fact, both the analysis of primordial nucleosynthesis and the CMB, give a non-zero number around 10^{-10} for the baryon-antibaryon asymmetry, which means that matter dominated over antimatter already in the very early universe. On the other hand, dark matter annihilation in almost all models occurs from a matter-antimatter symmetric initial state and thus equal amounts of matter and antimatter is created. This leads to an interesting possible new primary source of positrons and antiprotons (i.e. the stable anti-particles of protons) in the cosmic rays of dark matter halos, including the one where the Milky Way resides. (There is always a small amount of antimatter produced as secondary particles in collisions with galactic gas and dust by ordinary cosmic rays, of course.) As discussed extensively at conferences in 2009 (see, e.g., [128]) this was an extremely hot topic then. This was due to the PAMELA and FERMI collaborations just having discovered an anomalously high ratio of positrons over electrons up to 100 GeV [63], and sum of positrons and electrons up to 1 TeV [64], respectively. During the last two years, this anomaly, although possible to explain by dark matter annihilation, needs such large boost factors (e.g., from Sommerfeld enhancement to be discussed below), and somewhat contrived, leptophilic models, that these models are feeling severe pressure from other detection methods, e.g, γ-rays from the central parts of the Galaxy [129]. Alternative astrophysical explanations are on the other hand possible with quite standard assumptions. One cannot say that the dark matter explanation is yet completely ruled out, but it is in strong tension from other measurements.

Returning to more standard WIMP models, there have recently been improvements in the computations of the annihilation rate at low velocity as is the case in galaxies, where $v/c \sim 10^{-3}$. An amusing effect is caused due to the suppression of the $^3S_1$ for an initial state of two Majorana spinors (such as neutralinos) at zero velocity, due to the requirement of Fermi statistics. Namely, one cannot have two identical fermions in the same spin state. This means that annihilation only occurs from the pseudoscalar $^1S_0$ state where one of the particles has spin up, the other spin down. This causes for instance the annihilation amplitude into a light fermion-antifermion pair, like e^+e^-, to be suppressed by an explicit helicity factor of the fermion mass (as in the limit of zero mass, the vertices are helicity-preserving, and to cause a spin flip a mass term is needed). Direct annihilation into e^+e^- was thus thought to be very subdominant. However, it was realized [130, 131] (building on an old idea [132]), that a spin-flip by one of the Majorana fermions caused by emitting a photon could first of all relieve the helicity suppression of the process to a mere α/π ordinary radiative factor. And, in addition, the spectral shape of the emitted photon is very favorable for detection, causing a shoulder which peaks close to the dark matter particle mass. In particular, for heavy (TeV-scale) WIMPs this could be quite important, and using the radiative peak would help extracting the signal over background [133]. Recently, these radiative processes have been generalized also to emission of other gauge bosons, and have been shown to be quite important generally [134, 135].

The Sommerfeld Effect

The possibility of an enhanced annihilation rate due to DM halo substructure has been realized for a long time [112–114]. However, it seems hard to produce a boost factor of the order of a few hundred to a thousand in the solar neighborhood, as would be needed to explain the PAMELA and FERMI excesses. This is because substructure survives in numerical simulations mostly in the outer portions of the halo, due to tidal stripping in the inner part.

Another potentially very important effect, Sommerfeld enhancement, which may explain the large boost had been found a few years earlier. This effect, was computed for electromagnetism by Arnold Sommerfeld many years ago [136], but it was rediscovered [77, 97] in the quantum field theory of very heavy dark matter particles in the limit when the gauge particles, γ, Z^0 and $W^\pm$ are essentially massless, or at least have a Compton wavelength that is sufficiently large compared to the would-be bound state caused by the attractive gauge forces. (Of course, a bound state is never really formed due to the fast time scale of annihilation.)

In the quantum mechanical calculation of electron scattering and e^+e^- annihilation, Sommerfeld enhancement is caused by the distortion of the plane wave describing the relative motion of the annihilating particle pair through the near formation of a bound state caused by photon exchange. In the so-called ladder approximation for QED (where one sums only certain types of Feynman diagrams), one obtains this Sommerfeld effect, and the square of the wave function at the origin in relative coordinates $r_1 - r_2$, which enters into the probability for the short-distance process of annihilation, is increased by the factor [77]

$$S = \frac{|\Psi(0)|^2}{|\Psi_{(0)}(0)|^2} = \frac{\left(\frac{\pi\alpha}{\beta}\right)}{1 - e^{-\left(\frac{\pi\alpha}{\beta}\right)}}, \tag{170}$$

with α the fine-structure constant, and β the relative velocity. This can be expanded to $S_{QED} = \pi\alpha/\beta$ for small relative velocities. In the Milky Way halo, velocities are typically $\beta \sim 10^{-3}$, so this limit is certainly relevant. For smaller galaxies or DM substructure, velocities (as measured by velocity dispersions) are even smaller. Of course, there is no direct photon exchange between DM particles, since they are electrically neutral. However, if there are charged states nearby in mass, the neutral pair may momentarily, before annihilation, transform into a charged pair which in turn may exchange a photon between then. These are the basic processes that have to be summed to all orders in the ladder approximation, and which lead to Sommerfeld enhancement (see Fig. 8).

One could of course also have a Yukawa-like particle (i.e., spinless) of mass m_Y, mediating a weak attractive force with coupling constant α_Y between DM particles of mass m_χ. The small velocity limit of the enhancement then becomes

$$S_Y \propto \frac{\alpha_Y m_\chi}{m_Y}. \tag{171}$$
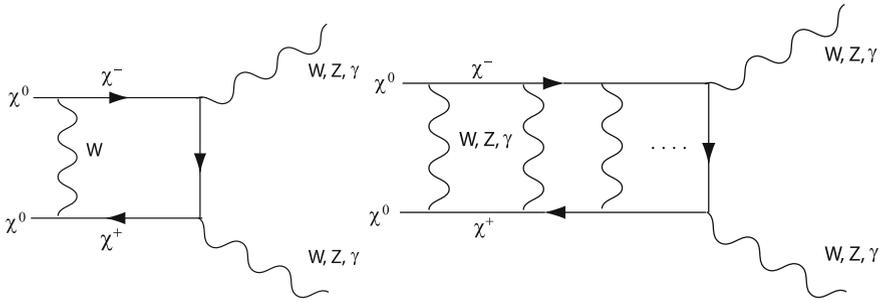
Fig. 8 Diagrams illustrating the field-theoretical reason for the Sommerfeld enhancement. The figure is drawn for a supersymmetric neutralino (which is the case where the effect was first found in dark matter physics [97]), but similar diagrams apply for any dark matter candidate which first of all is heavy compared to the exchanged particle in the t-channel (i.e. in the "ladder"), and where there is a near degeneracy between the neutral state being the dark matter and the virtual states (in this case charged particles, charginos). On the *left* is shown the lowest order contribution, which gets very important for large masses, and which is further enhanced by the ladder diagrams of the type shown on the *right*. The net result could be an "explosive annihilation", to quote [97]

In some cases, depending on the detailed nature of the mediating particles, the enhancement factor S can indeed be as high as several hundred to a few thousand, depending on the exact parameters. The effect is generally strongly velocity-dependent, depending on velocity as $1/\beta$ or even (near resonance) $1/\beta^2$ but in the Yukawa case the $1/\beta$ scaling is valid only for $\beta > m_Y/m_\chi$. At smaller velocities and outside resonances, the effect saturates at m_Y/m_χ [137].

Important bounds come from γ-rays, but also from the non-observation of energy distortions in the cosmic microwave background. It may still be possible to (marginally) fit the PAMELA/FERMI excess, if one takes astrophysical uncertainties into account [79].

It should be noted that the Sommerfeld effect has a solid theoretical backing and is important, if the mass and coupling parameters are in the right range. For supersymmetric models, however, it occurs only for very heavy neutralinos (generally higgsinos) and the phenomenology has only been partly investigated [138].

To conclude this section on detection methods of WIMPs, we have seen that supersymmetric particles, which are the theoretically most plausible WIMPs have many interesting features which may make them detectable in the not too distant future. Supersymmetry, in particular MSSM, invented already in the 1970s, and obtained as a phenomenological manifestation of the most realistic string theories, has since the early 1980s, when the CDM paradigm first won universal acclaim, been the prime template for a WIMP [103, 139].

Even in the MSSM, however, there are in principle more than a hundred free parameters, meaning that for practical reasons the templates, for instance used at the LHC experiments, are drastically simplified versions, like constrained MSSM (CMSSM) or the even more constrained minimal super gravity (mSUGRA), which do not, in contrast to the full MSSM, correspond very well to more recent thinking

about supersymmetry breaking [140]. This has to be kept in mind when discussing the impressive LHC limits. Even in still simplified versions, like the 19 to 24-parameter "phenomenological MSSM", pMSSM [141, 142], the bounds on particle masses given, e.g., by fulfilling the WMAP relic density, are not very constraining at the moment [143]. Of course, the outlook for the MSSM would be much bleaker if a light Higgs (with mass below roughly 130 GeV) were not to be established by the end of the 7 TeV run, in 2012.

With the freely available [144] DarkSUSY package [2], one can compute in detail the relic density, not only for supersymmetric models, but since the package has a modular design, one can insert any favourite model one has for WIMP-like dark matter. Of course, DarkSUSY is mostly used for the supersymmetric case, and it has been originally set up for a general pMSSM model, with large freedom in the choice of parameters.

7 Particular Dark Matter Candidates

7.1 WIMP Models

The particle physics connection is particularly striking in the WIMP scenario, namely that for typical gauge couplings and a mass at the weak interaction scale of a few hundred GeV, consistent with the relic density computed using standard big bang thermodynamics, as we saw in Sect. 3. This is rather well tested by the calculation of the abundances of hydrogen and helium in the early universe, through big bang nucleosynthesis. The calculation of these abundances turns out to be in amazingly good agreement with the measured ones. Using the same early universe thermodynamics and solving the Boltzmann equation for hypothetical dark matter particles of mass m_χ, we found that the annihilation rate $\langle \sigma v \rangle$ needed to explain $\Omega_\chi h^2 \sim 0.11$ (as determined by WMAP), naturally appears for ordinary gauge couplings and a mass between around 20 GeV to a few TeV—a WIMP.

Although this is not a completely convincing argument for WIMP dark matter—it may perhaps be a coincidence—it nevertheless gives WIMP candidates a flavour of naturalness. For non-WIMP candidates there is, on the other hand, usually a fine tuning involved, or use of non-standard cosmology, to obtain the correct relic density. Even limiting oneself to WIMP models for dark matter, the literature is extensive, and among some recent developments, which cannot be discussed in this review in any detail, can be mentioned:

7.2 Dark Stars

Since cosmological structure in WIMP models occurs hierarchically, starting from scales as small as 10^{-12}–$10^{-6} M_\odot$ [145], the idea has been put forward that the earliest dense, small structures created by dark matter may play a role in star formation and if the dark matter particles annihilate within the stars, unusual stellar evolution may result [146–149].

7.3 Inelastic Dark Matter

These are dark matter candidates which may be excited to a state with slightly higher mass and therefore cause a higher than usual direct detection rate [150–155], and also relieve the tension between the different direct detection experiments.

7.4 Dynamical Dark Matter

As it is not obvious that there is only one type of particle making up the dark matter (neutrinos should, for example contribute up to a few percent), an extreme solution could be to have a very large number, with different spins, masses, etc. [156].

7.5 Leptophilic Dark Matter

As we have mentioned, there was an almost explosion of suggestions of this kind of models in 2009, when the dark matter interpretation of the anomalous positron ratio measured by PAMELA [63] and FERMI [64] was proposed to be explained by dark matter annihilation. Leptophilic means that these dark matter particles annihilate mainly to leptons, for example by proceeding through axion-like particles below the pion mass [157–163]. Although the original motivation for these models has become somewhat weaker, the concept has established itself in the dark matter community.

7.6 Supersymmetric Models Beyond the MSSM

Of course, even though the minimal supersymmetric version of the standard model, the MSSM, has more than 100 free parameters, models having, e.g., motivation from new scenarios of supersymmetry breaking, are of course logically possible. These "beyond the MSSM" or BMSSM models [164–167] may among other things give a higher Higgs mass than the limit of 130 GeV given by minimal SUSY models. In the summer of 2011, this was perhaps a favoured scenario, as the first indications

of the Higgs mass was around 140 GeV. However, with more data, the preferred range (not yet significant enough by the end of 2011 to be called a discovery) is now 124–126 GeV which is more easily encompassed in the MSSM.

7.7 Asymmetric Dark Matter

This is a class of dark matter models which may also explain the baryon (or lepton) asymmetry of the universe [168–173]. This generally only works for masses around or below 10 GeV, and this mass range has been in focus recently due to a (possible) signal in direct detection experiments [58, 174–176], and maybe also in γ-ray detection in the direction near the Galactic centre [177]. However, it remains to see whether these indications will stand the test of time. A similar model is "emergent dark matter". This is a recent version of asymmetric DM with larger possible parameter range, such as a DM mass up to 100 GeV [168–173].

7.8 Kaluza-Klein Models

A candidate for dark matter, the so-called LKP (for lightest Kaluza-Klein particle) has been identified. This appears in theories with extra dimensions, and has a rich phenomenology which we will not enter into here (for a review, see [178]). The main difference with supersymmetry is that the dark matter candidate has spin-1, and can give the correct relic density for a mass in the range from 600 GeV to 1 TeV.

7.9 Inert Higgs Doublet

Interesting are also versions of the Standard Model with an enlarged Higgs sector. If there would be, for instance, a second Higgs doublet which does not couple directly to Standard Model particles (an "inert doublet"), there turns out to be a stable spin-0 state which then would be the dark matter particle (see [179], and references therein).

7.10 Non-WIMP Models

WIMPs are arguably the leading candidates for Dark Matter, due to lack of fine-tuning to get correct relic density. In most models, the annihilation cross section which sets the relic density also implies observable rates in various DM detection experiments.

A word of caution is in place here, however. There are many non-WIMP models that also have good particle physics motivation, and may be detectable, like: axions,

gravitinos, superWIMPS, non-thermal dark matter, decaying dark matter, sterile Neutrinos, Q-balls. . . The literature is extensive, but a good summary of both WIMP and non-WIMP models has recently appeared, namely, a 700-page book giving details of most dark matter scenarios [180].

7.11 The Axion

Another, rather different candidate [181] for dark matter is provided by the axion, a hypothetical light boson which was introduced for theoretical reasons to explain the absence of CP violation in the strong interactions (as far as we know, CP violations only take place in the weak interactions). It turns out that for a mass range between 10^{-6} and 10^{-3} eV, the axion could give a sizable contribution to Ω_M. It couples very weakly to ordinary matter, but it may be converted into a photon in a cavity containing a strong magnetic field (the basic coupling is to two photons, but here the magnetic field takes the role of one photon). Experiments in the USA and Japan are currently probing parts of the interesting mass region. A section about the axion should always be inserted when describing dark matter candidates, since the axion has, as has the lightest supersymmetric particle, a good particle physics motivation for its existence.

8 Dark Matter Detection: Status

As we have mentioned, there are basically three different, and complementary methods for detecting WIMPs. First, the dark matter particle may be directly produced at accelerators, in particular at the LHC, which today is the only high-energy accelerator running (although data from Fermilab's Tevatron collider will still be analyzed and may give surprises in the coming year or so). Of course, it is not clear that the particle will be kinematically allowed, and even if it is produced, one will not know that the lifetime is of the required cosmological order of magnitude. Anyway, detecting a candidate and determining its mass would be a great gain when combining with the other two search methods of dark matter, namely direct and indirect detection. In particular, direct detection experiments have seen an impressive gain of sensitivity during the last few years. The idea is to register rare events giving a combination of scintillation, ionization and nuclear recoil signals in chunks of matter shielded from cosmic rays in underground sites.

In indirect detection, one rather registers products of dark matter annihilation from regions in the surrounding universe with a high dark matter density like the galactic centre, dwarf spheroidal galaxies, or the interior of the Earth or the Sun. An interesting feature of indirect detection is that the expression for the local annihilation rate of a pair of DM particles χ (here assumed, like in supersymmetry, to be self-charge-conjugate, of relative velocity v_{rel}

$$\Gamma_{ann} \propto n_\chi^2 \sigma_{ann}(v_{rel}) v_{rel} \tag{172}$$

is the dependence on the square of the number density. Also, the cross section may depend in non-trivial ways on the relative velocity. In particular, for low velocities the rate may be much higher than at high velocity, for models containing an attractive force between the annihilating particles. This is in particular true for models with so-called Sommerfeld enhancement [97], a resonant enhancement by in some cases orders of magnitude (see Sect. 6.3). This means that dwarf galaxies (dark matter subhalos) may be particularly interesting objects to study, as they are completely dark matter dominated with low rate of cosmic-ray induced γ-rays, and their low mass means a relatively low velocity dispersion, meaning higher possible rates if Sommerfeld enhancement is active.

So far, indirect methods have not been as competitive as direct detection, but recently the FERMI collaboration has started to probe the interesting WIMP region by stacking data from several dwarf galaxies [182].

For non-WIMP dark matter, like sterile neutrinos (warm DM), the production rate in the early universe generally has to be tuned to give the observed relic density, but phenomenologically warm DM is possible, and according to some analyses even preferred in cosmological data [183]. However, the significance is weak and may be influenced by statistical bias [184]. Ordinary, active neutrinos have too small mass to contribute significantly to the dark matter density, although in the extreme case may contribute a couple of percent to the critical density today.

A very interesting effect for direct detection of dark matter WIMPs in terrestrial detectors comes about due to the motion of the solar system in the Galaxy [55]. This circular speed is around 200 km/s, and the direction of the "wind" of dark matter particles varies in between seasons. This is due to the detector following the Earth's motion around the Sun and sometimes (actually around June 2) having "headwind" of WIMPs and sometimes (December 2) "tailwind". As the cross section between a WIMP and the detector target depends strongly on their relative velocity, this causes a few percent annual modulation of the detection rate, something that is a very distinct signature. The DAMA/LIBRA experiment in the Gran Sasso tunnel [58] has in fact seen an annual modulation, which has a statistical significance of more than 8 standard deviations. However, since no other experiment has found the same effect (see Table 1), the effect can still not be taken as an established detection of dark matter. There have been attempts to interpret the DAMA signal as possibly being due to a neutralino of the MSSM [185, 186]. It seems premature, however, to draw strong conclusions from this experiment alone. Besides some cloudy experimental issues, the implied scattering rate seems somewhat too high for the MSSM or any other canonical WIMP, given the strong Higgs mass bounds from LEP and LHC unless one stretches the astrophysical and nuclear physics quantities. Also, it is disturbing that neither XENON100 nor CDMS-II see an effect despite their nominally higher sensitivity. Clearly, even more sensitive experiments, preferably also using NaI, seem to be needed to settle this issue. An interesting idea, DM-Ice [187], uses the IceCube site to deploy crystals of NaI with ice as a very calm surrounding medium. If an annual modulation could be measured also there one could check whether it has the

Table 1 Some of the recent experimental claims for possible dark matter detection, and a comment on the present status

Experiment	Status of claim
DAMA/LIBRA annual modulation [58]	Unexplained at the moment; not confirmed by other experiments [54, 188]
CoGeNT excess events and annual modulation [174, 175]	Tension with other data [54, 188]
EGRET excess of GeV photons [189, 190]	Due to instrument error (?)—not confirmed by FERMI [191]
INTEGRAL 511 keV γ-line from galactic centre region [192]	Does not seem to have spherical symmetry—shows an asymmetry which follows the disk (?) [193]
PAMELA: Anomalous ratio of cosmic ray positrons/electrons [63]	May be due to DM [194, 195], or pulsars [72, 73]—energy signature not unique for DM
FERMI positrons + electrons excess [64]	May be due to DM [194, 195], or pulsars [72, 73]—energy signature not unique for DM
FERMI γ-ray excess towards galactic centre [196, 197]	Unexplained at the moment—astrophysical explanations possible [198, 199], no statement from the FERMI collaboration
WMAP radio "haze" [200]	Has a correspondence in "FERMI bubbles" [201]—probably caused by outflow from the galactic center

same phase as that of DAMA, or if it rather follows the seasons (which are opposite on the southern hemisphere).

There have recently been a number of claimed possible detections of dark matter, see Table 1. Of the items in Table 1, it seems that only the positron excess at high energy (20 GeV–1 TeV) and the γ-ray excess towards the galactic center, inferred by an analysis of FERMI public data [196, 197], can be due to dark matter annihilation without tension from other data. However, they may both perhaps more naturally be explained by ordinary astrophysical processes. In addition, the DM explanation of the PAMELA and FERMI data as we have seen needs a leptophilic particle of TeV-scale mass and a very much boosted cross section. Although this may perhaps be obtained, stretching all uncertainties involved [202], and employing Sommerfeld enhancement [79], the remaining window seems quite tight.

The DAMA/LIBRA annual modulation is a statistically very strong signal (significance of the order of 8σ), however the lack of supporting data from other experiments is disturbing. The annual modulation hinted at by CoGeNT [174, 175] is statistically much weaker, and the purported excess unmodulated signal may in fact be incompatible with the level of modulation reported. Also, it seems that the DAMA/LIBRA and GoGeNT signals, if interpreted as being due to dark matter, may be in tension with each other, even if one uses freedom in isospin violation, inelastic scattering, and non-standard halo properties [203–206]. At the moment this is one of the unsolved, frequently debated issues in the dark matter community.

The recent improvement of the upper limits on the WIMP-nucleon scattering cross section reported by CDMS II [188] and, in particular, XENON100 [54] are

truly impressive. Not only does it cast some doubt on other reported experimental results, the sensitivity is also good enough to start probing the parameter space of supersymmetric models [2]. The new calibration of the sensitivity to low-energy recoils of Xenon adds to the credibility of the new limits. The very good news is also that the installation of the next stage, a 1 ton liquid Xenon detector, has already started in the Gran Sasso experimental halls in Italy.

Of course, a much more decisive claim of detection of dark matter would result if any of the other methods, like a suitable new particle candidate being detected at the LHC, or a signature in γ-rays from the Galactic dark matter halo would be discovered.

In the first runs at LHC, no signs of a Higgs particle, nor supersymmetry or any other of the prime candidates for dark matter, have been discovered. On the other hand, the mass region 115–130 GeV, interesting for the lightest Higgs boson in the simplest versions of supersymmetry, was yet to be investigated, and in fact a weak indication around 125 GeV seem to have been found.

One possible scenario might be that such a Higgs particle is indeed found, but the particles carrying non-trivial R-parity all have masses beyond reach with the LHC. This is not impossible, depending on the amount of fine-tuning one is willing to tolerate. In fact, if one puts no prior constraints on the supersymmetric parameter space other than one should have the WMAP-measured relic density, and fulfill all other experimental constraints (cf. [143]), a mass for the lightest supersymmetric neutralino in the TeV region is generic. For such heavy dark matter neutralinos, the rate for direct detection will also be small, and it would seem impossible to test such a scenario. However, for this particular case indirect detection through gamma rays turns out to have an interesting advantage, as the new imaging air Cherenkov arrays like CTA will have their peak sensitivity in the energy range between a few hundred GeV to a few TeV [207].

Depending on the particular model realized in nature, Sommerfeld enhancement of indirect detection may also be operative. However, these large arrays will be served by a large astrophysical community which will be very much interested in transient or periodic events, meaning that a "boring" search for a stationary dark matter spectral signature during hundreds or even thousands of hours seem out of question. One may therefore consider a dedicated particle physics experiment, the "Dark Matter Array", DMA [208] only used for dark matter search. This would have great, and complementary, potential to the large direct detection experiments that are presently being planned. In fact, we mentioned, and you heard at the lectures by F. Aharonian (see Sect. 2.3 in his part), that there are ideas [110] on how to decrease the lower threshold for detection, something that could increase the sensitivity for DM detection considerably (see Fig. 7). If such a working prototype could be built, this idea may materialize in the next decade as a new way to search for phenomena beyond the Standard Model—with an expensive dedicated detector, still far below the cost of a new high-energy accelerator.

Of course, LHC data has already started to exclude some regions of supersymmetric parameter space, although not very much. This may be surprising, but is in fact due to the relative independence of the squark and gluino sector of supersymmetry,

Fig. 9 Scan of the MSSM
parameter space showing the
direct detection cross section
versus indirect detection
through gamma rays. The
uppermost points are excluded
by XENON100, and points
which survive also the LHC
2011 data are shown in *black*

$$S = N(E_\gamma > 1\ GeV)\sigma v_{\gamma cont}/M^2\ (10^{-29} cm^3 s^{-1} GeV^{-2})$$
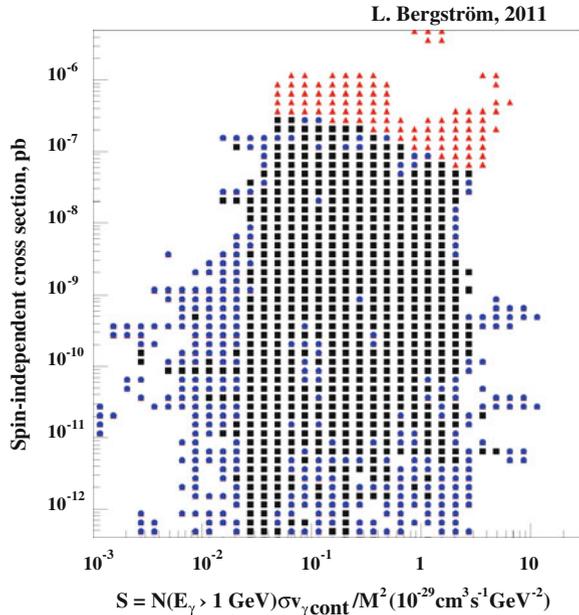
and the neutralino sector, which hosts the dark matter candidate. In fact, as mentioned, there are so-called split supersymmetry models, which have this dichotomy explicitly postulated [35].

The complementarity of direct and indirect detection is shown in Fig. 9, where also the effects on the parameter space caused by the XENON100 bounds and LHC 2011 bounds, respectively, are shown.

9 A Detailed Calculation: The Saas-Fee WIMP

An interesting question came up during the Saas-Fee Course: Could there be a cosmological contribution to the γ-ray spectrum making up the deficit in the diffuse γ-ray emission measured by FERMI? As we heard, this is not readily explained by adding the well-known sources like AGNs, millisecond pulsars and star-forming galaxies described by C. Dermer.

Here we will outline the simple steps in making the dark matter prediction for this flux, based on [209] (see [210] for a much more thorough treatment). We will see how that could lead us to predict a several hundred GeV dark matter particle—the Saas-Fee particle as we named it at the Saas-Fee Course in 2010. This was only published in online slides from my talk,[4] and should not be taken too seriously. However, as a

---

[4] Available on the website of the Course: http://isdc.unige.ch/sf2010/.

pedagogical example of a surprising effect of the accumulated dark matter structure in the universe it is quite instructive.

As we have seen, in the presently most successful model for structure formation, ΛCDM, most of the matter is in the form of non-relativistic cold dark matter (CDM), but with a contribution to the present-day energy density also from a cosmological constant (Λ). As shown by detailed N-body simulations (see, e.g., [59–61] and references therein), in such a picture large structures form by the successive merging of small substructures, with smaller objects generally being denser. N-body simulations also show that the dark matter density profile in clusters of galaxies and in single galaxies develops an enhancement of the density near the centre, although it is at present unclear how steep this increase is, and whether it even shows a cusp near the center like in the popular parametrization of NFW, Eq. (163), $\rho_{\mathrm{CDM}}(r) \sim r^{-\alpha}$ with α close to 1 [105] (a very similar profile, the Einasto profile, Eq. (164), does not show a singularity, but behaves rather similarly on moderate length scales).

At present, it is not clear whether these N-body predictions are in agreement or not with all available data (one recently acknowledged problem is, for example, the apparent lack of a halo mass threshold for dwarf galaxies [211]). On large scales, however, the ΛCDM scenario gives excellent agreement with observations. On smaller scales, the dynamical range of the simulations is not enough, and one of the main puzzles is how to properly include the non-linearities induced by baryonic matter in the form of supernova explosions and other feedback mechanisms.

Let us assume that the ΛCDM picture is basically correct and that structure forms hierarchically, with the number density of halos of mass M being distributed as $dN/dM \propto M^{-\beta}$ with $\beta \sim 1.9 - 2$, as predicted by the Press-Schechter theory [212] and also verified in N-body simulations. Furthermore, the concentration of halos grows in these simulations with decreasing mass.

It is interesting that the averaging involved in computing the integrated signal of annihilation γ-rays of unresolved cosmological dark matter gives results which are more robust to changes in the details of how the dark matter is distributed on small scales. (The same is actually also true for all sources which are completely encompassed by the angular resolution cone of a given γ-ray experiment, for the obvious source of the galactic centre, the prediction of fluxes differ by up to 4 orders of magnitude for different models: in particular they are very sensitive to the presence or not of a central cusp.)

Let us consider annihilation of a WIMP such as the lightest neutralino χ of the MSSM, as a template. The mass range is from around 20 GeV up to several TeV [2]. For the sake of pedagogy, let us start with the unrealistic case of all the dark matter is smoothly distributed with the density distribution being just a number, the average density, at all redshifts. The idea is that since the dark matter was more dense in the early universe, one may get a large (red-shifted) flux by integrating over the line of sight from 0 to very high redshifts. Actually, in papers preceding [209] this was the only effect considered. We will soon modify the results by introducing the effects of structure, which indeed increases the flux by many orders of magnitude.

9.1 The Flux in a Smooth Universe

The comoving number density n_c of WIMPS, after decoupling from chemical equilibrium ("freeze-out") at very large temperatures ($T \sim m_\chi/20$) is depleted only slightly due to self-annihilations, governed by the Boltzmann equation

$$\dot{n}_c = -\langle\sigma v\rangle(1+z)^3 n_c^2, \tag{173}$$

where $\langle\sigma v\rangle$ is the thermal- and angular-averaged annihilation rate, which, to an excellent approximation after freeze-out, is velocity independent, since the neutralinos move non-relativistically, and are always in a dominant S-wave component (at least for our supersymmetric WIMP templates).

Each pair of χ particles that disappears through annihilation give rise to N_γ photons on the average, with an energy distribution in the rest frame of the annihilation pair,

$$\frac{dN_\gamma(E)}{dE} = \frac{dN_{\text{cont}}}{dE}(E) + b_{\gamma\gamma}\delta\left(m_\chi - E\right). \tag{174}$$

Here the first term gives the average continuum gamma ray distribution per annihilating χ and we have also added a term for the possible $\gamma\gamma$ line contribution, with $b_{\gamma\gamma}$ being the branching ratio into $\gamma\gamma$ (one could also have an additional $Z\gamma$ channel).

A γ-ray observed today, at redshift $z = 0$, of energy E_0 would correspond to an energy at the emission redshift z of $E = (1 + z)E_0$. We can now track, using the Boltzmann equation, the number of WIMPs that have disappeared from redshift z until now, and fold in the energy distribution (174). Thus we get a first estimate of the level of the diffuse extragalactic γ-ray flux. As usual, H_0 is the Hubble parameter, and we use the relation between time and redshift (see, e.g., [1]) $d/dt = -H_0(1 + z)h(z)d/dz$ with

$$h(z) = \sqrt{\Omega_M(1+z)^3 + \Omega_K(1+z)^2 + \Omega_\Lambda} \sim \sqrt{\Omega_M(1+z)^3 + \Omega_\Lambda}. \tag{175}$$

Here Ω_M, Ω_Λ and $\Omega_K = 1 - \Omega_M - \Omega_\Lambda$ are the present fractions of the critical density given by matter, vacuum energy and curvature. We can here use the result from Sect. 2 that the universe to an excellent approximation is flat, $\Omega_K = 0$. We then obtain

$$\frac{dn_c(z)}{dz} = \frac{\langle\sigma v\rangle}{H_0}\left(\frac{(1+z)^2}{h(z)}\right)n_c(z)^2. \tag{176}$$

The differential energy spectrum of the number density n_γ of photons generated by WIMP annihilations is then given by

$$\frac{dn_\gamma}{dz} = N_\gamma\frac{dn_c}{dz} = \int_0^{m_\chi}\frac{dN_\gamma(E)}{dE}\frac{dn_c}{dz}dE. \tag{177}$$

Here, dn_c/dz can be computed directly from (176) to excellent accuracy, replacing the exact solution $n_c(z)$ by the present average number density of neutralinos n_0 on the right hand side. This we can do since the comoving number density does not change appreciably after freeze-out.

Neglecting the baryon contribution (as we will see, factors of order unity will not make a difference here, due to much larger uncertainties in structure formation), $\Omega_\chi \sim \Omega_M$, we obtain

$$n_0 = \rho_\chi/m_\chi = \rho_c \Omega_M/m_\chi. \tag{178}$$

Here $\rho_c = 1.06 \cdot 10^{-5}\, h^2$ GeV/cm^3 and h as before is the scaled Hubble parameter in units of 100 km s^{-1} Mpc^{-1}, $h \sim 0.7$. There are a few more effects we have to include. We have to use the fact that all photons move with velocity c and that the average flux is isotropic from each volume element where annihilation takes place, giving a factor $1/4\pi$ per steradian. The cross section times velocity average should, for Majorana particles, also be divided by 2, something which was missing in the original derivation [209], but added in [210] (see the published version). Some of the photons will be absorbed after travelling over cosmological distances. This can to the level of our approximate calculation be handled by introducing a simple energy- and redshift-dependent factor $e^{-z/z_{\max}}$ (or the more detailed calculation in [210] a more complicated factor depending on z and E_0).

The resulting γ-ray flux at the detector is then given by:

$$\phi_\gamma = \frac{c}{8\pi}\frac{dn_\gamma}{dE_0} = 4.2 \cdot 10^{-14}\ \text{cm}^{-2}\text{s}^{-1}\text{sr}^{-1}\text{GeV}^{-1}$$
$$\times \frac{\Gamma_{26}\Omega_M^2 h^3}{m_{100}^2} \int_0^{z_{up}} dz \frac{(1+z)^3 e^{-z/z_{\max}}}{h(z)} \frac{dN_\gamma(E_0(1+z))}{dE}. \tag{179}$$

where we defined $\Gamma_{26} = \langle\sigma v\rangle/(10^{-26}\ \text{cm}^3\,\text{s}^{-1})$ and m_{100} the mass in units of 100 GeV.

For the energies we are interested in, 1 GeV $< E_0 <$ 500 GeV, it is the starlight and (more poorly known) infrared background radiation which is most important, whereas the CMBR may be important for higher energies. An optical depth of order unity is reached for a redshift which in [210] was approximated by $z_{\max}^{\text{old}}(E_0) \sim 3.3(E_0/10\ \text{GeV})^{-0.8}$, which represented older results. However, the newer data discussed in the lectures by C. Dermer (see Sect. 7.3 in his part), indicate much less absorption. As a representative of the more recent evaluation of this absorption [213], we take instead the simple approximation $z_{\max}^{\text{new}}(E_0) \sim 2.3(E_0/50\ \text{GeV})^{-1.1}$.

The exponential form is a good approximation for small values of $z_{\max}$ as is dominant in most of the cases we study here. The upper limit of integration is given by kinematics, $z_{up} = m_\chi/E_0 - 1$, as the maximum rest frame energy of a photon in an annihilation event is $E = m_\chi$. The gamma line contribution to (179) is particularly simple, just picking out the integrand at $z + 1 = m_\chi/E_0$; it has the very distinctive and potentially observable signature of being asymmetrically smeared to lower energies (due to the redshift) and of suddenly dropping just above m_χ. Unfortunately,

for most models the branching ratio for this channel is too small to be measurable with present-day energy resolution, and we will drop it from now on. (This may however change when the high-resolution instrument GAMMA-400 [94] is operational towards the end of this decade. This is specified to have an energy resolution of 1 %, which will be a perfect instrument for searching for γ lines from annihilation, and also from models where dark matter decays radiatively [214].) The continuum emission will produce a characteristic, although less conspicuous feature, a smooth "bump" below around one tenth of the neutralino mass, and may be more difficult to detect. One should notice that there are particular models where radiative corrections ("internal bremsstrahlung") may give a significantly harder spectrum near $E_\gamma = m_\chi$, facilitating discrimination against most backgrounds [130, 131].

9.2 Including Effects of Cosmic Structure

To give an example of the results (which in [209] contained both obsolete SUSY models and not very accurate data from the old EGRET experiment), we take a generic model with mass 600 GeV, and the canonical WIMP averaged cross section times velocity of $\langle \sigma c \rangle = 3 \cdot 10^{-26}$ cm^3 s^{-1}, in the "concordance" cosmology $\Omega_M = 0.3, \Omega_\Lambda = 0.7, h = 0.7$. The continuous γ-ray rest frame energy distribution per annihilating particle comes mainly from hadronization and decay of π^0s and is conveniently parametrized to reasonable accuracy as (see Eq. (158))

$$dN_{\text{cont}}(E)/dE = (0.42/m_\chi)e^{-8x}/(x^{1.5} + 0.00014)$$

where $x = E/m_\chi$. This is valid for most quark jet final states, except for the top. Also, τ lepton decays give a somewhat harder γ-ray spectrum, and as mentioned internal bremsstrahlung may be important for certain types of models.

The most difficult, but also most important and interesting part of the calculation is to include the effects of structure formation. Following [209], we consider first a halo of mass M whose radial density profile can be described by $\rho_{DM}(r) = \rho'_{DM} f(r/a)$, with ρ'_{DM} being a characteristic density and a a length scale. These are found in N-body simulations not to be independent parameters, as smaller halos are generally associated with higher densities.

As a simple first model for structure formation, assume that the halo of mass M accreted from a spherical volume of radius R_M, determined by requiring that the average cosmological density times that volume is equal to M, $\rho_0 \cdot 4\pi R_M^3/3 = M$ (with $\rho_0 \sim 1.3 \cdot 10^{-6}$ GeV/cm^3). The increase of average squared overdensity per halo, which is what enters the annihilation rate, is given by:

$$\Delta^2 \equiv \left\langle \left(\frac{\rho_{DM}}{\rho_0}\right)^2 \right\rangle_{r < R_M} = \left(\frac{\rho'_{DM}}{\rho_0}\right) \frac{I_2}{I_1}, \tag{180}$$

where $I_n \equiv \int_0^{R_M/a} y^2 dy(f(y))^n$ with $y \equiv r/a$. Here the dependence on the limits of integration is rather weak, at least for profiles less cuspy than the NFW profile [105] (see Eq. (163)).

Computing I_2/I_1 numerically, and using values of ρ'_{DM}/ρ_0 as determined for Milky Way size halos we find values of Δ^2 of $1.5 \cdot 10^4$ for the NFW profile, and $7 \cdot 10^3$ for a cored, modified isothermal profile (modified so that the density falls as $1/r^3$ at large radii). The flux ratio, 2:1 for these two models should be compared with the ratios roughly 100:1 obtained within a few-degree cone encompassing the galactic center, showing the announced relative insensitivity to halo density profiles.

We should now also take into account that the number density of halos is scaling like $\sim 1/M^{1.9}$, and that small-mass halos are denser. We can resort to the highest-resolution N-body simulations available to date [215]. Fitting the concentration parameter of halos by

$$c \sim 100 \, (M_{\mathrm{vir}}/h^{-1} M_\odot)^{-0.08}, \tag{181}$$

one finds to a good approximation

$$\Delta^2 \sim 2 \cdot 10^5 M_{12}^{-0.2}, \tag{182}$$

where M_{12} is the halo mass in units of 10^{12} solar masses. This means that the total flux from a halo of mass M scales as $M^{0.8}$. Since the number density of halos goes as $M^{-1.9}$, the fraction of flux coming from halos of mass M scales as $M^{-1.1}$. Thus the γ-ray flux will dominantly come from the smallest CDM halos. In simulations, substructure has been found on all scales (being limited only by numerical resolution). For very small dark matter clumps, however, no gain in overdensity is expected, since once the matter power spectrum enters the k^{-4} region a constant density is expected. There are arguments [216] that structure is present in cold dark matter models all the way down to 10^{-6} or smaller [217]. We conservatively set $1 M_\odot$ as the minimal scale. In a more detailed treatment, one should also include effects of clumps within clumps, which increase the enhancement. However, destruction of DM clumps near large central densities of halos should also be included.

Finally, regarding redshift dependencies, we assumed in [209] a constant enhancement factor Δ^2 out to $z \sim 1$, and somewhat arbitrarily imposed quadratic growth in the enhancement factor from $z \sim 10$ to the fully developed epoch $z = 1$. (The computed flux is not very sensitive to this assumption.) Furthermore, in (179) we make the replacement $(1 + z)^3 \rightarrow 1$, reflecting the fact that the we are now dealing with a clumped, rather than a smooth distribution with density scaling $\sim(1 + z)^3$.

We thus arrive at the following expression for the flux including structure formation

$$\phi_\gamma = \frac{c}{8\pi} \frac{dn_\gamma}{dE_0} = 4.2 \cdot 10^{-14} \, \mathrm{cm}^{-2} \, \mathrm{s}^{-1} \, \mathrm{sr}^{-1} \, \mathrm{GeV}^{-1}$$
$$\times \frac{\Gamma_{26} \Omega_M^2 h^3}{m_{100}^2} \int_0^{z_{up}} dz \frac{\Delta^2(z) e^{-z/z_{max}}}{h(z)} \frac{dN_\gamma(E_0(1 + z))}{dE}. \tag{183}$$
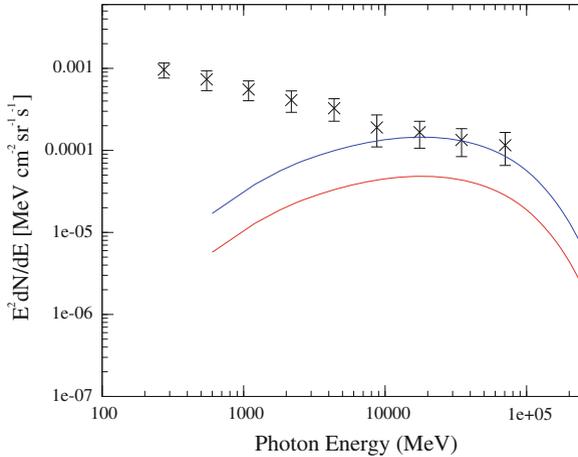
Fig. 10 The predicted diffuse extragalactic γ-ray flux computed using the methods described in the text, for a 600 GeV WIMP with a cross section compatible with the WMAP-inferred relic density and with different assumptions for the effects of structure. The diffuse extragalactic data was measured by FERMI-LAT [191]

9.3 The Saas-Fee WIMP

We find using (183) (see also [218]) that the flux from small halo structure is enhanced by roughly a factor $(4 - 10) \cdot 10^7$ compared to the smooth case, giving in the upper range observability for the annihilation parameters as used above. The uncertainties reside mainly in the still poorly known factor $\Delta^2(z)$ and its extrapolation to small halo masses (and also the effects of DM clumps within clumps, for instance).

In Fig. 10, we show the results for this 600 GeV WIMP model. The results are compared with the measurements from FERMI-LAT [191], and despite the fact that there is this uncertainty in the absolute rates, it is amusing, as I discussed at the Saas-Fee course, that the possible break in the FERMI data may be caused by a new contribution from 500–600 GeV mass annihilating dark matter ("The Saas-Fee WIMP", of which there would be one per 2 litres in the lecture hall at Les Diablerets as in all other places on Earth) setting in. It will obviously be interesting to follow the development of this data set during the next few years, to see if this model survives or even becomes more convincing.

It has of course to be remembered that the strength of the annihilation signal can be much lower than the proof-of-existence example chosen for Fig. 10 in which case a detection would be correspondingly more difficult. On the other hand, there may be particle physics effects (such as Sommerfeld enhancement) which could give a higher flux.

As a recent illustration of the importance of adding up all structure present, e.g., in galaxy clusters, can be mentioned the results of [219–223] and [224], where it was shown that by choosing particularly favourable, not too distant clusters, one is very

close to the current observational limits from FERMI-LAT. Indeed, there may even be (weak) indications from FERMI data of a signal [225].

A related type of analysis for the diffuse extragalactic case is performed in a similar way as when analyzing the angular fluctuations in the CMB. Also using this method of analysis, the conclusion is that with FERMI-LAT data one may be very near detection [226–229].

10 Primordial Black Holes as Dark Matter?

Sometimes one gets the question from lay persons, when telling that we are interested in the problem of dark matter: *Could it be black holes?* Black holes are in some sense dark: they do not emit light from within the event horizon, so the question is not completely out of context. However, the only black holes which we are relatively certain to exist are those around 2–20 solar masses, and the supermassive ones like the one residing at the Galactic center (of mass a few times $10^6 M_\odot$). We also know of even more massive ones (of mass up to a few times $10^9 M_\odot$) making up active galactic nuclei (AGNs). However, the galactic halos in which even the most massive, "supermassive", black holes reside have a factor of at least 1000 more total mass. Thus their influence on the cosmic energy balance is rather marginal. Also the solar mass type black holes which are produced as end results of stellar evolution constitute a small fraction, by far too small to explain the abundant and omnipresent dark matter. Finally, most black holes are in practice not very dark, as their concentrated mass has an effect on surrounding matter, causing various types of radiation processes, as ordinary matter is "eaten" by the black hole.

However, very massive black holes may be important for dark matter detection: if left undisturbed for a long time, they may adiabatically attract surrounding dark matter, changing the NFW-type distribution to a much more spiky cusp [230]. As the annihilation rate grows with the square of the dark matter density, this could give a dramatically increased rate of γ-rays, and in particular neutrinos (which are unlikely to be absorbed by surrounding matter). More extreme versions of this scenario are in fact ruled out already, due to the lack of unambiguous dark matter signals from the galactic centre. An interesting possibility is that intermediate mass black holes exist, where this type of signal could be close to detection, e.g., with FERMI-LAT [231].

10.1 Primordial Black Holes

There is also a small, be definite probability that small mass black holes would have been formed in the early universe. These "primordial" black holes would have to have been formed with a rather peculiar mass spectrum not to overclose the universe, and not to create too much radiation due to the interesting spontaneous emission of Hawking radiation, named after its discoverer (in theory).

Let us first remind ourselves of the metric surrounding a point mass M,

$$ds^2 = \left(1 - \frac{r_S}{r}\right) dt^2 - \frac{1}{1 - \frac{r_S}{r}} dr^2 - r^2 \left(d\theta^2 + \sin^2\theta d\varphi^2\right), \quad (184)$$

where r_S is the Schwarzschild radius, $r_S = 2GM$ (here G is Newton's constant, and as usual we set $c = 1$). A radial photon trajectory in this metric is given by $ds^2 = 0$, which gives

$$dt = \frac{dr}{1 - \frac{r_S}{r}}, \quad (185)$$

for $\theta = \phi = const$. The time for the photon to travel from r_i to r_f is thus

$$t_f - t_i = \int_{r_i}^{r_f} \frac{dr}{1 - \frac{r_S}{r}} = r_f - r_i + r_S \ln\left(\frac{r_f - r_S}{r_i - r_S}\right). \quad (186)$$

This diverges as $r_i \to r_S$, so that a light ray which starts at $r < r_S$ will never reach an outside observer: we have a black hole! If we define the sphere at $r = r_S$ as the areas A_{BH} of the black hole, we find

$$A_{BH} = 4\pi r_S^2 = 4\pi (2GM)^2 = 16\pi G^2 M^2. \quad (187)$$

It is interesting to contrast this with the behaviour of a solid sphere in classical physics, where $M \sim R^3$ which gives $R \sim M^{\frac{1}{3}}$, so that $A_{\text{class}} \sim M^{\frac{2}{3}}$. This difference is due to the strong curvature of space-time near the black hole.

As we noted, known black hole candidates either have a mass of a few solar masses (probably remnants of stellar collapse, as the maximal mass of a neutron star is somewhere between 1.4 solar masses—the Chandrasekhar mass—and a few solar masses), or a few million solar masses (Milky Way centre) to billions of solar masses (AGNs). There is no known present formation mechanism for BHs of mass less than a solar mass, so these, if they exist, must be primordial (PBHs), i.e. produced in the early universe, e.g. at some phase transition after inflation. There are various limits restricting formation scenarios, in general one has to "cook up" a power spectrum of density fluctuations which peaks at a particular mass length scale. When the horizon passes that scale, copious production of BHs may occur in such a scenario. An example can be found in a recent paper [232] where one tries to explain all of dark matter with PBHs, by having a power spectrum with a huge peak ($\delta\rho/\rho \sim 0.1$) at a scale corresponding to a black hole mass of $10^{-7} M_\odot$.

10.2 Hawking Radiation

If PBHs exist, one may detect them through Hawking radiation, as Hawking discovered in a remarkable 1975 paper [233] that a black hole should emit thermal radiation. This can be explained as a tunneling phenomenon [234].

Let us make an extremely simplified heuristic version of the derivation. Let us say that we have an isolated black hole. We can then for certain say that it is inside the Schwarzschild radius r_S. This uncertainty in the position of the radiation gives an uncertainty of the time $\Delta t \sim r_S/c = 2GM/c^3$, but the uncertainty relation between time and energy, $\Delta E \Delta t \sim \hbar/2$ gives

$$\Delta E \sim \frac{\hbar c^3}{4GM} \sim E_{th} = k_B T \to k_B T \sim \frac{\hbar c^3}{4GM} = \frac{1}{4GM}. \tag{188}$$

Thus, in our units, where also $k_B = 1$, the temperature $T = 1/(4GM)$. This is only a factor of 2π different from Hawking's result:

$$T_H = \frac{1}{8\pi GM}. \tag{189}$$

Of course, Hawking's derivation is much more beautiful, by explaining the radiation by mixing of positive and negative energy states due to the strong space-time curvature near the black hole. Another way to understand the process is that for a virtual particle pair created just at the horizon, one of the particles will be dragged into the black hole, releasing gravitational binding energy to the other particle, which can then appear as a real propagating state outside the horizon.

An interesting consequence of Hawking radiation and the equivalence principle is that an uniformly accelerated observer, with acceleration a, in empty space should see a thermal distribution of particles—the Unruh effect. The Unruh temperature is $T = a/2\pi$. Attempts have been made to measure the Unruh effect at accelerators, but present-day accelerations are not large enough. It has been argued, however, that the so-called Sokolov-Ternov effect (depolarization of electrons in a storage ring due to synchrotron radiation) really has the same origin as the Unruh effect—and it has been experimentally verified [235].

10.3 Thermodynamics of Black Holes

If we regard the Hawking temperature as a true thermodynamical temperature $T(M) = T(E)$, there should also be an entropy (Bekenstein entropy) associated with the BH:

$$T(E) = \frac{1}{8\pi GE}; \quad dS = \frac{dE}{T(E)} \rightarrow S = \int_0^M 8\pi GE dE = 4\pi GM^2 = \frac{1}{4}\frac{A_{BH}}{G}.$$
(190)

If we remember that $G = 1/M_{Pl}^2 = l_{Pl}^2$, we see that each "Planck area" of the surface of the BH contributes one quarter unit of entropy, and one gets huge numbers. This is still mysterious—what are the degrees of freedom describing the black hole, and why does ordinary matter that forms a BH suddenly increase its entropy enormously?

To describe black hole evaporation, it is useful to remember the form of a thermal distribution for a particle species

$$f_i(\mathbf{p}) = \frac{1}{e^{\frac{E_i - \mu_i}{k_B T}} \pm 1} = \frac{1}{e^{\frac{E_i - \mu_i}{T}} \pm 1}.$$
(191)

This means that for the rate of mass loss we can analogously write [236]

$$\frac{dM}{dt} = -\sum_j \frac{1}{2\pi} \int_{m_j}^\infty \Gamma_j \frac{E dE}{e^{8\pi GME} \pm 1} = \ldots = -5 \cdot 10^{25} f(M) M^{-2} \, gs^{-1}. \quad (192)$$

Here Γ_j is the absorption rate for particle of type j and the sum is over all particle-antiparticle pairs. This gives the evaporation time

$$\tau_{evap} \sim \int_{M_{min}}^{M_{max}} \frac{M^2}{f(M)} dM \sim \frac{6 \cdot 10^{-27}}{f(M_i)}\left(\frac{M_i}{1\,g}\right)^3 \, s.$$
(193)

Thus, only BHs with mass $> 10^{15}$ g are stable on cosmological time scales (so don't worry about BHs produced at LHC—they would evaporate immediately—if they exist!) Upper limits of γ-ray emission from EGRET and FERMI-LAT gives the approximate bound for light PBHs:

$$\Omega_{PBH}(M < 10^{15}\,g) < 10^{-8}.$$
(194)

Actually, since the temperature increases with decreasing mass, all particles, even more massive than those presently produced at accelerators, may give a contribution in the final "Hawking explosion". In particular, if supersymmetry is realized in nature, the end-time evolution may have interesting differences from the scenario with only Standard Model particles [237].

10.4 Formation of Primordial Back Holes

Let us discuss how primordial black holes formed in the early universe (see [238]). The relevant length scale is the particle horizon length, so that

$$M = \gamma M_{\text{particle horizon}} = 2 \cdot 10^5 \gamma \left(\frac{t}{1\,s} \right) M_\odot, \tag{195}$$

where $\gamma \sim 0.2$ depends on the detailed formation scenario. We can now compute the fraction of total energy density in black holes at formation:

$$\beta(M) \equiv \frac{\rho_{PBH}(t_i)}{\rho(t_i)} = 8 \cdot 10^{-29} \frac{1}{\sqrt{\gamma}} \left(\frac{g_i}{106.75} \right)^{\frac{1}{4}} \left(\frac{M}{M_\odot} \right)^{\frac{3}{2}} \left(\frac{n_{PBH}(t_0)}{1\,\text{Gpc}^{-3}} \right). \tag{196}$$

This means a contribution to Ω today of

$$\Omega_{PBH} = \frac{M n_{PBH}(t_0)}{\rho_c} = \left(\frac{\beta(M)}{1.2 \cdot 10^{-8}} \right) \sqrt{\gamma} \left(\frac{g_i}{106.75} \right)^{-\frac{1}{4}} \left(\frac{M}{M_\odot} \right)^{-\frac{1}{2}}. \tag{197}$$

The WMAP bound (the PBHs would behave gravitationally as cold dark matter) $\Omega < 0.25$, gives

$$\beta(M) < 2 \cdot 10^{-18} \frac{1}{\sqrt{\gamma}} \left(\frac{g_i}{106.75} \right)^{\frac{1}{4}} \left(\frac{M}{10^{15}\,g} \right)^{\frac{1}{2}}. \tag{198}$$

(This is valid for BHs that have not evaporated today, i.e., for $M > 10^{15}$ g.) It is convenient to divide out the cosmology/formation factors and consider the simpler expression for the energy density limit from WMAP:

$$\beta'(M) < \sqrt{\gamma} \left(\frac{g_i}{106.75} \right)^{-\frac{1}{4}} \beta(M) = 2 \cdot 10^{-18} \left(\frac{M}{10^{15}\,g} \right)^{\frac{1}{2}}. \tag{199}$$

Limits on $\beta'(M)$ can be obtained from a variety of data, from BBN and CMB in the early universe to the galactic and diffuse extragalactic γ-ray emission, gravitational lensing data and large scale structure. The limits we just computed on Ω_{PBH} is also important in the region $M \sim 10^{15}$–10^{27} g (for a detailed summary of the situation, see [238]).

To conclude: PBHs of mass less than around 10^{15} g cannot be the dark matter due to important constraints from the absence of Hawking radiation in 1–100 MeV γ-rays, but may still be a subdominant component. It is worthwhile to look for γ-ray signatures—a discovery of Hawking radiation would be truly wonderful!

At all masses, there are important bounds from a variety of methods. In principle, there are mass ranges where PBHs can still be the dark matter—all of dark matter, but one needs contrived production mechanisms such as a strongly peaked, and fine-tuned, density power spectrum.

11 Gravitational Waves

We will now, in view of the multi-messenger aspects of this lecture series, discuss one more type of radiation which is deeply linked to the theory of general relativity on which modern cosmology rests: gravitational radiation.

Due to the nonlinearity of Einstein's equations, it is virtually impossible to find exact solutions to the metric tensor $g^{\mu\nu}(\mathbf{r}, t)$ corresponding to the dynamics, for example, of a massive star which collapses to a black hole near the strong gravitational field of the star (using supercomputers, numerical studies can, however, be made). Far from the source of the gravitational field, it is on the other hand reasonable to use a first-order approximation. The gravitational deformation of space-time near celestial bodies like the Earth or the Sun due to conceivable astrophysical processes happening elsewhere in the Galaxy is indeed as we will see extremely tiny, which justifies such a perturbative approach. The same smallness of the effect unfortunately also make detection of gravitational radiation very challenging.

11.1 The Gauge Choice for Electromagnetism

Recall the way one derives the existence of electromagnetic waves in Maxwell's theory. One inserts the vector potential A^μ in the equations of motion for a vanishing current j^μ (since we are dealing with propagation in vacuum) to obtain

$$\Box A^\mu - \partial^\mu \left(\partial_\nu A^\nu \right) = 0 \tag{200}$$

Through the use of the gauge freedom $A^\mu \to A^\mu + \partial^\mu f$, we can choose A^μ to fulfill $A^0 = 0$ and also the so-called Lorentz condition $\partial_\nu A^\nu = 0$. This leads to the wave equation

$$\Box A^\mu = 0 \tag{201}$$

with solutions of the form

$$A^\mu(\mathbf{r}, t) = \epsilon^\mu e^{\pm i(\omega t - \mathbf{k} \cdot \mathbf{r})} = \epsilon^\mu e^{\pm i k^\mu x_\mu} \tag{202}$$

where $k^\mu k_\mu = 0$ (light-like propagation) and the gauge conditions $A^0 = 0$ and $\partial_\nu A^\nu = \nabla \cdot \mathbf{A} = 0$ translate into $\epsilon^0 = 0$ and $\mathbf{k} \cdot \epsilon = 0$. This means that the two physical degrees of freedom are transverse to the direction of propagation, and there is no time-like mode of propagation (this is deeply connected to the masslessness of the photon).

11.2 Gauge Choice for the Metric Perturbation

In the case of gravity waves in Einstein's theory of general relativity, we can similarly make a first-order expansion of the dynamical degrees of freedom, which are the components of the metric tensor field $g_{\mu\nu}$, around the constant Minkowski metric $\eta_{\mu\nu}$:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{203}$$

and work to the first non-vanishing order in $h_{\mu\nu}$.

Now we have a spin-2 field $h_{\mu\nu}$ instead of the vector quantity A^μ, but again we can use a gauge-like invariance (which in this case rather is re-parametrization invariance)

$$x_\mu \to x_\mu + \xi_\mu(x) \tag{204}$$

translating into

$$h_{\mu\nu} \to h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu. \tag{205}$$

Using this we may impose the so-called traceless gauge condition

$$h^\mu_{\ \mu} = 0 \tag{206}$$

The analogy of $A_0 = 0$ is

$$h_{0\nu} = h_{\nu 0} = 0, \tag{207}$$

and of the transversality condition

$$\nabla \cdot \mathbf{A} = 0 \tag{208}$$

is

$$\nabla_i h^{i\nu} = \nabla_i h^{\nu i} = 0. \tag{209}$$

The Einstein equation (neglecting back-reaction, i.e. the contribution to the energy-momentum tensor by the gravitational field itself) becomes simply

$$\Box h_{\mu\nu} = 0. \tag{210}$$

11.3 Solutions to the Wave Equation

Exactly like for photons we can write for the wave solutions to Einstein's equation

$$h_{\mu\nu} = E_{\mu\nu} e^{\pm i(\omega t - \mathbf{k} \cdot \mathbf{x})} \tag{211}$$

with $k^2 = \omega^2$, i.e. massless propagation, with the speed of light. (There have been brave attempts to replace Einstein's gravity with a massive theory, with the extra component having extremely small mass. This would lead to many interesting differences, perhaps even explaining the small value of the cosmological constant. So far, there has not appeared any generally accepted way to do this, however.)

We can represent $E_{\mu\nu}$ by a 4 × 4 matrix, which, exactly like for A_μ, should reflect the gauge choice. We know already that the $E_{0\nu}$ row and $E_{\mu 0}$ columns are zero. Also E has to be symmetric in the two indices (since the metric is). Further, $k^i E_{i\nu} = k^j E_{\mu j} = 0$, meaning that also the elements of the $E_{3\nu}$ column and $E_{\mu 3}$ row are zero for a wave propagating in the z-direction. So, we really just have zeros for our perturbative solution apart from a symmetric, traceless 2 × 2 matrix. A general such matrix is a linear combination of

$$E^+_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{212}$$

and

$$E^\times_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{213}$$

For a given value of the 3-component z, and at time t, we can then write

$$E_{\mu\nu}(t) = h_+(t) E^+_{\mu\nu} + h_\times(t) E^\times_{\mu\nu}. \tag{214}$$

Look at the case

$$h_+ \neq 0, \quad h_\times = 0. \tag{215}$$

At a given time, we have in the unperturbed case

$$\Delta s^2 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu = (\Delta t)^2 - \sum_i (\Delta x^i)^2 = -\sum_i (\Delta x^i)^2 \tag{216}$$

For two diametrically opposed points on the unit circle,

$$\Delta x^i = (2\cos\theta, 2\sin\theta, 0) \tag{217}$$

and their distance is

$$d_0 = \sqrt{-\eta_{ij}(\Delta x^i)(\Delta x^j)} = 2\sqrt{\sin^2\theta + \cos^2\theta} = 2. \tag{218}$$

In the perturbed case (i.e., if a gravity wave passes)
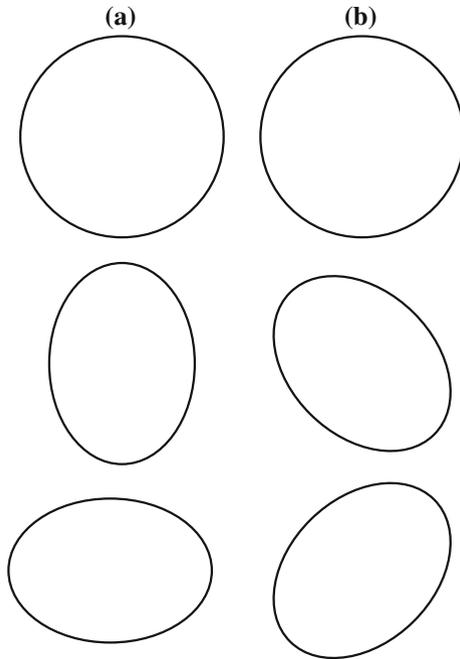
Fig. 11   a The deformation of the unit circle caused by gravity waves proportional to the polarization amplitude h_+. Shown are the unperturbed circle and the maximally stretched configurations along the two axes of symmetry, the x and y axes. **b** The corresponding pattern for the orthogonal polarization state described by the amplitude $h_\times$. Note that the axes along which stretching and compression occur form 45-degree angles to the x and y axes

$$d_+ = \sqrt{-(\eta_{ij} + h_+ E_{ij}^+)\Delta x^i \Delta x^j} = \sqrt{4 - h_+(t)4(\cos^2\theta - \sin^2\theta)} \qquad (219)$$

$$\simeq 2 - h_+(t)(\cos^2\theta - \sin^2\theta) = 2 - h_+(t)\cos 2\theta. \qquad (220)$$

For simplicity, we may work with real h by combining as usual the waves with the two signs in the exponential, giving

$$h_{\mu\nu}^+ = E_{\mu\nu}^+ h_+(t) = E_{\mu\nu}^+ \cos(\omega t - \mathbf{k} \cdot \mathbf{r}) \qquad (221)$$

and we see that the unit circle will be successively "compressed" or "squeezed" depending on the sign of the last factor (see Fig. 11, where the corresponding deformation caused by $h_\times$ is also shown).

These are the two independent quadrupole deformations of a circle. This means that the source of the gravitational field giving gravity waves has to have a quadrupole moment, Q. From dimensional reasoning,

$$h \sim \frac{G_N \ddot{Q}}{d} \sim \frac{4 G_N E_{\text{kin}}}{d}, \tag{222}$$

which is obtained by the crude estimate

$$Q \sim M l^2 \Rightarrow \dot{Q} = M 2 l \dot{l} = 2 M l v \Rightarrow \ddot{Q} \sim 2 M v^2 = 4 E_{kin}. \tag{223}$$

For objects in the Milky Way, typically $d \sim 10$ kpc, and with $E_{\text{kin}} \sim M_\odot$ we find

$$h \sim 10^{-17}. \tag{224}$$

On the other hand, for the distance appropriate for the Virgo galaxy cluster,

$$h \sim 10^{-20}. \tag{225}$$

These extremely tiny deformations is the reason for the non-detection so far of gravitational radiation, although there are promising objects like coalescing neutron stars which should have amplitudes nearing the experimental upper limits.

In a sense, gravity waves have already been indirectly detected, however, by comparing the slowing-down of the rotation rate of the binary pulsar system PSR 1913–16 by Hulse and Taylor (Nobel Prize of 1993):

$$\frac{dP}{dt} = (-2.4225 \pm 0.0056) \cdot 10^{-12}, \tag{226}$$

with the general relativistic calculation (with energy loss due to gravitational radiation):

$$\frac{dP_{GR}}{dt} = -2.40 \cdot 10^{-12}. \tag{227}$$

This excellent agreement has put severe limits on possible modifications of Einstein gravity. But effects of gravity waves have so far never been detected directly on Earth, despite an impressive increase in sensitivity of the LIGO experiment in the US, and VIRGO in Italy. Actually by combining several experiments and searching for time-coincident effects, one may both decrease various noise sources and increase the sensitivity for a signal. This is presently done by LIGO, VIRGO and GEO600 in Germany. All three detector are presently being upgraded to more advanced versions. However, it may be that a space experiment, LISA, will be needed to detect a significant signal. Its status is, however, at present unclear due to the difficult financial situation in most countries of the world.

We finally remind that there is also a possibility of detecting gravitational waves that are relics of dramatic processes in the early universe, such as during the epoch of inflation or during the formation of cosmic strings, if such exist. In that case, the most promising method is through analyzing the imprints they have made in the CMBR. As gravitational waves carry a quadrupole moment it is possible to distinguish their effects through studies of CMBR polarization. With the Planck satellite it is possible

to search for gravitational waves of very long wavelength generated through these hypothetical processes. Results are expected in early 2013.

12 Conclusions

This finishes our trip through the universe, looking at fundamental processes of similar interest to particle physicists, cosmologists, astrophysicists and astroparticle physicists alike. As hopefully has becomes clear, by combining information from all messengers that we have available: photons of all wavelengths, neutrinos, antimatter and perhaps gravitational waves, we may study from the Earth some of the most energetic and interesting processes in the universe. If we are lucky, we may even solve the problem of the nature of the dark matter, which has been with us since the times of Fritz Zwicky. Let us remind ourselves of his prophetic words from 1933 [8], after observing the rapid movement of the galaxy members of the Coma cluster, which pointed to an overdensity of matter in the cluster:

If this over-density is confirmed we would arrive at the astonishing conclusion that dark matter is present with a much greater density than luminous matter...

References

1. L. Bergström, A. Goobar, *Particle Astrophysics and Cosmology*, 2nd edn. (Springer, Dordrecht, 2006). ISBN-13: 978-3540329244
2. P. Gondolo, J. Edsjö, P. Ullio, L. Bergström, M. Schelke, E.A. Baltz, JCAP **0407**, 008 (2004) . [astro-ph/0406204]
3. L. Bergström, G. Bertone, in *Particle Dark Matter*, ed. by Bertone, G. (Cambridge University Press, Cambridge, 2010) pp. 491–506. ISBN-13: 978-0521763684
4. L. Bergström, invited talk at the EPS-HEP Conference, to appear in the Proceedings (Grenoble, 2011)
5. L. Bergström, *Dark Matter and Imaging Air Cherenkov Arrays*, to appear in Astroparticle Physics, special issue for the Cherenkov Telescope Array (2012)
6. M. Blomqvist, E. Mortsell, JCAP **1005**, 006 (2010). [arXiv:0909.4723 [astro-ph.CO]]
7. E. Komatsu et al. WMAP Collaboration. Astrophys. J. Suppl. **192**, 18 (2011). [arXiv:1001.4538 [astro-ph.CO]]
8. F. Zwicky, Helv. Phys. Acta **6**, 110 (1933)
9. D. Clowe, M. Bradac, A.H. Gonzalez, M. Markevitch, S.W. Randall, C. Jones, D. Zaritsky, Astrophys. J. **648**, L109–L113 (2006). [astro-ph/0608407]
10. P. Binetruy, G. Girardi, P. Salati, Nucl. Phys. B **237**, 285 (1984)
11. J. Edsjö, P. Gondolo, Phys. Rev. D **56**, 1879 (1997). [hep-ph/9704361]
12. A. Linde, Phys. Lett. **B129**, 177 (1983)
13. C. Itzykson, J.B. Zuber, *Quantum Field Theory*, (McGraw-Hill, New York, 1980)
14. L.J. Reinders, H. Rubinstein, S. Yazaki, Phys. Rept. **127**, 1 (1985)

15. J.B. Kogut, Rev. Mod. Phys. **55**, 775 (1983)
16. G. Altarelli, G. Parisi, Nucl. Phys. B **126**, 298 (1977)
17. M. Shaposhnikov, I. Tkachev, Phys. Lett. B **639**, 414 (2006). [hep-ph/0604236]
18. G.G. Raffelt, *Stars as Laboratories for Fundamental Physics*, (University of Chicago Press, Chicago, 1996)
19. K. Nakamura et al., Particle data group collaboration. J. Phys. G G **37**, 075021 (2010)
20. A. Osipowicz et al., KATRIN Collaboration. hep-ex/0109033
21. K. Abe et al., Super-Kamiokande collaboration. Phys. Rev. D **83**, 052010 (2011).[arXiv:1010.0118 [hep-ex]]
22. B. Aharmim et al., SNO collaboration. Phys. Rev. C **81**, 055504 (2010). [arXiv:0910.2984 [nucl-ex]]
23. S. Abe et al., KamLAND collaboration. Phys. Rev. Lett. **100**, 221803 (2008). [arXiv:0801.4589 [hep-ex]]
24. R. Acquafredda et al., OPERA collaboration. New J. Phys. **8**, 303 (2006). [hep-ex/0611023]
25. P. Adamson et al., MINOS collaboration. Phys. Rev. Lett. **107**, 011802 (2011). [arXiv:1104.3922 [hep-ex]]
26. F. Halzen, D. Hooper, New J. Phys. **11**, 105019 (2009). [arXiv:0910.4513 [astro-ph.HE]]
27. M. Cirelli, A. Strumia, M. Tamburini, Nucl. Phys. B **787**, 152 (2007). [arXiv:0706.4071 [hep-ph]]
28. K.A. Intriligator, N. Seiberg, Class. Quant. Grav. **24**, S741 (2007). [hep-ph/0702069]
29. R.R. de Austri, R. Trotta, L. Roszkowski, JHEP **0605**, 002 (2006). [hep-ph/0602028]
30. P. Scott, J. Conrad, J. Edsjö, L. Bergström, C. Farnier, Y. Akrami, JCAP **1001**, 031 (2010). [arXiv:0909.3300 [astro-ph.CO]]
31. C. Strege, G. Bertone, D.G. Cerdeno, M. Fornasa, R.R. de Austri, R. Trotta, [arXiv:1112.4192 [hep-ph]]
32. G. Jungman, M. Kamionkowski, K. Griest, Phys. Rept. **267**, 195 (1996). [hep-ph/9506380]
33. G. Bertone, D. Hooper, J. Silk, Phys. Rept. **405**, 279 (2005) [hep-ph/0404175]
34. L. Bergström, New J. Phys. **11**, 105006 (2009) [arXiv:0903.4849 [hep-ph]]
35. G. F. Giudice, A. Romanino, Nucl. Phys. B **699**, 65 (2004). [Erratum-ibid. B **706** (2005) 65] [hep-ph/0406088]
36. K. Griest, M. Kamionkowski, Phys. Rev. Lett. **64**, 615 (1990)
37. J. Ellis, T. Falk, K.A. Olive, Phys. Lett. **B444**, 367 (1998)
38. J. R. Ellis, T. Falk, G. Ganis, K.A. Olive, M. Srednicki, Phys. Lett. B **510**, 236 (2001). [hep-ph/0102098]
39. G.W. Bennett et al., Muon G-2 collaboration. Phys. Rev. D **73**, 072003 (2006). [hep-ex/0602035]
40. L. Covi, H. -B. Kim, J.E. Kim, L. Roszkowski, JHEP **0105**, 033 (2001). [hep-ph/0101009]
41. A. Kusenko, M.E. Shaposhnikov, Phys. Lett. B **418**, 46 (1998). [hep-ph/9709492]
42. M.I. Tsumagari, [arXiv:0910.3845 [hep-th]]
43. R. Catena, P. Ullio, JCAP **1008**, 004 (2010). [arXiv:0907.0018 [astro-ph.CO]]
44. J. Ellis, R. Flores, Phys. Lett. **B263**, 259 (1991)
45. J. Engel, Phys. Lett. **B264**, 114 (1991)
46. L. Bergström, P. Gondolo, Astropart. Phys. **5**, 263 (1996)
47. M. Drees, M. Nojiri, Phys. Rev. **D48**, 3483 (1993)
48. D. Bailin, G.V. Kraniotis, A. Love, Nucl. Phys. **B556**, 23 (1999)
49. L. Baudis et al., Heidelberg-Moscow collaboration. Phys. Rev. **D59**, 022001 (1999)
50. R.W. Schnee et al., CDMS collaboration, Phys. Rep. **307**, 283 (1998)
51. P. Belli et al., DAMA collaboration. Phys. Rep. **307**, 269 (1998)
52. P. Smith et al., UKDM collaboration. Phys. Rep. **307**, 275 (1998)
53. Z. Ahmed et al., CDMS-II collaboration. Phys. Rev. Lett. **106**, 131302 (2011).[arXiv:1011.2482 [astro-ph.CO]]
54. E. Aprile et al., XENON100 collaboration. Phys. Rev. Lett. **105**, 131302 (2010). [arXiv:1005.0380 [astro-ph.CO]]
55. A.K. Drukier, K. Freese, D.N. Spergel, Phys. Rev. D **33**, 3495 (1986)

56. A. Bottino, F. Donato, N. Fornengo, S. Scopel, Astropart. Phys. **13**, 215 (2000). [hep-ph/9909228]
57. S. Andreas, T. Hambye, M.H.G. Tytgat, JCAP **0810**, 034 (2008). [arXiv:0808.0255 [hep-ph]]
58. R. Bernabei et al., DAMA collaboration. Eur. Phys. J. **C56**, 333–355 (2008). [arXiv:0804.2741 [astro-ph]]
59. V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J.F. Navarro, C.S. Frenk et al., Mon. Not. Roy. Astron. Soc. **391**, 1685 (2008). [arXiv:0809.0898 [astro-ph]]
60. V. Springel, S.D.M. White, C.S. Frenk, J.F. Navarro, A. Jenkins, M. Vogelsberger, J. Wang, A. Ludlow et al., Nature **456N7218**, 73–80 (2008)
61. J. Diemand, M. Kuhlen, P. Madau, M. Zemp, B. Moore, D. Potter, J. Stadel, Nature **454**, 735 (2008). [arXiv:0805.1244 [astro-ph]]
62. S.W. Barwick et al., Phys. Rev. Lett. **75**, 390–393 (1995)
63. O. Adriani et al., PAMELA collaboration. Nature **458**, 607 (2009). [arXiv:0810.4995 [astro-ph]]
64. A.A. Abdo et al., The fermi LAT collaboration. Phys. Rev. Lett. **102**, 181101 (2009).[arXiv:0905.0025 [astro-ph.HE]]
65. D. Müller, K.K. Tang, Astrophys. J. **312**, 183 (1987)
66. R.L. Golden et al., Astrophys. J. **457**, L103 (1996)
67. H. Gast, J. Olzem, S. Schael, *Proceedings of XLIst Rencontres de Moriond. Electroweak Interactions and Unified Theories*, (La. Thuile, Italy, 2006), pp. 421–428
68. S.W. Barwick et al., Astrophys. J. **482**, L191 (1997)
69. J.J. Beatty et al., Phys. Rev. Lett. **93**, 241102 (2004)
70. I.V. Moskalenko, A.W. Strong, Astrophys. J. **493**, 694 (1998). [arXiv:astro-ph/9710124]
71. O. Adriani et al., Phys. Rev. Lett. **102**, 051101 (2009). [arXiv:0810.4994 [astro-ph]]
72. F.A. Aharonian, A.M. Atoyan, H. J. Völk, Astron. Astrophys. **294**, L41–L44 (1995)
73. D. Grasso et al., FERMI-LAT collaboration. Astropart. Phys. **32**, 140–151 (2009). [arXiv:0905.0636 [astro-ph.HE]]
74. S. Profumo, Cent. Eur. J. Phys. **10**, 1 (2011). [arXiv:0812.4457 [astro-ph]]
75. D. Malyshev, I. Cholis, J. Gelfand, Phys. Rev. D **80**, 063005 (2009). [arXiv:0903.1310 [astro-ph.HE]]
76. L. Bergström, T. Bringmann, J. Edsjö, Phys. Rev. D **78**, 103520 (2008).[arXiv:0808.3725 [astro-ph]]
77. N. Arkani-Hamed, D.P. Finkbeiner, T.R. Slatyer, N. Weiner, Phys. Rev. **D79**, 015014 (2009) [arXiv:0810.0713 [hep-ph]]
78. Y. Nomura, J. Thaler, Phys. Rev. D **79**, 075008 (2009). [arXiv:0810.5397 [hep-ph]]
79. D.P. Finkbeiner, L. Goodenough, T.R. Slatyer, M. Vogelsberger, N. Weiner, JCAP **1105**, 002 (2011) [arXiv:1011.3082 [hep-ph]]
80. K. Lubelsmeyer, A. Schultz von Dratzig, M. Wlochal, G. Ambrosi, P. Azzarello, R. Battiston, R. Becker, U. Becker et al., Nucl. Instrum. Meth. A **654**, 639 (2011)
81. F. Donato, N. Fornengo, P. Salati, Phys. Rev. D **62**, 043003 (2000). [hep-ph/9904481]
82. J.E. Koglin, T. Aramaki, S.E. Boggs, W.W. Craig, H. Fuke, F. Gahbauer, C.J. Hailey, N. Madden et al., J. Phys. Conf. Ser. **120**, 042011 (2008)
83. H.U. Bengtsson, P. Salati, J. Silk, Nucl. Phys. **B346**, 129 (1990)
84. T. Sjostrand, S. Mrenna, P.Z. Skands, JHEP **0605**, 026 (2006). [hep-ph/0603175]
85. M. Srednicki, S. Theisen, J. Silk, Phys. Rev. Lett. **56**, 263 (1986). Erratum-ibid **56**, (1986) 1883
86. L. Bergström, H. Snellman, Phys. Rev. **D37**, 3737 (1988)
87. A. Bouquet, P. Salati, J. Silk, Phys. Rev. **D40**, 3168 (1989)
88. G. Jungman, M. Kamionkowski, Phys. Rev. **D51**, 3121 (1995)
89. M. Urban et al., Phys. Lett. **B293**, 149 (1992)
90. L. Bergström, J. Kaplan, Astropart, Phys. **2**, 261 (1994)
91. P. Ullio, L. Bergström, Phys. Rev. **D57**, 1962 (1998)
92. L. Bergström, P. Ullio, J.H. Buckley, Astropart. Phys. **9**, 137–162 (1998). [astro-ph/9712318]

93. A.A. Abdo, M. Ackermann, M. Ajello, W.B. Atwood, L. Baldini, J. Ballet, G. Barbiellini, D. Bastieri et al., Phys. Rev. Lett. **104**, 091302 (2010). [arXiv:1001.4836 [astro-ph.HE]]
94. A.M. Galper, R.L. Aptekar, I.V. Arkhangelskaya, M. Boezio, V. Bonvicini, B.A. Dolgoshein, M.O. Farber, M.I. Fradkin et al., Astrophys. Space Sci. Trans. **7**, 75 (2011)
95. L. Bergström, P. Ullio, Nucl. Phys. **B504**, 27 (1997)
96. Z. Bern, P. Gondolo, M. Perelstein, Phys. Lett. **B411**, 86 (1997)
97. J. Hisano, S. Matsumoto, M.M. Nojiri, O. Saito, Phys. Rev. **D71**, 063528 (2005). [hep-ph/0412403]
98. L. Bergström, Phys. Lett. B **225**, 372 (1089)
99. L. Bergström, T. Bringmann, M. Eriksson, M. Gustafsson, Phys. Rev. Lett. **95**, 241301 (2005). [arXiv:hep-ph/0507229]
100. A. Birkedal, K.T. Matchev, M. Perelstein, A. Spray, [arXiv:hep-ph/0507194]
101. L. Bergström, T. Bringmann, M. Eriksson, M. Gustafsson, Phys. Rev. Lett. **94**, 131301 (2995). [arXiv:astro-ph/0410359]
102. T. Bringmann, L. Bergström, J. Edsjö, JHEP **0801**, 049 (2008). [arXiv:0710.3169 [hep-ph]]
103. H. Goldberg, Phys. Rev. Lett. **50**, 1419 (1983)
104. E.A. Baltz, L. Bergström, Phys. Rev. D **67**, 043516 (2003). [arXiv:hep-ph/0211325]
105. J.F. Navarro, C.S. Frenk, S.D.M. White, Astrophys. J. **462**, 563 (1996)
106. A. Abramowski et al., H.E.S.S. collaboration. Phys. Rev. Lett. **106**, 161301 (2011).[arXiv:1103.3266 [astro-ph.HE]]
107. J. Aleksic et al., The MAGIC collaboration. JCAP **1106**, 035 (2011). [arXiv:1103.0477 [astro-ph.HE]]
108. M. Wood, G. Blaylock, S.M. Bradbury, J.H. Buckley, K.L. Byrum, Y.C.K. Chow, W. Cui, I.D.l.C. Perez et al., [arXiv:0801.1708 [astro-ph]]
109. T. Bringmann, M. Doro, M. Fornasa, JCAP **0901**, 016 (2009). [arXiv:0809.2269 [astro-ph]]
110. F.A. Aharonian, A.K. Konopelko, H.J. Volk, H. Quintana, Astropart. Phys. **15**, 335–356 (2001). [astro-ph/0006163]
111. E. A. Baltz, B. Berenji, G. Bertone, L. Bergström, E. Bloom, T. Bringmann, J. Chiang, J. Cohen-Tanugi et al., JCAP **0807**, 013 (2008). [arXiv:0806.2911 [astro-ph]]
112. J. Silk, A. Stebbins, Astrophys. J. **411**, 439 (1993)
113. L. Bergström, J. Edsjö, P. Gondolo, P. Ullio, Phys. Rev. D **59**, 043506 (1999). [astro-ph/9806072]
114. B. Moore, S. Ghigna, F. Governato, G. Lake, T.R. Quinn, J. Stadel, P. Tozzi, Astrophys. J. **524**, L19 (1999). [astro-ph/9907411]
115. P. Gondolo, J. Silk, Phys. Rev. Lett. **83**, 1719 (1999)
116. S. Ritz, D. Seckel, Nucl. Phys. **304**, 877 (1988)
117. J. Edsjö, Ph.D. Thesis, Uppsala University, hep-ph/9504205
118. F. Halzen, Comments Nucl. Part. Phys. **22**, 155 (1997)
119. A. Bottino, N. Fornengo, G. Mignola, L. Moscoso, Astropart. Phys. **3**, 65 (1995)
120. V. Berezinsky et al., Astropart. Phys. **5**, 333 (1996)
121. L. Bergström, J. Edsjö, P. Gondolo, Phys. Rev. **D55**, 1765 (1997)
122. M. Blennow, J. Edsjö, T. Ohlsson, JCAP **0801**, 021 (2008). [arXiv:0709.3898 [hep-ph]]
123. L. Bergström, J. Edsjö, P. Gondolo, Phys. Rev. **D58**, 103519 (1998)
124. T.K. Gaisser, F. Halzen, T. Stanev, Phys. Rep. **258**, 173 (1995)
125. M. Kamionkowski, G. Jungman, K. Griest, B. Sadoulet, Phys. Rev. Lett. **74**, 5174 (1995)
126. J. Edsjö, P. Gondolo, Phys. Lett. **B357**, 595 (1995)
127. L. Bergström, J. Edsjö, M. Kamionkowski, Astropart. Phys. **7**, 147 (1997)
128. A. Strumia, PoS **EPS-HEP2009**, 012 (2009)
129. L. Bergström, G. Bertone, T. Bringmann, J. Edsjö, M. Taoso, Phys. Rev. D **79**, 081303 (2009). [arXiv:0812.3895 [astro-ph]]
130. T. Bringmann, L. Bergström, J. Edsjö, JHEP **0801**, 049 (2008). [arXiv:0710.3169 [hep-ph]]
131. L. Bergström, T. Bringmann, J. Edsjö, Phys. Rev. **D78**, 103520 (2008). [arXiv:0808.3725 [astro-ph]]
132. L. Bergström, Phys. Lett. **B225**, 372 (1989)

133. T. Bringmann, F. Calore, G. Vertongen, C. Weniger, [arXiv:1106.1874 [hep-ph]]
134. N.F. Bell, J.B. Dent, A.J. Galea, T.D. Jacques, L.M. Krauss, T.J. Weiler, [arXiv:1104.3823 [hep-ph]]
135. P. Ciafaloni, M. Cirelli, D. Comelli, A. De Simone, A. Riotto, A. Urbano, JCAP **1110**, 034 (2011). [arXiv:1107.4453 [hep-ph]]
136. A. Sommerfeld, Annalen der Physik **403**, 257 (1931)
137. M. Kamionkowski, S. Profumo, Phys. Rev. Lett. **101**, 261301 (2008). [arXiv:0810.3233]
138. A. Hryczuk, R. Iengo, P. Ullio, JHEP **1103**, 069 (2011). [arXiv:1010.2172 [hep-ph]]
139. J.R. Ellis, J.S. Hagelin, D.V. Nanopoulos, K.A. Olive, M. Srednicki, Nucl. Phys. **B238**, 453–476 (1984)
140. P. Meade, N. Seiberg, D. Shih, Prog. Theor. Phys. Suppl. **177**, 143–158 (2009). [arXiv:0801.3278 [hep-ph]]
141. C.F. Berger, J.S. Gainer, J.L. Hewett, T.G. Rizzo, JHEP **0902**, 023 (2009). [arXiv:0812.0980 [hep-ph]]
142. J.A. Conley, J.S. Gainer, J.L. Hewett, M.P. Le, T.G. Rizzo, [arXiv:1103.1697 [hep-ph]]
143. L. Bergström, Phys. Lett. **B225**, 372 (1989)
144. For downloading DarkSUSY, visit http://www.fysik.su.se/~edsjö/ds
145. T. Bringmann, New J. Phys. **11**, 105027 (2009a). [arXiv:0903.0189 [astro-ph.CO]]
146. D. Spolyar, K. Freese, P. Gondolo, Phys. Rev. Lett. **100**, 051101 (2008). [arXiv:0705.0521 [astro-ph]]
147. F. Iocco, Astrophys. J. **677**, L1–L4 (2008). [arXiv:0802.0941 [astro-ph]]
148. P. Scott, M. Fairbairn, J. Edsjö, Mon. Not. Roy. Astron. Soc. **394**, 82 (2009). [arXiv:0809.1871 [astro-ph]]
149. S. Sivertsson, P. Gondolo, Astrophys. J. **729**, 51 (2011). [arXiv:1006.0025 [astro-ph.CO]]
150. D. Tucker-Smith, N. Weiner, Phys. Rev. **D64**, 043502 (2001). [hep-ph/0101138]
151. J. March-Russell, C. McCabe, M. McCullough, JHEP **0905**, 071 (2009). [arXiv:0812.1931 [astro-ph]]
152. Y. Cui, D.E. Morrissey, D. Poland, L. Randall, JHEP **0905**, 076 (2009). [arXiv:0901.0557 [hep-ph]]
153. D.P. Finkbeiner, T. Lin, N. Weiner, Phys. Rev. **D80**, 115008 (2009). [arXiv:0906.0002 [astro-ph.CO]]
154. D.S.M. Alves, M. Lisanti, J.G. Wacker, Phys. Rev. **D82**, 031901 (2010). [arXiv:1005.5421 [hep-ph]]
155. M.T. Frandsen, F. Kahlhoefer, J. March-Russell, C. McCabe, M. McCullough, K. Schmidt-Hoberg, Phys. Rev. **D84**, 041301 (2011). [arXiv:1105.3734 [hep-ph]]
156. K.R. Dienes, B. Thomas, [arXiv:1106.4546 [hep-ph]]
157. Y. Nomura, J. Thaler, Phys. Rev. **D79**, 075008 (2009). [arXiv:0810.5397 [hep-ph]]
158. P.J. Fox, E. Poppitz, Phys. Rev. **D79**, 083528 (2009). [arXiv:0811.0399 [hep-ph]]
159. B. Kyae, JCAP **0907**, 028 (2009). [arXiv:0902.0071 [hep-ph]]
160. D. Spolyar, M.R. Buckley, K. Freese, D. Hooper, H. Murayama, [arXiv:0905.4764 [astro-ph.CO]]
161. A. Ibarra, A. Ringwald, D. Tran, C. Weniger, JCAP **0908**, 017 (2009). [arXiv:0903.3625 [hep-ph]]
162. T. Cohen, K.M. Zurek, Phys. Rev. Lett. **104**, 101301 (2010). [arXiv:0909.2035 [hep-ph]]
163. N. Haba, Y. Kajiyama, S. Matsumoto, H. Okada, K. Yoshioka, Phys. Lett. **B695**, 476–481 (2011). [arXiv:1008.4777 [hep-ph]]
164. A. Menon, D.E. Morrissey, C.E.M. Wagner, Phys. Rev. **D70**, 035005 (2004). [hep-ph/0404184]
165. M. Dine, N. Seiberg, S. Thomas, Phys. Rev. **D76**, 095004 (2007). [arXiv:0707.0005 [hep-ph]]
166. N. Bernal, A. Goudelis, JCAP **1003**, 007 (2010). [arXiv:0912.3905 [hep-ph]]
167. M. Berg, J. Edsjö, P. Gondolo, E. Lundström, S. Sjors, JCAP **0908**, 035 (2009). [arXiv:0906.0583 [hep-ph]]
168. D.B. Kaplan, Phys. Rev. Lett. **68**, 741–743 (1992)

169. D.E. Kaplan, M.A. Luty, K.M. Zurek, Phys. Rev. **D79**, 115016 (2009). [arXiv:0901.4117 [hep-ph]]
170. A.L. Fitzpatrick, D. Hooper, K.M. Zurek, Phys. Rev. **D81**, 115005 (2010). [arXiv:1003.0014[hep-ph]
171. M.T. Frandsen, S. Sarkar, Phys. Rev. Lett. **105**, 011301 (2010). [arXiv:1003.4505 [hep-ph]]
172. M.L. Graesser, I.M. Shoemaker, L. Vecchi, [arXiv:1103.2771 [hep-ph]]
173. S. Profumo, L. Ubaldi, JCAP **1108**, 020 (2011). [arXiv:1106.4568 [hep-ph]]
174. C. E. Aalseth et al., CoGeNT collaboration. Phys. Rev. Lett. **106**, 131301 (2011). [arXiv:1002.4703 [astro-ph.CO]]
175. C.E. Aalseth, P.S. Barbeau, J. Colaresi, J.I. Collar, J. Diaz Leon, J.E. Fast, N. Fields, T.W. Hossbach et al., Phys. Rev. Lett. **107**, 141301 (2011). [arXiv:1106.0650 [astro-ph.CO]]
176. G. Angloher, M. Bauer, I. Bavykina, A. Bento, C. Bucci, C. Ciemniak, G. Deuter, F. von Feilitzsch et al., [arXiv:1109.0702 [astro-ph.CO]]
177. D. Hooper, T. Linden, Phys. Rev. D **84**, 123005 (2011). [arXiv:1110.0006 [astro-ph.HE]]
178. D. Hooper, S. Profumo, Phys. Rept. **453**, 29 (2007). [hep-ph/0701197]
179. M. Gustafsson, E. Lundstrom, L. Bergström, J. Edsjö, Phys. Rev. Lett. **99**, 041301 (2007). [astro-ph/0703512 [ASTRO-PH]]
180. G.F. Bertone (ed.), *Particle Dark Matter*, (Cambridge University Press, Cambridge, 2010).
181. S.J. Asztalos et al., The ADMX collaboration. Phys. Rev. Lett. **104**, 041301 (2010). [arXiv:0910.5914 [astro-ph.CO]]
182. M.L. Garde, J. Conrad, J. Cohen-Tanugi, M. Kaplinghat, G. Martinez, For the Fermi-LAT collaboration. [arXiv:1111.0320 [astro-ph.HE]]
183. J. Hamann, S. Hannestad, G.G. Raffelt, I. Tamborra, Y.Y.Y. Wong, Phys. Rev. Lett. **105**, 181301 (2010). [arXiv:1006.5276 [hep-ph]]
184. A.X. Gonzalez-Morales, R. Poltis, B.D. Sherwin, L. Verde, [arXiv:1106.5052 [astro-ph.CO]]
185. A. Bottino, F. Donato, N. Fornengo, S. Scopel, Phys. Lett. **B423**, 109 (1998)
186. R. Arnowitt, P. Nath, Phys. Rev. **D60**, 044002 (1999)
187. J. Cherwinka, R. Co, D.F. Cowen, D. Grant, F. Halzen, K.M. Heeger, A. Karle et al., [arXiv:1106.1156 [astro-ph.HE]]
188. Z. Ahmed et al., The CDMS-II collaboration. Science **327**, 1619–1621 (2010). [arXiv:0912.3592 [astro-ph.CO]]
189. S.D. Hunter, D.L. Bertsch, J.R. Catelli, T.M. Digel, S.W. Dingus, J.A. Esposito, C.E. Fichtel, C.E. Hartman et al., Astrophys. J. **481**, 205–240 (1997)
190. W. de Boer, C. Sander, V. Zhukov, A.V. Gladyshev, D.I. Kazakov, Astron. Astrophys. **444**, 51 (2005). [astro-ph/0508617]
191. A.A. Abdo et al., The Fermi-LAT collaboration. Phys. Rev. Lett. **104**, 101101 (2010). [arXiv:1002.3603 [astro-ph.HE]]
192. J. Knodlseder, P. Jean, V. Lonjou, G. Weidenspointner, N. Guessoum, W. Gillard, G. Skinner, P. von Ballmoos et al., Astron. Astrophys. **441**, 513–532 (2005). [astro-ph/0506026]
193. G. Weidenspointner, G. Skinner, P. Jean, J. Knodlseder, P. von Ballmoos, G. Bignami, R. Diehl, A.W. Strong et al., Nature **451**, 159–162 (2008)
194. P. Meade, M. Papucci, A. Strumia, T. Volansky, Nucl. Phys. **B831**, 178–203 (2010). [arXiv:0905.0480 [hep-ph]]
195. L. Bergström, J. Edsjö, G. Zaharijas, Phys. Rev. Lett. **103**, 031103 (2009). [arXiv:0905.0333 [astro-ph.HE]]
196. D. Hooper, L. Goodenough, Phys. Lett. **B697**, 412–428 (2011). [arXiv:1010.2752 [hep-ph]]
197. D. Hooper, T. Linden, [arXiv:1110.0006 [astro-ph.HE]]
198. M. Chernyakova, D. Malyshev, F.A. Aharonian, R.M. Crocker, D.I. Jones, Astrophys. J. **726**, 60 (2011). [arXiv:1009.2630 [astro-ph.HE]]
199. A. Boyarsky, D. Malyshev, O. Ruchayskiy, [arXiv:1012.5839 [hep-ph]]
200. D. Hooper, D.P. Finkbeiner, G. Dobler, Phys. Rev. **D76**, 083012 (2007). [arXiv:0705.3655 [astro-ph]]
201. M. Su, T.R. Slatyer, D.P. Finkbeiner, Astrophys. J. **724**, 1044–1082 (2010). [arXiv:1005.5480 [astro-ph.HE]]

202. G. Bertone, M. Cirelli, A. Strumia, M. Taoso, JCAP **0903**, 009 (2009). [arXiv:0811.3744 [astro-ph]]
203. T. Schwetz, J. Zupan, JCAP **1108**, 008 (2011). [arXiv:1106.6241 [hep-ph]]
204. P.J. Fox, J. Kopp, M. Lisanti, N. Weiner, [arXiv:1107.0717 [hep-ph]]
205. M. Farina, D. Pappadopulo, A. Strumia, T. Volansky, [arXiv:1107.0715 [hep-ph]]
206. M.T. Frandsen, F. Kahlhoefer, C. McCabe, S. Sarkar, K. Schmidt-Hoberg, [arXiv:1111.0292 [hep-ph]]
207. Cherenkov Telescope Array Consortium, [arXiv:1008.3703 [astro-ph.IM]]
208. L. Bergström, T. Bringmann, J. Edsjö, Phys. Rev. **D83**, 045024 (2011). [arXiv:1011.4514 [hep-ph]]
209. L. Bergström, J. Edsjö, P. Ullio, Phys. Rev. Lett. **87**, 251301 (2001). [astro-ph/0105048]
210. P. Ullio, L. Bergström, J. Edsjö, C.G. Lacey, Phys. Rev. D **66**, 123502 (2002). [astro-ph/0207125]
211. I. Ferrero, M.G. Abadi, J.F. Navarro, L.V. Sales, S. Gurovich, [arXiv:1111.6609 [astro-ph.CO]]
212. W.H. Press, P. Schechter, Astrophys. J. **187**, 425 (1974)
213. J.D. Finke, S. Razzaque, C.D. Dermer, Astrophys. J. **712**, 238 (2010) [arXiv:0905.1115 [astro-ph.HE]]
214. W. Buchmuller, A. Ibarra, T. Shindou, F. Takayama, D. Tran, JCAP **0909**, 021 (2009). [arXiv:0906.1187 [hep-ph]]
215. M. Boylan-Kolchin, V. Springel, S.D.M. White, A. Jenkins, G. Lemson, Mon. Not. Roy. Astron. Soc. **398**, 1150 (2009). [arXiv:0903.3041 [astro-ph.CO]]
216. S. Hofmann, D.J. Schwarz, H. Stoecker, Phys. Rev. D **64**, 083507 (2001). [astro-ph/0104173]
217. T. Bringmann, New J. Phys. **11**, 105027 (2009). [arXiv:0903.0189 [astro-ph.CO]]
218. A.A. Abdo et al., Fermi-LAT collaboration. JCAP **1004**, 014 (2010). [arXiv:1002.4415 [astro-ph.CO]]
219. A. Pinzke, C. Pfrommer, L. Bergström, Phys. Rev. Lett. **103**, 181302 (2009).[arXiv:0905.1948 [astro-ph.HE]]
220. T.E. Jeltema, J. Kehayias, S. Profumo, Phys. Rev. **D80**, 023005 (2009). [arXiv:0812.0597 [astro-ph]]
221. M.A. Sanchez-Conde, M. Cannoni, F. Zandanel, M.E. Gomez, F. Prada, [arXiv:1104.3530 [astro-ph.HE]]
222. A. Pinzke, C. Pfrommer, L. Bergström, [arXiv:1105.3240 [astro-ph.HE]]
223. L. Gao, C.S. Frenk, A. Jenkins, V. Springel, S.D.M. White, [arXiv:1107.1916 [astro-ph.CO]]
224. L. Gao, C.S. Frenk, A. Jenkins, V. Springel, S.D.M. White, Mon. Not. Roy. Astron. Soc. **419**, 1721 (2012). [arXiv:1107.1916 [astro-ph.CO]]
225. J. Han, C.S. Frenk, V.R. Eke, L. Gao, S.D.M. White, [arXiv:1201.1003 [astro-ph.HE]]
226. S. 'i. Ando, E. Komatsu, Phys. Rev. D **73**, 023521 (2006). [astro-ph/0512217]
227. P.D. Serpico, E. Sefusatti, M. Gustafsson, G. Zaharijas, [arXiv:1109.0095 [astro-ph.CO]]
228. A. Cuoco, T. Linden, M.N. Mazziotta, J.M. Siegal-Gaskins, V. Vitale, E. Komatsu, f. t. F. -L. collaboration. [arXiv:1110.1047 [astro-ph.HE]]
229. J. Zavala, M. Vogelsberger, T.R. Slatyer, A. Loeb, V. Springel, Phys. Rev. D **83**, 123513 (2011). [arXiv:1103.0776 [astro-ph.CO]]
230. P. Gondolo, J. Silk, Phys. Rev. Lett. **83**, 1719 (1999). [astro-ph/9906391]
231. G. Bertone, A.R. Zentner, J. Silk, Phys. Rev. D **72**, 103517 (2005). [astro-ph/0509565]
232. P.H. Frampton, M. Kawasaki, F. Takahashi, T.T. Yanagida, JCAP **1004**, 023 (2010). [arXiv:1001.2308 [hep-ph]]
233. S.W. Hawking, Commun. Math. Phys. **43**, 199 (1975). [Erratum-ibid. **46** (1976) 206]
234. M.K. Parikh, F. Wilczek, Phys. Rev. Lett. **85**, 5042 (2000). [hep-th/9907001]
235. E.T. Akhmedov, D. Singleton, Int. J. Mod. Phys. A **22**, 4797 (2007). [hep-ph/0610391]
236. J.H. MacGibbon, B.J. Carr, Astrophys. J. **371**, 447 (1991)
237. M. Sahlen, [astro-ph/0304478]
238. B.J. Carr, K. Kohri, Y. Sendouda, J. 'i. Yokoyama, Phys. Rev. D **81**, 104019 (2010). [arXiv:0912.5297 [astro-ph.CO]]

Sources of GeV Photons and the Fermi Results

Charles D. Dermer

The Fermi Gamma-ray Space Telescope, now in its fourth year after launch, continues to make important discoveries and establish new results in all directions: pulsar astronomy, cosmic-ray physics, AGN and black-hole astrophysics, galactic astronomy, gamma-ray bursts (GRBs), limits on dark matter and Lorentz invariance violations, γ-ray astronomy of the Sun, moon, and Earth, etc. In this chapter, I survey results at medium energy γ rays, from some tens of MeV (at energies above nuclear de-excitation γ-ray lines) to ≈ 100 GeV where the ground-based γ-ray Cherenkov detector arrays become more sensitive. As shorthand, Fermi and medium-energy γ-ray astronomy is referred to here as "GeV astronomy," and ground-based Cherenkov γ-ray astronomy $\gtrsim 100$ GeV as "VHE astronomy."

The Fermi results already provided considerably more material than could be presented in the nine lectures that I gave on this subject at the Saas-Fee school on "Astrophysics at Very High Energies," held 15–20 March 2010 in Les Diablerets, Switzerland. Happily, though, Professor Lars Bergström gave a brilliant series of lectures that covered dark matter, so the absence here of extended discussion on dark matter and new physics in GeV astronomy reflects Prof. Bergström's better capabilities to address this subject. My lectures and this book chapter are therefore restricted to astrophysical and astroparticle sources of GeV radiation rather than to γ rays with origins in exotic particle physics and dark matter.

Even while the school was in progress, news appeared of a new type of γ-ray emitter of GeV photons that was identified with the symbiotic binary Nova V407 Cygni. This extraordinary system reveals an explosive shock evolving on a timescale of days to weeks, rather than the hundreds of years for supernova remnants (SNRs). I showed the 2010 March 18 ATEL [2] announcing the Fermi result in my last lecture, capturing the real-time recognition of a new type of galactic γ-ray source triggered by a thermonuclear explosion on a white dwarf fed by its binary red giant's wind.

C. D. Dermer (✉)
United States Naval Research Laboratory, Code 7653, 4555 Overlook Ave.,
Washington, DC 20375-5352, USA
e-mail: charles.dermer@nrl.navy.mil

The discovery has been published by the Fermi Collaboration [1], and is already triggering a new line of research strongly tied to MeV line astronomy and white-dwarf physics (cf. citations to [1]). So these lectures and this write-up capture an early glimpse of the state of knowledge of astrophysical sources of $\gtrsim 100$ MeV and GeV radiation obtained with the Large Area Telescope (LAT) on Fermi at ≈ 2–3 years into the mission, weighted by the extragalactic interests of the author.

Alongside the LAT on Fermi is the Gamma-ray Burst Monitor (GBM), sensitive to GRBs and bright transients in the 10 keV–30 MeV range. This review can only briefly mention important GRB results made with the GBM—by itself and with the LAT—and related GRB science employing Swift, INTEGRAL, and other detectors. Indeed, multiwavelength science is value-added science, and the possibilities to uncover the underlying physics of the powerful compact systems that are at the heart of high-energy astronomy are multiplied by radio/microwave/sub-mm/IR/optical/UV/ X-ray/MeV/TeV/neutrino/gravitational-wave data correlated with the GeV window, now observed with unprecedented clarity due to the LAT on Fermi.

The GeV field is in full discovery mode, not only due to Fermi but also thanks to AGILE, an EGRET-like sentinel of bright γ-ray transients, and to ground-based VHE observatories, particularly HESS, VERITAS, and MAGIC.[1] The νF_ν spectral energy distributions (SEDs) based on simultaneous and overlapping data sets are providing valuable information about Galactic sources, blazars and radio galaxies, and starburst and normal galaxies. GeV astronomy with Fermi is still in the midst of an active phase as Fermi accumulates data and increasing time makes faint sources visible and detection of rare cosmic transients more likely.

This chapter will be divided into sections that follow the course of lectures delivered at the Saas-Fee course. Though now somewhat out-of-date, these lectures can be found on my website in ppt format.[2] The topics of the lectures and the corresponding sections of this chapter are:

1. GeV instrumentation and the GeV sky with the Fermi Gamma-ray Space Telescope
2. First Fermi Catalog of Gamma Ray Sources and the Fermi Pulsar Catalog
3. First Fermi AGN Catalog
4. Relativistic jet physics and blazars
5. γ rays from cosmic rays in the Galaxy
6. γ rays from star-forming galaxies and clusters of galaxies, and the diffuse extragalactic γ-ray background
7. Microquasars, radio galaxies, and the extragalactic background light
8. Fermi observations of GRBs
9. Fermi acceleration, ultra-high energy cosmic rays, and black holes

---

[1] AGILE: Astro-rivelatore Gamma a Immagini LEggero (Gamma-ray Light Imaging Detector); HESS: High Energy Stereoscopic System, in Namibia. VERITAS: Very Energetic Radiation Imaging Telescope Array System, in Arizona. MAGIC: Major Atmospheric Gamma-ray Imaging Cherenkov Telescope in La Palma, Canary Islands.

[2] http://heseweb.nrl.navy.mil/gamma/~dermer/default.htm

Besides a discussion of the results of the Fermi Gamma-ray Telescope, I also include here some high-energy astrophysical theory essential for analysis of γ-ray data that builds on the research presented in my book with Govind Menon, Ref. [4]: "High Energy Radiation from Black Holes: Gamma Rays, Cosmic Rays, and Neutrinos," published by Princeton University Press in 2009. The book itself is focused on theory rather than observation, and develops the hypothesis that *the most energetic and powerful radiations in nature are made by particles energized through Fermi acceleration processes in black-hole jets powered by rotation.*

It is not possible, even at this early stage of the Fermi mission, to adequately summarize all the results from Fermi. But together with the accompanying lectures and book, this article provides a broad overview of some recent astrophysical advances in the Fermi era of γ-ray astronomy.

1 GeV Instrumentation and the GeV Sky with the Fermi Gamma-Ray Space Telescope

1.1 Historical Introduction

The year 2010 represents a highpoint in high-energy astronomy. Astronomical observatories at multiwavelength electromagnetic, neutrino, cosmic-ray, and gravitational wavebands are operating and collecting data. X-ray astronomy has sensitive pointing X-ray telescopes, Chandra and XMM-Newton, deployed in space to observe catalogued sources. Broadband X-ray and soft γ-ray observatories like Swift, INTEGRAL and Suzaku are available to measure the X/γ spectra of compact objects. The Rossi X-ray Timing Explorer (RXTE) continues to operate and provide a monitor of the brightest X-ray sources in the sky. Already mentioned were AGILE and the ground-based γ-ray air Cherenkov telescopes. The Cherenkov Telescope Array (CTA) consortium [3] is planning to start building as early as 2014 with array completion as early as 2018. Lacking at the moment, however, is an operating all-sky water-based Cherenkov telescope successor to MILAGRO in the 1–100 TeV range. This gap will soon be filled by the High Altitude Water Cherenkov (HAWC) experiment [5] on the Sierra Negra mountain near Puebla, Mexico, which uses 300 tanks rather than a single pond as utilized by MILAGRO.

The Pierre Auger Observatory, located in the Mendoza province of Argentina and covering an area the size of Rhode Island, had its third data release in 2010 [6], giving the spectrum and composition of ultra-high energy cosmic rays (UHECRs) with energies $E \gtrsim 10^{18}$ eV. The IceCube Neutrino Observatory, most sensitive to astrophysical neutrinos with energies of 100 TeV–100 PeV ($\approx 10^{14}$–10^{18} eV), has deployed all 86 of its strings in the latest austral summer (December, 2010), and has developed the DeepCore subarray that is sensitive to lower energy, ≈ 10–100 GeV, neutrinos. LIGO, the Laser Interferometry Gravitational-wave Observatory, hoping to detect gravitational radiation from coalescing compact objects, is operating at

design sensitivity. Development of an order-of-magnitude more sensitive Advanced LIGO has been approved, with completion expected for 2017 or thereafter. The NASA Laser Interferometry Space Antenna (LISA) is supported in the recent Astro-2010 study, though ESA is developing a separate space-based gravitational wave facility. Likewise, Constellation-X has evolved into IXO/ATHENA.[3] The large-area X-ray timing mission RXTE will soon be ended, with the ESA Large Observatory for X-ray Timing (LOFT) mission taking its place. Here in the US, a focusing hard X-ray telescope, NuSTAR (Nuclear Spectroscopic Telescope Array), in the 5–80 keV range, and the Gravity and Extreme Magnetism Small Explorer, GEMS, a NASA mission to study X-ray polarization of astrophysical sources in the 2–10 keV range, will soon be launched.

Brief History of GeV Astronomy

The progress of GeV astronomy in the range from ≈ 10 MeV to ≈ 100 GeV followed a period of remarkable advances starting over 40 years ago that culminated with the launch of Fermi. Prior to the Fermi-LAT, the most important detectors and some of their achievements in the development of medium-energy γ-ray astronomy are the following:

- 1967–1968, OSO-3, the Third Orbiting Solar Observatory, carried a Cherenkov counter experiment sensitive to >50 MeV γ rays, of which 621 were detected [19]. It discovered one source, the extended γ-ray emission of the Milky Way.
- 1972–1973, SAS-2, the Small Astronomy Satellite-2 [16], a spark chamber experiment sensitive to γ rays with energies between ≈ 30 and 200 MeV (and an integral flux above 200 MeV). It detected ≈ 8000 celestial γ-rays, making the first γ-ray identifications of the Crab and Vela pulsars, Geminga ($\gamma 195 + 5$, then unidentified), and the Cygnus region, and an association with Cygnus X-3 was suggested. A north-south asymmetry in the Galactic γ-ray plane emission was noted and attributed to the massive stars in the Gould belt. An isotropic γ-ray background radiation was first reported from analysis of SAS-2 data [15, 33].
- 1975–1982, COS-B, the Cosmic ray Satellite (option B), a magnetic-core, wire-matrix spark chamber sensitive to γ rays with energies from ≈ 30 MeV to ≈ 5 GeV, with an effective area of ≈ 50 cm^2 at 400 MeV [29]. Its orbit resulted in a large and variable background of charged particles. During its lifetime, it detected ≈ 200000 γ-rays, with the COS-B Caravane Collaboration announcing the discovery of 25 sources, most along the Galactic plane. These included 2CG 135+01, now identified with LSI +61° 303, and the first extragalactic source of $\gtrsim 100$ MeV γ rays, 3C 273 [18].
- 1991–2000, EGRET, the Energetic Gamma Ray Experiment Telescope on the *Compton Gamma Ray Observatory*, was a spark chamber experiment with large, $\gtrsim 1200$ cm^2 effective area between 200 MeV and 1 GeV, accompanied by excellent

background rejection [32]. During its 9 year mission lifetime, it collected $>2 \times 10^6$ γ rays and discovered that blazars and GRBs are luminous γ-ray sources. Because of the importance of this experiment in the development of GeV astronomy, we describe EGRET in more detail below.

- 2007–, AGILE, with a Gamma Ray Imaging Detector (GRID) and small calorimeter giving γ-ray sensitivity from 30 MeV to 50 GeV, roughly comparable to EGRET. Along with the GRID is the accompanying 18–60 keV super-AGILE hard X-ray survey instrument [14]. The first AGILE GRID catalog consists of 47 >100 MeV γ-ray sources with significance $>4\sigma$ from data taken between July 2007 and June 2008 [25]. The brightest sources in the catalog, e.g., Vela, Crab and, during periods of outburst, 3C 454.3, exceed integral photon fluxes of 2×10^{-6} ph(>100 MeV)/cm^2-s, while the weakest are fainter than 50×10^{-8} ph(>100 MeV)/cm^2-s.

The EGRET Experiment on *CGRO*

The *Compton Gamma Ray Observatory*, or *CGRO*, was a pioneering γ-ray space observatory (5 April 1991–4 June 2000) consisting of 4 instruments, OSSE, the Oriented Space Scintillator Experiment (sensitive from $\approx$50 keV to 10 MeV); BATSE, the Burst And Transient Source Experiment ($\approx$20 keV–few MeV); COMPTEL, the Compton Telescope ($\approx$800 keV–30 MeV); and EGRET, the Energetic Gamma Ray Experiment Telescope ($\approx$30 MeV–10 GeV). EGRET's detection method utilized a gas-filled spark chamber that tracked the γ ray after it converted to an electron-positron pair by pair-production interactions with nuclei in the thin Ta foils in the gas-filled spark chamber. Directional information was obtained by time-of-flight coincidence with a scintillator array in the lower spark chamber assembly. Charged particles were vetoed by an anticoincidence shield and total γ-ray energy was measured with EGRET's total absorption scintillator counter, the TASC, consisting of crystal scintillators and photomultipliers [20].

The Third EGRET (3EG) catalog [17], made from data accumulated between 1991 April 22 and 1995 October 3, consists of 271 sources of >100 MeV emission, including a Solar flare in 1991, the Large Magellanic Cloud, five pulsars, one probable radio galaxy, namely Centaurus A, and 66 high-confidence detections of blazars. The catalog also lists 27 lower confidence potential blazar detections, and contains 170 unidentified EGRET sources lacking associated sources. Five GRBs were detected with the spark chamber on EGRET. Figure 1 shows a skymap of the 3EG sources.

The EGRET Field of View (FoV), defined roughly by the solid angle within which the effective area is greater than 1/2 of the on-axis effective area, was $\approx$0.5 sr, or $\approx$1/24th of the full sky. EGRET operated in a pointing mode, and targeted one region of the sky for two weeks, representing $\sim$10^6 s after time for Earth occultation and time spent in the South Atlantic Anomaly are subtracted. The Point Spread Function (PSF) at 100 MeV was $\approx$5.7°, with the PSF improving roughly as $\approx E^{-1/2}$ [31, 32]. The first 18 months of the mission were devoted to a full-sky survey.
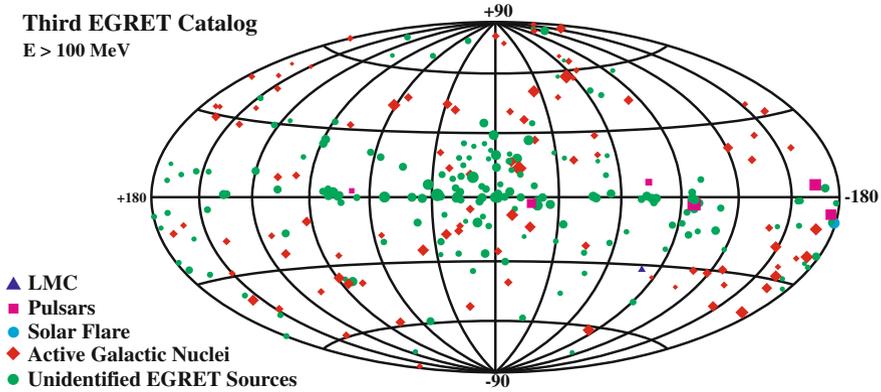
Fig. 1 EGRET all-sky map of γ-ray sources [17]

Important results from analysis of the full 6 years of data (degradation of spark chamber gas led to smaller apertures after 4 years into the mission) include the following:

1. Diffuse extragalactic background with intensity ≈ 1.5 keV/cm^2-s-sr, corresponding to an energy density $u_\gamma \approx 10^{-17}$ erg/cm^3;
2. The galactic flux is as much as $50\times$ brighter than the extragalactic flux, and much softer, but not as soft as expected if the γ rays are formed by secondary nuclear production by Galactic cosmic rays with the same spectrum as observed locally (the so-called "EGRET excess");
3. Typical fluxes of EGRET sources are between $\sim (10^{-7}$ and $10^{-6})$ ph(>100 MeV) /cm^2-s, with a typical 2-week on-axis limiting flux at $\approx (15\text{--}25) \times 10^{-8}$ ph($>$ 100 MeV)/cm^2-s;
4. Galactic sources, including young radio pulsars;
5. Sources with significant flux variability.

Temporal variability is an essential characteristic of GRBs and blazars, Solar flares, and now V407 Cyg. The large FoVs of the LAT and Swift, and larger still with BATSE and GBM, are crucial for study of γ-ray transients.

Point Source Sensitivity of EGRET

The best description of results leads to the question of units in GeV astronomy. Because medium-energy and high-energy γ-ray astronomy is challenged by limited signal counts, an integral photon flux is the most natural unit. For the EGRET experiment, units of 10^{-6} ph(>100 MeV)/cm^2-s are suitable, as this value roughly separates signal-dominated and noise-dominated detection in EGRET, as we now show. Units 10^{-8} ph(>100 MeV)/cm^2-s are more suitable for Fermi sources.

As measured with EGRET, the diffuse, or at least unresolved intensity of photons at 100 MeV, is measured to be $\epsilon I_\epsilon \approx 1.5\,\mathrm{keV/cm^2}$-s-sr at $E \equiv \epsilon m_e c^2 \approx 100\,\mathrm{MeV}$. Writing

$$\epsilon I_\epsilon = 1.5 \left(\frac{\epsilon}{\epsilon_{100}} \right)^{2-\Gamma_\gamma} \frac{\mathrm{keV}}{\mathrm{cm^2\text{-}s\text{-}sr}}, \tag{1}$$

with photon number index Γ_γ and $\epsilon_{100} = 100\,\mathrm{MeV}/m_e c^2$, implies

$$\frac{dN(>\epsilon)}{dA\,dt\,d\Omega} = \int_\epsilon^\infty d\epsilon' \, \frac{\epsilon' I_{\epsilon'}}{m_e c^2 \epsilon'^2} = \frac{1.5 \times 10^{-5}}{\Gamma_\gamma - 1} \left(\frac{\epsilon}{\epsilon_{100}} \right)^{1-\Gamma_\gamma} \frac{\mathrm{ph}(>100\,\mathrm{MeV})}{\mathrm{cm^2\text{-}s\text{-}sr}}. \tag{2}$$

EGRET measured a hard spectrum, with $\Gamma_\gamma \approx 2.1$ [26], whereas LAT measures a softer spectrum, with $\Gamma_\gamma \approx 2.4$ [7], and with a 100 MeV intensity only 60 % as large as the EGRET intensity.[4] For the EGRET intensity, therefore, $dN(>\epsilon_{100})/dA\,dt\,d\Omega \approx 1.5 \times 10^{-5}\,\mathrm{ph}(>100\,\mathrm{MeV})/\mathrm{cm^2}$-s-sr.

The EGRET PSF at 100 MeV, as previously noted, is $\approx 5.7°$. In comparison, the LAT PSF (68 % containment radius) at 100 MeV is $\approx 3.5°$, with the PSF dropping to $\approx 0.6°$ for 1 GeV photons, and $\approx 0.15°$ for 10 GeV photons (for conversion in the thin layers) [10]. The EGRET PSF at 100 MeV represents about $\pi(5.7°)^2 \approx 100$ square degrees, or $\approx 1/400$th of the full sky. Thus the flux from each patch corresponding to the EGRET PSF is $\approx 5 \times 10^{-7}\,\mathrm{ph}(>100\,\mathrm{MeV})/\mathrm{cm^2}$-s. A γ-ray source is *signal dominated* for EGRET when its flux is $\gtrsim 10^{-6}\,\mathrm{ph}\,(>100\,\mathrm{MeV})/\mathrm{cm^2}$-s. By contrast, it is *noise dominated* when its flux is $\ll 10^{-6}\,\mathrm{ph}(>100\,\mathrm{MeV})/\mathrm{cm^2}$-s.

The time Δt needed to accumulate 100 photons with EGRET, consisting of 50 signal S and 50 background B photons, and to give a detection at the $\approx S/\sqrt{2B} \cong 5\sigma$ level, is found through an expression for the integral photon flux $F(>\epsilon)$ of the source. For a source at the level of

$$F_{-8} \equiv \frac{F(>\epsilon_{100})}{10^{-8}\,\mathrm{ph}(>100\,\mathrm{MeV})/\mathrm{cm^2\text{-}s}}, \tag{3}$$

the number of detected photons is $\approx 10^{-8} F_{-8} \times \Delta t \times 1000\,\mathrm{cm^2}$, so that $\approx 50\,\mathrm{ph}$ can be detected from a bright source at the level of $F_{-8} \sim 10^2$ during an EGRET observation period of $\Delta t \approx$ one day (during half this time, the Earth is occulted). A nominal 2-week observation period with $\Delta t \approx 10^6$ s on source gives a 5σ limiting sensitivity for sources with $F(>\epsilon) \approx 2 \times 10^{-7}\,\mathrm{ph}(>100\,\mathrm{MeV})/\mathrm{cm^2}$-s.

These sorts of arguments can be used to estimate the time needed to make a detection and resolve temporal variability with EGRET, LAT, and counter detectors with broad FoVs, including neutrino telescopes. The Fermi LAT becomes noise dominated at much lower flux levels than EGRET, with sources regularly detected

---

[4] A number index $\Gamma_\gamma = 2.41 \pm 0.05$ and intensity normalization $I(>100\,\mathrm{MeV}) = (1.03 \pm 0.17) \times 10^{-5}/\mathrm{cm^2}$-s-sr for the intensity spectrum of the isotropic diffuse γ-ray background are measured from the first year Fermi data [7]. The lower diffuse extragalactic flux measured with Fermi compared to EGRET is partially but not entirely due to resolving more point sources out.

at $F_{-8} < 1$. Note that the background is smaller at higher energies as a result of the smaller PSF, but the smaller flux usually reduces significance for a given observation time except for hard-spectrum sources. A better approach for characterizing detection significance is likelihood analysis [21] (described below), but the detection significance can be simply estimated as outlined above [12].

The above estimates apply to high galactic latitude, $|b| > 10°$ sources where the γ-ray sky is dominated by extragalactic sources and unresolved isotropic γ-ray background. At lower galactic latitudes, the diffuse γ-ray emission from cosmic-ray interactions with gas and dust makes source detection more difficult. Subtraction of the diffuse emission and nonuniform and variable background requires a Galactic model for cosmic-ray/gas interactions. The sensitivity of source detection to background model is seen in the alternate 3EGR analysis of the EGRET data by Casandjian & Grenier [11]. They do not confirm 107 3EG sources, most in the vicinity of the Gould Belt, and find 30 new sources in the full 9-year data set.

Long after EGRET's effective lifetime had expired (though BATSE, OSSE, and COMPTEL were still collecting valuable data), *CGRO* was deorbited into the Pacific Ocean in 2000. The launch of *INTEGRAL* in October 2002, Swift on November 20, 2004, and *AGILE* on April 23, 2007, helped fill the γ-ray gap leading to Fermi.

1.2 Fermi Gamma-Ray Space Telescope

The Fermi Gamma-ray Space Telescope consists of two major detector systems, namely

1. The Large Area Telescope (LAT), a pair-conversion telescope sensitive between ≈ 20 MeV and $\gtrsim 300$ GeV, with the higher energy limit a result of the vanishing small detection rate given Fermi's ~ 1 m^2 aperture. The Fermi-LAT has opened the previously unexplored ≈ 10–100 GeV window, as self-vetoing due to particle shower backsplash in the anti-coincidence detector may have reduced EGRET's effective area above ≈ 5 GeV [27]. Fermi nominally operates in a survey mode, and given its large FoV, scans the entire sky every 3 h.
2. The Gamma-ray Burst Monitor (GBM), sensitive in the 8 keV–40 MeV range, consisting of 12 NaI detectors sensitive between 8 keV and 1 MeV, and 2 BGO detectors sensitive in the 0.15–40 MeV range. The scintillator detectors surround the base of the LAT, and view the entire unocculted sky which, at the nominal 565 km altitude of Fermi, represents $\approx 2/3$rd of the full sky.

Some of the science questions that Fermi was designed to answer are:

- How do supermassive black holes in Active Galactic Nuclei (AGN) create powerful jets of material moving at nearly light speed? What are the jets made of?
- What are the mechanisms that produce GRB explosions? What is the energy budget?
- What is the origin of the cosmic rays that pervade the Galaxy?

- How does the Sun generate high-energy γ rays in flares?
- How has the amount of starlight in the universe changed over cosmic time?
- What are the unidentified γ-ray sources found by EGRET?

GLAST Becomes Fermi

After a number of delays, a Delta II 7920-H(eavy) rocket carrying the Gamma-ray Large Area Space Telescope—GLAST—payload was launched from Cape Canaveral Air Station on 2008 June 11 at 12:05 pm EDT into a 565 km altitude circular orbit with 25.6° inclination and a 96 min period. GLAST completed a 60 day checkout period by early August, and released its first-light image on 26 August 2008, when it was renamed the Fermi Gamma-ray Space Telescope, after the Italian-American physicist Enrico Fermi (1901–1954).

The first-light image, based on only 4 days of observation, already reveals dozens of sources to the eye, with the blazar 3C 454.3 almost as bright as the Vela pulsar, which is the brightest persistent γ-ray source in the sky ($F_{-8} \cong 1060$). Limiting fluxes of EGRET sources are at the level of $F_{-8} \approx 20$–30 for a two-week observation, with a corresponding all-sky flux limit of $F_{-8} \approx 15$–30 for a year of observing (given the FoV of EGRET). For a source with a flat νF_ν SED, or a number index $\Gamma_\gamma = -2$, Fermi reaches $F_{-8} \sim 1$ in one year over the entire high Galactic latitude $|b| > 10°$ sky. Due to its energy-dependent effective area and PSF, limiting fluxes are strongly dependent on both source and background spectrum. Hard-spectrum sources with number index ≈ -1.5 are detected with comparable significance as soft spectrum sources, but at integral photon fluxes as low as $F_{-8} \approx 0.1$.

The LAT operates in a nominal scanning mode whereby the spacecraft rocks about the zenith. The rocking angle was equal to 39° in the first part of the Fermi mission, and then increased to 50° after 3 September 2009. The larger rocking angle gives a more uniform exposure, but with the loss of data due the increased fluorescence γ-ray emission from cosmic-ray bombardment of the Earth's atmosphere.[5] The LAT observes the entire sky every two orbits, or ≈ 3 h, by rocking north and south of the zenith on alternate orbits, with each point in the sky receiving ≈ 30 min exposure during this time. Onboard triggering has also been enabled for the Fermi spacecraft to autonomously slew. As of October 2011, more than 30 autonomous repoint requests (ARRs) have taken place, resulting in 5-h pointing mode observations in response to bright GRBs detected with the GBM (the length of the ARR repoint was reduced in the third year of the mission). Several dedicated Targets of Opportunity (ToO) pointings have been executed, including one to 3C 454.3, two to the flaring Crab, one to Cyg X-3, and two Solar pointings.[6] A regular schedule of nadir pointings to look for γ rays from terrestrial γ-ray flashes, or TGFs, is underway.

The Fermi/GLAST LAT Collaboration is an international organization originally comprised of institutions in France, Italy, Japan, Sweden, and the United States.

---

[5] This is also called, colloquially and inaccurately, "γ-ray albedo," or just "albedo".

[6] http://fermi.gsfc.nasa.gov/ssc/observations/timeline/posting/

There are about 120 full Collaboration members, and a total of about 200 scientific members, including affiliated scientists, postdocs, and students. The principal investigator of the project is Professor Peter Michelson of Stanford, and the project scientist during the development, commissioning, and early science phase was Dr. Steven Ritz, now professor at UC Santa Cruz. Dr. Julie McEnery of NASA's Goddard Space Flight Center is now Fermi LAT project scientist. A review of the first-year LAT results is given by Michelson, Atwood, and Ritz [22], and the detailed instrument description can be found in the paper by Atwood et al. (2009) [10].

In recognition of the importance of the Fermi mission to high-energy astronomy, Peter Michelson, Bill Atwood, and the Fermi Gamma Ray Space Telescope LAT team were awarded the 2011 Rossi Prize of the High Energy Astrophysics Division of the American Physical Society, "for enabling, through the development of the Large Area Telescope, new insights into neutron stars, supernova remnants, cosmic rays, binary systems, active galactic nuclei, and gamma-ray bursts." UCSC physicist Bill Atwood was also recently awarded the 2012 W. K. H. Panofsky Prize in Experimental Particle Physics by the American Physical Society.

LAT Instrument Description

The LAT is a pair-conversion telescope composed of 16 tracker-converter towers, a calorimeter, an anticoincidence shield, and an electronics system to filter signal from background. The conversion of γ-ray photons to electron-positron pairs takes place when a γ ray produces a pair in one of the 16 high-Z tungsten planes in each tower. The photon direction is measured with a precision silicon-strip tracker with a $228\,\mu$ pitch (separation between strips), totaling 8.8×10^5 channels distributed over 18 tracker planes in each of 16 tracker modules. The Si tracker follows the e^+-e^- pairs to give direction, and the photon energy is measured in a by a cesium-iodide calorimeter in a hodoscopic assembly that images the shower, allowing for additional rejection of non-photon-like events. A segmented anticoincidence detector (ACD) rejects background of charged cosmic rays. Segmentation of the ACD helps prevent self-vetoing at high energies (Fig. 2).

Table 1 gives the performance characteristics of the Fermi LAT. Depending on the source type, different classes of events are adopted to maximize sensitivity. Three hierarchical analysis classes based upon analysis cuts and event selections are generally used for Fermi analysis. For the most stringent *diffuse* event class used to identify γ rays, the on-axis effective area increases from $\approx 1500\,\mathrm{cm}^2$ at 100 MeV to $\approx 8000\,\mathrm{cm}^2$ at 1 GeV, and is roughly constant at higher energies. The *transient* event class is used for the study of GRBs and brief transients where the amount of background can be greater due to the smaller time over which data is taken. The *source* class, intermediate to the other two, is where the residual instrumental plus environmental background is comparable to the extragalactic diffuse emission measured with EGRET.

Fig. 2 Schematic of the Large Area Telescope on Fermi (from Atwood et al. 2009 [10])
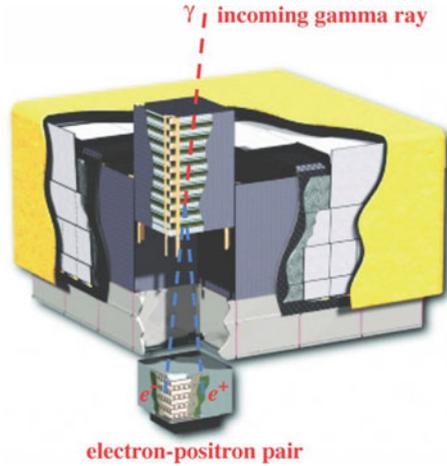
Table 1 Performance characteristics of the Fermi LAT [10]

Parameter	Value or range
Energy range	20 MeV–300 GeV
Effective area at normal incidence (60° off-axis)	
100 MeV	3700 (700) cm^2
1 GeV	$\leq$8300 (2900) cm^2
10 GeV	$\leq$8400 (3100) cm^2
Energy resolution (1σ, on-axis)	
100 MeV–1 GeV	15 %–9 %
1 GeV–10 GeV	15 %–9 %
10 GeV–300 GeV	8.5 %–18 %
Single photon angular resolution (68 % containment radius)	
100 MeV	3.5°
1 GeV	0.6°
>10 GeV	<0.15°
Field of View (FoV)	2.4 sr
Timing accuracy	300 ns
Event readout time (deadtime)	26.5 μs

LAT Instrument Response

Even though two tracker towers and three calorimeter modules of the LAT were beam tested, the original science tools provided for instrument response were the Pass 6_v3 response functions based on extensive GEANT4 Monte Carlo simulations of the satellite that followed secondary particles and photons through the digitization and software filters to determine—assuming tracker azimuthal symmetry—effective

area, energy uncertainty, and PSF (or containment radius). The event reconstruction, filters, and different event classes are described in [10].

The P6_v11 corrections to the P6_v3 response functions using a few months of on-orbit data [8] showed that the earlier response functions gave smaller angles for 68 % and 95 % containment radii above ≈ 5 GeV than given by the on-orbit calibration. The PSF is best fit with a King model profile of the form

$$f_{\text{King}}(\theta, \sigma, g) = \frac{1}{2\pi\sigma^2} \left(1 - \frac{1}{g}\right)\left(1 + \frac{\theta^2}{2\sigma^2 g}\right)^{-g}, \tag{4}$$

where θ is the angle between the incident, "true" photon direction and the reconstructed direction, and g is a fitting parameter. The wings of the PSF follow a power-law rather than exponential behavior, and are well fit with the sum of two King functions.

The first-year analyses and the 1FGL, and second-year analyses leading to the 2FGL, generally employ the P6 and P7 instrument response functions, respectively. Updated instrument performance that improves on pre-flight Monte Carlo and muon calibrations by using inflight corrections to the instrument response functions are found at[7].

Utilization of Fermi-LAT to energies as low as 30 MeV is possible with the LAT Low-Energy (LLE) technique [24]. This approach was developed for transients, such as Solar flares and GRBs, where a direction is known and a short time window can be defined. During this period, less discriminating event selections are used, the analysis includes all photons with reconstructed arrival directions within 20° of the target, and the criteria for >100 keV vetoing in the ACD is relaxed.

1.3 Energy, Flux, and Luminosity

Consistent with the notation of [4], we define the dimensionless photon energy, in units of electron rest mass, as

$$\epsilon = \frac{h\nu}{m_e c^2} = \frac{E_\gamma}{m_e c^2}. \tag{5}$$

Flux density F_ν is usually reported in units of Jansky (1 Jy $= 10^{-23}$ erg/cm^2-s-Hz), so that the quantity νF_ν is an energy flux F (units of erg/cm^2-s, or Jy-Hz, noting that 10^{10} Jy-Hz $= 10^{-13}$ erg/cm^2-s). The luminosity distance d_L for a steady, isotropically emitting source is defined so that the energy flux F is related to the source luminosity L_* (erg/s) according to the Euclidean expression

---

[7] www.slac.stanford.edu/exp/glast/groups/canda/lat_Performance.htm and http://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/Cicerone_LAT_IRFs

$$F = \frac{L_*}{4\pi d_L^2}. \tag{6}$$

If $\phi(\epsilon)$ is the measured spectral photon flux (units of photons $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\epsilon^{-1}$), then $\nu F_\nu = m_e c^2 \epsilon^2 \phi(\epsilon)$. Henceforth we use the notation

$$f_\epsilon = \nu F_\nu \tag{7}$$

for the νF_ν flux. From the definitions of F and f_ϵ,

$$F = \int_0^\infty d\epsilon \, \frac{f_\epsilon}{\epsilon}. \tag{8}$$

Considering Eq. (6), the luminosity radiated by a source between measured photon energies ϵ_1 and ϵ_2, or between source frame photon energies $\epsilon_1(1+z)$ and $\epsilon_2(1+z)$, is therefore given by

$$L_*[\epsilon_1(1+z), \epsilon_2(1+z)] = 4\pi d_L^2 m_e c^2 \int_{\epsilon_1}^{\epsilon_2} d\epsilon \, \epsilon \, \phi(\epsilon) = 4\pi d_L^2 \int_{\ln \epsilon_1}^{\ln \epsilon_2} d(\ln \epsilon) f_\epsilon. \tag{9}$$

Equation (9) shows that if the νF_ν spectrum is flat with value f_ϵ^0, corresponding to a photon flux $\phi(\epsilon) \propto \epsilon^{-2}$, then the apparent power of the source over one decade of energy is $\approx (\ln 10) L_0 \approx 2.30 L_0$, where $L_0 = 4\pi d_L^2 f_\epsilon^0$.

From these relations, one obtains the mean apparent isotropic γ-ray energy release

$$\mathcal{E}_* = \frac{4\pi d_L^2 F \langle \Delta t \rangle}{(1+z)}, \tag{10}$$

for a source at redshift z that releases average energy flux F during observing timescale Δt. The apparent versus the absolute energy releases and luminosities depends on the jet structure and variability behavior. For a steady, two-sided top-hat jet with uniform emission within angle $\theta \leq \theta_j$ of the jet axis, the absolute luminosity

$$L_{abs} = f_b L_{iso}, \tag{11}$$

where the beaming factor $f_b = 2 \times 2\pi \int_{\mu_j}^1 d\mu_j / 4\pi = 1 - \mu_j$, and $\mu_j = \cos\theta_j$. For $\theta_j = 0.1$ (5.7°), $f_b \cong 1/200$, whereas if $\theta_j = 0.01$ (0.57°), $f_b \cong 1/20000 = 5 \times 10^{-5}$. What is reported is the apparent isotropic luminosity $L_{*,iso} = 4\pi d_L^2 F$, with absolute luminosity implied by arguments for the jet opening angle and beaming factor.

Variability Information

The Schwarzschild radius of a black hole of mass M is

$$R_S = \frac{2GM}{c^2} \cong 3.0 \times 10^5 \left(\frac{M}{M_\odot}\right) \text{ cm} \cong 10^{-4} \, M_9 \text{ pc}, \tag{12}$$

defining $M_9 = M/(10^9 \, M_\odot)$. Variations in the source flux by a large factor ($\gtrsim 2$) over a time scale Δt must, from causality arguments for a stationary source, originate from an emission region of size $R \lesssim c\Delta t/(1+z)$. Incoherent superpositions of emission from larger size scales and from regions that are not in causal contact would usually (though not always) wash out large-amplitude fluctuations. For high-quality data from bright flares, large amplitude variations in source flux on timescale Δt would, from this argument, imply a black-hole mass

$$M_9 \lesssim \frac{(\Delta t/10^4 \text{ s})}{1+z}. \tag{13}$$

This relation is even preserved in systems with relativistic outflows, unless specific model-dependent provisions are made on the method of wind dissipation. Variability timescales far shorter than the light-crossing time across an $\approx 10^9 M_\odot$ black hole have been measured in the TeV BL Lac objects PKS 2155-304, Mrk 501, and Mrk 421, discussed in more detail below. Another interesting fiducial is the ratio of the Eddington luminosity to the light-crossing time across a Schwarzschild black hole [13], namely

$$\frac{L_{Edd}}{t_S} = 4\pi \frac{m_p c^4}{\sigma_T} \cong 2.5 \times 10^{43} \text{ erg s}^{-2}. \tag{14}$$

We take this criterion as separating the extreme universe from the moderate universe. During intense 3C 454.3 flaring activity [9], the ratio of the apparent isotropic luminosity and the source variability timescale strongly violate this limit, making this an extreme event.

Extragalactic Background Light (EBL)

In intergalactic space, the energy density of the EBL is dominated by that of the cosmic microwave background radiation (CMBR), with present temperature $T_{CMB} = 2.72$ K and energy density $u_{CMBR} \cong 0.25$ eV/cm^3 $\cong 4 \times 10^{-13}$ erg/cm^3. The intensity of the EBL at infrared frequencies is difficult to measure directly because of foreground radiations, such as Galactic electron synchrotron radiation and zodiacal light scattered by dust in our Solar system. Photons observed at 10–100 GeV interact primarily with EBL photons at optical and UV energies. The energy density of the dust and stellar components of the EBL is $\approx 10\%$ of the CMBR energy density at the present epoch.
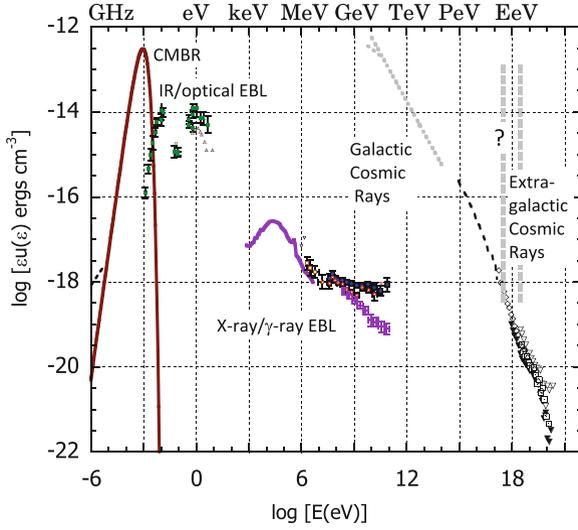
Fig. 3 Spectral energy densities in intergalactic space of various radiations, including the CMB, the infrared (IR) and optical, X-ray, γ ray, and the extragalactic cosmic ray energy density. Also shown is the energy density of cosmic rays measured near Earth; the transition energies between the modulted and unmodulated, and galactic and extragalactic components remains uncertain

Figure 3 shows the energy density of photons in intergalactic space and cosmic rays in outer space in the Solar cavity near Earth. Fermi measurements [7] give a lower intensity for the diffuse extragalactic γ-ray background (EGB) than EGRET data [26, 28], probably due to EGRET miscalibration above $\approx 5\,$GeV [27].[8] The cosmic-ray particle spectrum is modulated at low, $\sim$GeV/nucleon energies by changing Solar activity. Cosmic-ray source origin changes from a Galactic to extragalactic origin at an energy scale currently under debate. The case for an extragalactic origin of the UHECRs can be made by comparing Galactic size with the Larmor radius

$$r_L = \frac{E}{QB} = \left(\frac{A}{Z}\right) \frac{m_p c^2 \gamma}{eB} \cong 100 \frac{E/10^{20}\,\text{eV}}{Z B_{\mu G}}\,\text{kpc} \tag{16}$$

of an ion with energy E, Lorentz factor γ, atomic charge Z and atomic mass A. The characteristic magnetic field $B = 10^{-6} B_{\mu G}$ G of the Milky Way is a few μG in the disk and probably much less in the Galactic halo.

---

[8] The spectral energy density of the isotropic background radiation field at $100\,$MeV as measured with EGRET is

$$\epsilon u(\epsilon) = \frac{4\pi}{c}\,\epsilon I_\epsilon \cong \frac{4\pi}{c}\,\frac{\text{keV}}{\text{cm}^2\text{-s-sr}} \cong 10^{-18}\text{erg cm}^{-3}. \tag{15}$$

Because the ϵI_ϵ spectrum is essentially flat, the total energy density over the $100\,$MeV–$100\,$GeV energy range gives a factor $\ln 10^3 \cong 6.9$; thus the EGRET energy density of the EGB is nearly 5 orders of magnitude below the present CMBR energy density.

1.4 Limits to the Extreme Universe

The largest luminosity is obtained if the entire rest mass energy of the object with mass M is transformed into energy on a light-crossing time scale for the gravitational radius of the object, so that

$$L_{max} = \frac{Mc^2}{GM/c^3} = \frac{c^5}{G} = 3.6 \times 10^{59} \text{ erg s}^{-1} \qquad (17)$$

[23, 30]. The most luminous blazar source yet detected with the Fermi LAT is 3C 454.3, with apparent isotropic γ ray luminosity reaching $L_\gamma \approx 2 \times 10^{50}$ erg/s. The most luminous GRBs detected with the Fermi-LAT, namely GRB 080916C and GRB 090510A, reached $L_\gamma \approx 10^{53}$ erg/s.

2 Fermi Gamma-Ray Source Catalogs and Fermi Pulsars

Now we take a global view, and look at the Fermi Large Area Telescope First Source Catalog (1FGL) [43] as a *catalog*, or systematic collection of objects, before focusing on γ-ray emission from neutron stars and the first Fermi pulsar catalog [44]. The 1FGL catalog, taken with 11 months of data between 2008 August 4 and 2009 July 4, expands the 3 month bright source list (BSL, or 0FGL) [35], just as the First LAT AGN Catalog (1LAC) [45], the subject of the next lecture, expands the LBAS (LAT Bright AGN Sample) [36]. The 2FGL and 2LAC build on the 1FGL and 1LAC, respectively. The 1FGL and 2 FGL catalogs are described in this section.

Source detection significance for the 1FGL and 2FGL catalogs is based on the likelihood test statistic (TS). The likelihood ratio compares the likelihood of a null hypothesis with the likelihood of the model. According to Wijk's theorem, -2 times the logarithm of this ratio approaches χ^2 for large χ^2. The likelihood $\mathcal{L}$ is the probability of observing n_{ij} counts in pixel i, j given the expected number of counts λ_{ij} in pixel i, j. Poisson statistics gives

$$\mathcal{L} = \prod_{ij} p_{ij} = \prod_{ij} \frac{\lambda_{ij}^{n_{ij}} \exp(-\lambda_{ij})}{n_{ij}!}. \qquad (18)$$

Dropping a term independent of model, the logarithm of the likelihood is

$$\log \mathcal{L} = \sum_{ij} n_{ij} \log(\lambda_{ij}) - \sum_{ij} \lambda_{ij}, \qquad (19)$$

and $TS = -2 \log(\mathcal{L}/\mathcal{L}_0) \to \chi^2$, where $\mathcal{L}_0$ is the likelihood of the null hypothesis for the given data set. Equation (19) gives a prescription for calculating TS when

analyzing γ-ray data. Mattox and colleagues [21] originally applied this method to EGRET data.

The Fermi catalogs are based on a limiting test statistic, which is $TS > 25$ in the 1FGL, 2FGL, 1LAC and 2LAC, and $TS > 100$ in the BSL and the LBAS. This introduces systematic effects that should be taken into account when interpreting the 1FGL. For example, a cut on TS produces integral flux thresholds strongly dependent on the source spectral hardness. Because of the steeply falling Galactic background at high galactic latitudes, high-latitude hard spectrum sources will have a large TS for much smaller integral fluxes than a soft-spectrum source detected at the same significance level.

2.1 First Fermi Catalog of Gamma-Ray Sources: 1FGL

The 1FGL comprises 1451 sources with $TS > 25$, corresponding to a significance $\gtrsim 4\sigma$, and represents 21.22 Ms (or $\approx 73.3\,\%$ livetime) of data. Most of the deadtime loss is due to passage through the SAA South Atlantic Anomaly ($\approx 13\,\%$) and to readout deadtime (9.2 %). The exposure is uniform to within a factor ≈ 1.25 between north and south and the detections are based on integrated data rather than shorter bright periods or flares. An improved background and calibration model is used in the 1FGL compared to the 0FGL.

The 1FGL is not a flux-limited survey. Although the exposure is relatively uniform—an excellent feature of the GLAST design in nominal scanning mode— there are large differences in integral photon fluxes of sources depending on source location and the strong Galactic diffuse emission. Consequently, the 1LAC catalog is drawn from sources above 10° latitude where the Galactic diffuse intensity is low.[9] Any conclusions are weakened when source identifications are not complete, for example due to the limits of a counterpart catalog used to make associations, or to uncertainty in distance or lack of redshift, as this incompleteness can bias results.

The 1FGL catalog gives source location regions and association criteria defined in terms of elliptical fits to the 95 % confidence regions, and power-law spectral fits as well as flux measurements in five energy bands for each source. In addition, monthly light curves are provided. Firm identifications with sources found in other astronomical catalogs are based on correlated variability, e.g., rotation or orbital period in the case of pulsars and X-ray binaries, or morphological similarities with counterparts at other wavelengths, as in the case of SNRs.

For the catalogs and association criteria used, 630 of the sources in the 1FGL are unassociated. Due to the sensitivity of the results to the model of interstellar diffuse γ-ray emission used to model the bright foreground, 161 sources at low Galactic latitudes towards bright local interstellar clouds are flagged as having properties that are strongly dependent on the Milky Way gas model.

---

[9] The Galactic diffuse flux at $10° \leq |b| \leq 20°$, averaged over longitude, is still a factor of 2–3 greater than the extragalactic diffuse at higher latitudes [42].

The principal source classes found in the 1FGL and listed in Table 2 are

- 3 high mass X-ray Binaries
- Rotation-powered and millisecond pulsars
- Supernova remnants
- Globular clusters
- >600 blazars
- 28 Non-blazar AGNs and Radio Galaxies
- Dozens of AGNs of uncertain type
- 2 starburst Galaxies

In addition, the Sun, moon, and Earth, hundreds of GBM GRBs, and more than a dozen LAT GRBs had been detected by the time of publication of the 1FGL.

Association depends on the underlying catalog and the degree of confidence assigned in the probability of association. For blazars, four catalogs with substantial overlap were used, giving 689 1FGL associations with sources found in at least one of these catalogs. A fuller discussion of the AGN catalog is deferred to the next lecture. Though no extended radio lobes of radio galaxies were reported in the 1FGL, the lobes of Centaurus A have since been imaged [46].

Table 2 LAT 1FGL and 2FGL source classes [43, 70]

Description	Designator	Assoc. (ID) 1FGL	Assoc. (ID) 2FGL
Pulsar, X-ray or radio, identified by pulsations	psr (PSR)	7 (56)	0 (83)
Pulsar, radio quiet (LAT PSR, subset of above)	PSR	24–	25–
Pulsar wind nebula	pwn (PWN)	2 (3)	0 (3)
Supernova remnant	snr[a] (SNR)	41 (3)	62[c] (6)
Globular cluster	glc (GLC)	8 (0)	11 (0)
High Mass X-ray binary	hxb (HXB)	0 (2)	0 (4)
Micro-quasar object: X-ray binary (black hole or neutron star) with radio jet	mqo (MQO)	0 (1)	0 (1)
Nova	nov(NOV)		0 (1)
BL Lac type of blazar	bzb (BZB)	295 (0)	428 (7)
FSRQ type of blazar	bzq (BZQ)	274 (4)	353 (17)
Non-blazar active galaxy	agn (AGN)	28 (0)	10 (1)
Radio galaxy			10 (2)
Active galaxy of uncertain type	agu (AGU)	92 (0)	257 (0)
Normal galaxy	gal (GAL)	6 (0)	4 (2)
Starburst galaxy	sbg (SBG)	2 (0)	4 (0)
Seyfert galaxy	sey (SEY)		5 (1)
Unassociated		630	576 + 1[d]
Total		1478[b]	1873[e]

[a] Indicates a potential association with a SNR or PWN
[b] 779 + 630 + (69) = 1409 + (69) = 1478. Greater than 1451 because of multiple class assignments
[c] Some of the 62 sources may also be associated with PWNe
[d] 576 unassociated plus one with uncertain class
[e] 1169 + 577 + (127) = 1746 + (126) = 1873

Regarding other source classes in the 1FGL, 41 Fermi sources are associated with SNRs or non-pulsed γ-ray emission, and three are sufficiently physically extended that their morphology counts as an identification (W44 W51C, and IC 443; see Lecture 5). Fermi reported three high-mass X-ray binary systems in the 1FGL, namely LS 5039, LSI +61° 303, and Cyg X-3 (Lecture 7), and a fourth was discovered later. In the 1FGL, it was unclear whether γ rays were made by massive O stars, but since then η Carinae has been been established as a γ-ray source [49]. Ten BSL candidates out of 205 in the 0FGL, do not show up in the 1FGL, illustrating the highly variable nature of some γ-ray sources (though note that all these sources are found in the Galactic ridge, $|l| < 60°$, where background modeling is especially important). On top of this, 630 sources were unassociated in the 1FGL. The unidentified Fermi sources are clustered towards the plane of the Galaxy where diffuse background Galactic emission and higher source density makes associations more tentative, though there are still many high-latitude unidentified Fermi sources.

2.2 Second Fermi Catalog of Gamma-Ray Sources: 2FGL

The 2FGL catalog [70] was released at the time of writing. This source catalog was derived from data taken during the first 24 months of the science phase of the mission, which began on 2008 August 4. Source detection is based on the average flux over the 24-month period. The 2FGL includes source location regions and fits to model spectral forms. Also included are flux measurements in 5 energy bands and light curves on monthly intervals for each source. Twelve sources in the 2FGL are found to be spatially extended.

The 2FGL contains 1873 sources detected and characterized in the 100 MeV–100 GeV range, of which 127 are firmly identified and 1170 are reliably associated with counterparts of known or likely γ-ray source classes. Although the diffuse Galactic and isotropic models used in the 2FGL analysis are improved compared to the 1FGL catalog, caution flags for 162 sources indicate possible confusion, given the uncertainty in the underlying diffuse model. Table 2 lists the number of sources of various types in the 2FGL.

Some important improvements compared to the 1FGL catalog are:

1. The 2FGL catalog is based on data from 24 months of observations.
2. The data and Instrument Response Functions use Pass 7 event selections, rather than the Pass 6 event selections used in the 1FGL.
3. The 2FGL employs a new, higher-resolution model of the diffuse Galactic and isotropic emissions.
4. Spatially extended sources and sources with spectra other than power laws are incorporated into the analysis.
5. The source association process has been refined and expanded.

2.3 Fermi Pulsars

In the 1FGL, 56 pulsars are identified by their γ-ray pulsations, and another 7 associations are based on expectations from pulsar catalogs, e.g., $\dot{E}/d^2$ ranking. Six 1FGL sources are associated with pulsar wind nebulae (PWNe) that lack known pulsars, and 8 1FGL sources are associated with globular clusters illuminated, most likely, by the superposition of millisecond pulsar (MSP) γ-ray emissions. The number of Geminga-like pulsars that lack detectable radio emission, has grown by a large factor. There are 83 γ-ray pulsars, now all identified, in the 2FGL.

A catalog of 46 γ-ray pulsars is presented in the First LAT Pulsar Catalog, Ref. [44]. This catalog includes 16 "blind-search" pulsars discovered by searching for pulsed γ-ray emission at the position of bright LAT sources. Pulsed γ-ray emission from 24 known pulsars were discovered using ephemerides derived from monitoring radio pulsars, of which 8 are MSPs. The remaining 6 γ-ray pulsars were known previously.

EGRET Pulsars

The Crab and Vela pulsars were known prior to *CGRO*, and Geminga was known as a bright γ-ray point source. The pulsar nature of Geminga was only established by detection of X-ray pulsations in ROSAT data early in the EGRET era [64]. The EGRET pulsars, with EGRET fluxes above 100 MeV and 2FGL fluxes between 1 and 100 GeV, ordered by brightness, are listed in Table 3.

PSR 1951+32, located in the confusing Cygnus arm region, is associated with the SNR CTB 80 discovered by Kulkarni et al. [68]. It has a phase-averaged νF_ν spectrum rising $\propto \epsilon^{0.2}$ and peaking at ≈ 2 GeV, with a slower than exponential decline at higher energies [51].

Table 3 γ-ray fluxes of EGRET pulsars [17, 44, 70]

Pulsar	Period (ms) P	Age (kyr) $P/2\dot{P}$	EGRET F_{-8}	Pulsar catalog F_{-8}	2FGL F(1–100 GeV)[a]
0833-45, Vela	89.3	11.3	834.3 ± 11.2	1061 ± 7.0	135.8 ± 0.4
J0633+1746, Geminga	237	340	352.9 ± 5.7	305.3 ± 3.5	72.9 ± 0.3
0531+21,[b] Crab	33	1.25	226.2 ± 11.2	209 ± 4	18.3 ± 0.15
1706-44	102	17.6	111.2 ± 6.2	149.8 ± 4.1	19.1 ± 1.7
1055-52	197	540	33.3 ± 3.82	30.45 ± 1.7	5.0 ± 0.09
1951+32[c]	39.5	110	16 ± 2	17.6 ± 1.9	2.1 ± 0.07
1509-58,[c] Circinus	88.9	150	–	8.7 ± 1.4	1.45 ± 0.08

[a] Also in units of 10^{-8} ph cm^{-2} s^{-1}

[b] Associated with SN 1054

[c] Pulsars not reported in the 3EG [71]

The Circinus pulsar was detected with COMPTEL but not EGRET, and is unusual in having an inferred polar magnetic field $B_p = 15.77$ TG, compared to ∼TG (10^{12} G) fields for the others. Geminga is unusually close, at $d \cong 160$ pc [57], whereas the others are more likely to be at ∼kpc distances.

Elementary Pulsar Physics

Over 1900 pulsars are known today, mostly through radio surveys. Pulsars are rapidly rotating, highly magnetized neutrons star. The neutron stars themselves are formed by explosions made by the collapsing cores of massive stars (core-collapse supernovae), or through accretion-induced collapse of white dwarfs. Misalignment of the magnetic axis and rotation axis makes in the simplest configuration a rotating dipole-field geometry. Emission beamed along favorably defined field lines naturally makes a wide range of pulse structures for different obliqueness and inclination and gap sizes. The additional range of parameters associated with period and period derivative, along with poorly understand radiation physics, allows for purely empirical and kinematic pulse profile fitting.

Neutron stars are predicted to have masses near $1.4 M_\odot$ and radii ∼15 km. The two properties of the pulsar that can be very precisely measured are the period P and period derivative $\dot{P}$. Besides these observables, theory gives the mass of the neutron star, M_{NS}, and its radius R_{NS} [77]. The uncertain equation of state of nuclear matter determines whether neutron stars with given masses can exist, whether collapse to a black hole occurs, or whether other degenerate quark phases exist.

Electron degeneracy pressure in a white dwarf cannot support the gravitational force of a degenerate core exceeding the Chandrasekhar mass $M_C = 1.4 M_\odot$. Either through nuclear burning or accretion, neutron stars can be formed. Neutron stars with masses $M > M_C$ can test black-hole formation theory, but the uncertainty in orbital inclination usually makes for large uncertainties in the neutron-star mass. By means of radio timing measurements of the Shapiro delay[10] of the binary millisecond pulsar J1614-2230, a $1.97 \pm 0.04 M_\odot$ pulsar mass was recently measured [58], which rules out some exotic hyperon and boson condensate equations of state.

Noting that velocity $\boldsymbol{v} = \boldsymbol{\Omega} \times \boldsymbol{R}$, and the angular speed $\Omega = 2\pi/P$, then the light-cylinder radius at which the speed $v = c$ is

$$R_{LC} = Pc/2\pi. \tag{20}$$

Even the rotation of a simple misaligned dipole field leads to unusual geometric features, for example, the footprint on the neutron star surface of the field lines that open into the light cylinder [60].

An expression for the surface polar magnetic field can be obtained by equating the rotational spindown energy loss rate with the magnetic dipole radiation power.

---

[10] The Shapiro delay is a regular general-relativistic change in the light travel time from the pulsar as the radio photon travels through the gravitational field of the binary system; see [81].

The former is

$$-\frac{dE_{rot}}{dt} = \frac{d}{dt}\left(\frac{1}{2}I\Omega^2\right) = \frac{4\pi^2}{P^3}I\dot{P}, \tag{21}$$

with the moment of inertia I depending on the mass and equation of state of neutron star matter. This implies a characteristic age

$$\tau = P/2\dot{P}, \tag{22}$$

for P much longer than the initial spin period. The magnetic dipole power can be estimated from the Poynting flux of a dipole radiation at the light cylinder, namely,

$$-\frac{dE_{md}}{dt} = \frac{B^2(R_{LC})}{8\pi}\left(4\pi R_{LC}^2 c\right) = \frac{1}{2}B_{NS}^2\left(\frac{R_{NS}^6}{R_{LC}^4}\right)c \propto \frac{B_{NS}^2}{P^4}. \tag{23}$$

Thus

$$B_{NS} \propto \sqrt{P\dot{P}}. \tag{24}$$

The magnetic field B_{LC} at R_{LC} is therefore

$$B_{LC} = \left(\frac{24\pi^4 I\dot{P}}{c^3 P^5}\right)^{1/2}. \tag{25}$$

We can imagine a time-independent configuration for a spinning, magnetized neutron star if we wait long enough until all current flows have asymptotically relaxed to their steady-state values. This defines the condition of a force-free magnetosphere, where the vanishing of the Lorentz force $\boldsymbol{F} = Q[(\boldsymbol{v}/c) \times \boldsymbol{B} + \boldsymbol{E})] = 0$ implies the existence of a magnetosphere filled with plasma with Goldreich-Julian [62] density

$$\rho_{GJ} = -\frac{\boldsymbol{\Omega} \cdot \boldsymbol{B}}{2\pi}. \tag{26}$$

Pulsar models build on these elementary concepts. For generic *polar cap* models, charge depletion in the polar field lines that open to the light cylinder, compared with the Goldreich-Julian density, generates strong electric fields that induce vacuum breakdown and current flow. In *outer gap* models, strong fields are generated at the gaps created by the surface defined by the condition $\boldsymbol{\Omega} \cdot \boldsymbol{B} = 0$. The primary γ-ray production and attenuation processes are curvature radiation, Compton radiation, and magnetic pair production. *Slot gap* models study a more sophisticated realization of the pair-starved region and how a slot gap forms along field lines with different acceleration potentials. Besides discriminating between magnetospheric models, pulsar studies help disentangle the geometry of a pulsar. For more on pulsar models, see [56, 60, 76].

Properties of Fermi Pulsars

Two different approaches are taken to discover pulsars, either by period folding with data from a pulsar established at other frequencies, or to perform a blind search for the pulsation. In the first method, the timing parameters of pulsars already known at radio or X-ray frequencies are used to search for evidence of γ-ray pulsations in the γ-ray data. For the blind-search technique, spectral power is calculated from time-tagged γ-ray photon event data by searching through likely values of P and $\dot{P}$ values. This is extremely computationally intensity, and not feasible for MSPs found in binary systems. More computationally efficient techniques compute the arrival time differences of γ rays, which cluster at values corresponding to multiples of the pulse period [55]. For γ-ray astronomy, with limited photon statistics, only the time-differencing technique for blind searches is effective at finding pulsars.

The first blind-search detection with Fermi was a pulsar with period $P = 0.317\,\mathrm{s}$ in the shell of the SNR CTA1, coincident with unidentified X-ray and 3EG J0010+7309 γ-ray sources [34]. The spin-down power of this pulsar is $\approx 4.5 \times 10^{35}\,\mathrm{erg\,s^{-1}}$, and the pulsar's inferred age is 14000 years, which is consistent with the SNR age. Of the 46 Fermi pulsars reported in the First Fermi Pulsar Catalog, 16 of them were found using blind-search techniques in sources that are positionally coincident with unidentified EGRET sources and supernova remnants [37]. These 16 pulsars are all young and highly magnetized, with inferred magnetic fields between ≈ 1–$10\,\mathrm{TG}$.

The rotational energy-loss rates of pulsars in the First Fermi Pulsar Catalog range from $\sim 3 \times 10^{33}\,\mathrm{erg\,s^{-1}}$ to $5 \times 10^{38}\,\mathrm{erg\,s^{-1}}$, with the young, highly magnetized pulsars typically having rotational energy-loss rates exceeding $\gtrsim 10^{35}\,\mathrm{erg\,s^{-1}}$. Comparing with the phase-averaged apparent isotropic γ-ray luminosity implies efficiencies for the conversion of spin-down power to γ-ray energy in the range from $\sim 0.1\,\%$ to $\approx 100\,\%$. About 75 % of the γ-ray pulses have two peaks, but a third emission structure, P3, shows up in the Vela pulse profile and moves to later phases with increasing photon energy [41]. As Fig. 4 shows, the main peak, P1, which is dominant from at least UV to GeV energies, becomes less intense compared to P2 at multi-GeV energies.

The Fermi LAT phase-averaged spectrum of γ-ray pulsars can be well fit by a generalized exponentially cutoff power law, given by

$$N(E) \propto E^{-\Gamma_\gamma} \exp[-(E/E_c)^b], \tag{27}$$

with hard photon number indices Γ_γ generally in the range $1 \lesssim \Gamma_\gamma \lesssim 2$, and cutoff energies E_c between ≈ 1 and $5\,\mathrm{GeV}$. This form has been found to apply to all types of pulsars, whether radio or γ-ray selected, or normal or millisecond pulsars. For Vela itself, as seen in Fig. 5, $\Gamma_\gamma = 1.38 \pm 0.03$, $E_c = 1.36 \pm 0.15\,\mathrm{GeV}$, and $b = 0.69 \pm 0.02$ (quoting statistical errors only) [50]. The sub-exponential ($b < 1$) rather than super-exponential ($b > 1$) cutoff that would be produced if the γ rays were made deep within the polar cap of the neutron star magnetosphere, and the detection of emission at tens of GeV, is evidence against a polar-cap model. A phase-averaged spectrum fitted with Eq. (27) is the composite phase-resolved spectra that individually
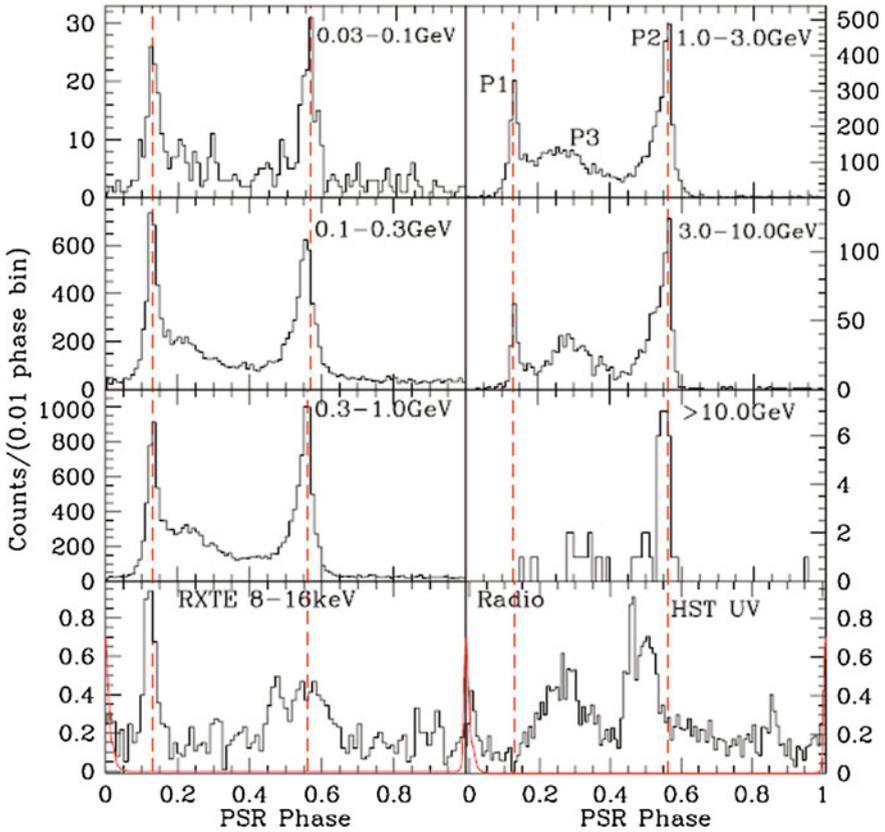
Fig. 4 Vela light curves at optical, X-ray, and γ-ray energies [41], binned to 0.01 of the pulsar phase. The main peaks P1, P2 and P3 are labeled in the *top right panel*. The *bottom left panel* shows the 8–16 keV *RXTE* light curve [65] along with the radio pulse profile (*dashed lines*). At *lower right*, the 4.1–6.5 eV *HST*/STIS NUV light curve [75] is shown

can be fit by simple exponential behaviours [50, 51]. Significant variations of cutoff energy with phase are needed to reproduce the phase-averaged spectra.

Millisecond Pulsars and Globular Clusters

Pulsed γ-ray emission was detected with the Fermi LAT from J0030+0451, making it the first firm detection of a MSP in γ rays [40], although there was a marginal EGRET detection of PSR J0218+4232 [67] that has been confirmed with the LAT [40]. Nine months into science operations, the Fermi-LAT Collaboration reported 8 γ-ray MSPs [39], establishing a new population of γ-ray emitters. As noted above, it is not computationally feasible to perform blind searches for binary MSPs, which introduces too many possibilities into the timing solutions. This makes it harder to
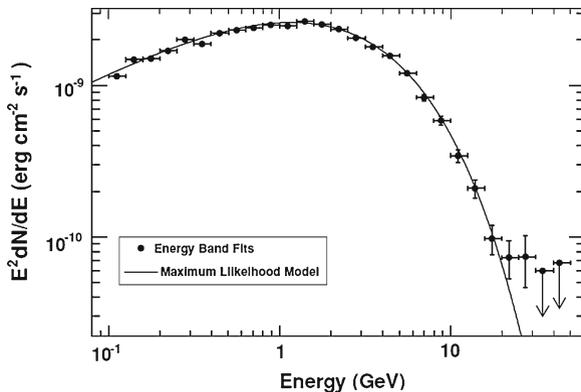
Fig. 5 Phase-averaged spectrum for $0.1 < E\,(\text{GeV}) < 60$ [50]

know what fraction of the unidentified sources are MSPs, or if there are any radio-quiet MSPs.

Globular clusters are quasi-spherical stellar systems consisting of hundreds of thousands of old (ages of $\approx 10^{10}$ years) metal-poor stars. The high-stellar density at the centers of these systems makes for unusual multi-body encounters. According to the most common scenario, binary neutron-star systems in low-mass X-ray binaries can be spin up by accretion over long time spans to periods of a few ms, provided that the polar surface magnetic field is $\lesssim 10^9$ G, which is $\sim$3 orders of magnitude smaller than the typical TG fields of young γ-ray pulsars. Even though they have very different magnetic field and rotation rates, the γ-ray spectra of MSPs and young pulsars are similar, and well-represented by a hard power-law with a modified exponential cutoff given by Eq. (27).

The Fermi-LAT collaboration has reported the detection of the famous globular cluster 47 Tucanae at a significance of 17σ [38]. The γ-ray spectral shape is consistent with a large populations of MSPs. With a typical luminosity of $\lesssim 3 \times 10^{33}$ erg s^{-1} in γ-rays, $\sim$20–40 MSPs are required to give the measured γ-ray luminosity. Using 587 days of data, the number of Fermi sources associated with globular clusters has grown to 8 [47], and now 11 in the 2FGL [70].

Both γ-ray detected "garden-variety" pulsars and MSPs have the highest values of magnetic field at the light cylinder, B_{LC}, suggesting that similar emission mechanisms operate. The γ-ray luminosity grows with spin-down energy $L_\gamma \propto \dot{E}$ at $\dot{E} \lesssim 10^{35}$ erg s^{-1} and $L_\gamma \propto \sqrt{\dot{E}}$ at $\dot{E} \gtrsim 10^{35}$ erg s^{-1}, with large scatter in this relation due, at least, to distance uncertainties (Fig. 6 in [44]).

Pulsar Wind Nebulae

Given the accuracy of timing parameters for a pulsar, the rotational energy-loss rate is known to the uncertainty of the moment of inertia of the neutron star. The spin

energy not coming out as γ rays, which usually accounts for no more than $\sim 10\,\%$ of the spin-down energy, must come out in a different form, e.g., as field energy in the form of a relativistic Poynting wind with such small baryon loading. The wind Lorentz factors attains values of $\sim 10^8$. The interaction of the outflowing MHD wind with the ISM makes a termination shock where additional particle acceleration can take place. The termination shock separates the cold upstream MHD wind with a turbulent flow downstream into the ISM. The boundary between the wind and ISM is highly structured because of various fluid instabilities. Cold particles can be directly injected into the ISM with the wind Lorentz factor, or accelerated at the termination shock.

Identification of pulsar wind nebulae (PWNe) with the Fermi LAT depends on finding steady extended GeV radiation. Three PWNe were reported with Fermi LAT using 16 months of survey data, namely the Crab nebula, Vela-X, and the PWN inside MSH 15-52 [53]. Searching in the off-pulse emission of the pulsar yields a candidate PWN related to PSR J1023-5746 and coincident with HESS J1023-575. The sources with GeV-detected PWNe have the highest spin-down flux and tend to have the highest spin-down power, with $\dot{E}_{rot} \gtrsim 10^{37}\,\mathrm{erg\,s^{-1}}$.

Crab Nebula and Flares

The Crab pulsar and nebula is associated with a supernova that took place in 1054 CE at a distance of ≈ 2 kpc. It is the best example of a center-filled plerionic SNR, where the interior is illuminated by synchrotron emission powered by the wind of a young, 33 ms neutron star. The spin-down energy of the Crab pulsar is $4.6 \times 10^{38}\,\mathrm{erg\,s^{-1}}$, of which $\approx 30\,\%$ is converted into the radiant power of the nebula. The mean spectral index of the nebula in the LAT range is $\cong -4$ between ≈ 100 and $400\,\mathrm{MeV}$, hardening to an index of $\cong -1.64$ at higher energies [48] (see Fig. 6). The harder, high energy component is the synchrotron self-Compton component (predicted by Gould [63]), and well-explained as SSC if the mean magnetic field is $\approx 200\,\mu\mathrm{G}$ [48, 54].

Year-scale variations by a several percent are found in Crab nebular fluxes measured with RXTE, Swift BAT, and INTEGRAL. The fluxes also declined during the first two years of the Fermi mission, consistent with measurements of an $\approx 7\,\%$ decrease between 12 and 500 keV as measured with the GBM [82]. In the period between 2008 August and 2009 April, no variations in the 100 MeV–30 GeV flux were found with the LAT [48].

Although some evidence for flux variations had been noted during the EGRET era [59], Fermi-LAT detected $\gtrsim 100\,\mathrm{MeV}$ flaring episodes from the Crab lasting for 16 days in February 2009 and 4 days in September 2010 [52]. The September 2010 flare was first announced by the AGILE team [78], and they also reported an earlier October 2007 flare [79]. Typical γ-ray flare luminosities are $\approx 4 \times 10^{36}\,\mathrm{erg/s}$, and reach an integral photon number flux above 100 MeV that is $\approx 10\times$ brighter than the mean flux.

The PSF of front events of the Fermi-LAT is $\approx 0.12°$–$0.15°$ at $E \gtrsim 10\,\mathrm{GeV}$, so imaging can be achieved to $\approx 0.12°/8 \sim 1'$ (at 2 kpc, $1'$ and $1''$ correspond to
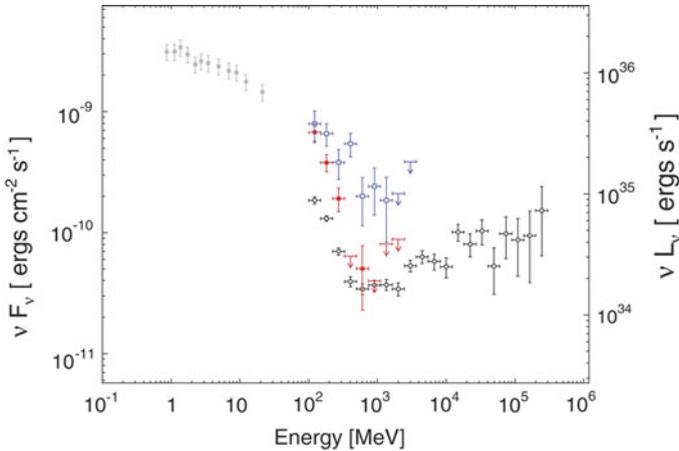
Fig. 6 COMPTEL 0.7–30 MeV data (*gray filled squares*) and long-term average Fermi-LAT 100 MeV–300 GeV γ-ray data (*open circles*), and February 2009 (*open squares*) and September 2010 (*filled squares*) flares of the Crab nebula [52]

$\approx$0.6 pc and $\approx 3 \times 10^{16}$ cm, respectively). The extent of the radio nebula is $\approx 10'$, and the extent of the Chandra X-ray torus from which perpendicular jets emerge is $\approx 1.3'$. A bright inner X-ray ring in the Chandra image is $\approx 0.5'$ in extent. Zooming into the central region with both Chandra and HST[11] reveals a dynamic wind, wisps, and $\sim 1''$ knots that brighten and dim. For comparison, one light day corresponds to $\approx 3 \times 10^{15}$ cm, far smaller than the smallest resolvable knots, yet the γ-ray flares radiate $\approx 1\%$ of the spin-down energy.

The crucial feature of the Crab spectrum may well be that the quiescent spectrum displays a strong softening at ≈ 20 MeV. This value is a factor ≈ 10 smaller than the maximum synchrotron frequency at ≈ 200 MeV obtained by balancing the timescale for electron synchrotron losses with the Larmor timescale on which particles gain energy through Fermi particle acceleration mechanisms (see Sect. 5.5), *for a non-relativistic flow*. A mildly relativistic flow can enhance the emission by large factors and could be compatible with changes in the flow profile at the termination shock when the ultra-relativistic wind slows to mildly relativistic speeds. Considerable theoretical interest has focused on perturbations of the flow induced by statistical fluctuations [83], by the generation of large amplitude MHD waves [69], and by outflowing knots formed at an oblique termination shock [66]. No evidence for the flare energy generation is found in the pulsar spin-down behavior [52].

---

[11] http://chandra.harvard.edu/photo/2002/0052/animations.html

Pulsar Physics with Fermi

Three years after launch, over 100 γ-ray pulsars are known, and γ-ray emission is associated with 8 globular clusters counting 47 Tuc [47]. In NGC 6624, a single MSP dominates the γ-ray energy output [61]. The γ-ray pulsars are divided into 31 radio-selected MSPs, 37 blind search γ-ray pulsars, and 38 young radio-selected γ-ray pulsars. Of the 37 blind search pulsars, extensive follow-up observations reveal pulsed radio emission in only 3 cases. Radio followup of Fermi pulsars is described, e.g., in [72].

Some of the interesting open questions relate to the field geometry and radiation mechanisms for pulsar γ rays. The significant fraction of Geminga-like blind-search pulsars is generally interpreted in terms of a larger γ-ray than radio cone. The detection of pulsed emission in the Crab to $\approx$125 GeV [80] means that this emission has to be made high in the magnetosphere. Emission to these energies requires extreme parameters for a curvature radiation origin, and may imply Compton-scattering in the Crab's magnetosphere. (The curvature radiation mechanism remains a feasible mechanism for pulsars older than the Crab.) The $L_\gamma \propto \dot{E}^a$ behaviour is not well-explained from the lowest MSP powers to the highest pulsar powers. Pulsar γ-ray emission can make a significant, tens of %, fraction of the Galactic diffuse emission. This fraction could be very different in early-type galaxies with only MSPs now active.

For reviews of Fermi pulsars, see [73] and [74].

3 Fermi AGN Catalogs

The discovery of the γ-ray blazar class is one of EGRET's lasting legacies. Of the 66 high-confidence sources associated with AGNs in the 3EG catalog [17], all except one—the nearby radio galaxy Centaurus A—were associated with blazars.

Blazars are sources that exhibit violent optical variability (e.g., changing in flux by $\sim$50 % in one day), high optical polarization (exceeding several percent), and flat radio spectra with radio spectral index $\alpha_r < 0.5$ at GHz frequencies ($F_\nu \propto \nu^{-\alpha_r}$). Superluminal motion at radio frequencies and highly luminous and variable γ-ray emissions are also typical blazar properties. Blazars themselves are interpreted to be relativistic jet sources powered by a supermassive black hole, like radio galaxies, though with the observer looking nearly along the jet axis. The variety of multi-wavelength SEDs displayed by blazars and their misaligned populations can often, though not exclusively, be attributed to orientation effects amplified by the Doppler boosting of the jetted radiation (Sect. 4).

3.1 LAT Bright AGN Sample (LBAS) and First LAT AGN Catalog (1LAC)

Three lists of AGNs detected with Fermi have now been published by the Fermi Collaboration. These are the LAT Bright AGN Sample (LBAS) [36], and the First and Second LAT AGN Catalogs, 1LAC [45] and 2LAC [92], respectively.

The LBAS is based on 3 months of science observations taking place between 2008 August 4 and 2008 October 30, and consists of 106 high Galactic latitude ($|b| >$ 10°) sources associated with AGNs. These sources have a test statistic $TS > 100$, corresponding to $\gtrsim 10\sigma$ significance, and are a subset of the 205 sources listed in the BSL [35].

By comparison, the 3EG [17] and the EGR [11] list 31 sources with significance $> 10\sigma$, of which 10 are at high latitude. Remarkably, 5 of the $> 10\sigma$ EGRET sources are not found in the BSL. These are the flaring blazars NRAO 190, NRAO 530, 1611+343, 1406-076 and 1622-297, the latter of which is the most luminous blazar detected with EGRET [108].

The 1LAC [45] is a subset of the 1451 sources in the 1FGL [43] derived from analysis of the first 11 months of LAT science data. There are 1043 1FGL sources at high latitudes, of which 671 are associated with 709 AGNs, with the larger number of AGNs than 1LAC sources due to multiple associations. Associations are made by comparing the localization contours with counterparts in various source catalogs, for example, the flat-spectrum 8.4 GHz CRATES (Combined Radio All-Sky Targeted Eight GHz Survey; [104]) and the Roma BZCAT blazar catalog [109]. The probability of association is calculated by comparing the likelihood of chance associations with catalog sources if randomly distributed.

Of the 671 associations, 663 are considered "high-confidence" associations due to more secure positional coincidences. The "clean" sample is a subset of the high-confidence associations consisting of 599 AGNs with no multiple associations or other analysis flags, for example, evidence for extended emission. As listed in Table 4, these subdivide into 275 BL Lac objects, 248 flat spectrum radio quasars, 26 other AGNs, and 50 AGNs of unknown types. The "New Classes" category contains non-blazar AGNs, including starburst galaxies and various types of radio galaxies, e.g., narrow line and broad line. An AGN is classified as an "unknown" type either because it lacks an optical spectrum, or the optical spectrum has insufficient statistics to determine if it is a BL Lac objects or a flat spectrum radio quasar (FSRQ). In comparison with the 671 AGNs in the 1LAC, EGRET found 66 high-confidence ($> 5\sigma$) and another 27 lower-confidence detections with significance between 4σ and 5σ, as noted earlier. Thus the 1LAC already represents an order-of-magnitude increase in the number of AGNs over EGRET. There are ≈ 300 unassociated and therefore unidentified high-latitude Fermi sources in the 1LAC.

Table 4 Classes of γ-ray emitting AGNs and galaxies in the 1LAC and 2LAC "clean" samples

Class	Number in 1LAC (2LAC)	Characteristics	Prominent members
All	599 (885)		
BL Lac objects	275 (395)	Weak emission lines	AO 0235+164
…LSP	64 (61)	$\nu_{pk}^{syn} < 10^{14}$ Hz	BL Lacertae
…ISP	44 (81)	10^{14} Hz $< \nu_{pk}^{syn} < 10^{15}$ Hz	3C 66A, W Comae
…HSP	114 (160)	$\nu_{pk}^{syn} > 10^{15}$ Hz	PKS 2155-304, Mrk 501
FSRQs	248 (310)	Strong emission lines	3C 279, 3C 354.3
…LSP	171 (221)		PKS 1510-089
…ISP	1 (3)		
…HSP	1 (0)		
New Classes[a]	26 (24)		
…Starburst	3 (2)	Active star formation	M82, NGC 253
…MAGN	7 (8)	Steep radio spectrum AGNs	M87, Cen A, NGC 6251
…RL-NLS1s	4 (4)	Strong FeII, narrow permitted lines	PMN J0948+0022
…NLRGs	4 (–)[c]	Narrow line radio galaxy	4C +15.05
…other sources[b]	9 (11)		
Unknown	50 (156)		

[a]Total adds to 27, because the RL-NLS1 source PMN J0948+0022 is also classified as FSRQ in the 1LAC
[b]Includes PKS 0336-177, BZU J0645+6024, B3 0920+416, CRATES J1203+6031, CRATES J1640+1144, CGRaBS J1647+4950, B2 1722+40, 3C 407, and 4C +04.77 in 1LAC clean sample
[c]Class designation deprecated in 2LAC

3.2 Classification of Radio-Emitting AGNs and Unification

Different classes of extragalactic AGNs are defined according to observing frequency. We have already noted the association of Fermi sources with BL Lac objects and FSRQs, which are based on an optical classification. The precise definition used by the Fermi team is that an AGN is a BL Lac object if the equivalent width of the strongest optical emission line is <5 Å, and the optical spectrum shows a Ca II H/K break ratio <0.4 in order to ensure that the radiation is predominantly nonthermal (the Ca II break arises from old stars in elliptical galaxies). The wavelength coverage of the spectrum must satisfy $(\lambda_{max} - \lambda_{min})/\lambda_{max} > 1.7$ in order that at least one strong emission line would have been detected if present. This helps guard against biasing the classification for AGNs at different redshifts where the emission lines could be redshifted out of the relevant wavelength range. For sources exhibiting BL Lac or FSRQ characteristics at different times, the criterion adopted is that if the optical spectrum conforms to BL Lac properties at any time, then it is classified as a BL object.

The criterion for classification of radio galaxies according to their radio properties stems from the remarkable correlation between radio morphology and radio luminosity [101]. The twin-jet morphology of radio galaxies is seen in low-power

radio galaxies, whereas the lobe and edge-brightened morphology is found in high-power radio galaxies, with a dividing line at $\approx 2 \times 10^{25}$ W/Hz-sr at 178 MHz, or at a radio luminosity of $\approx 2 \times 10^{41}$ erg s^{-1} for the current cosmology. Besides a radio-morphology/radio-power classification, radio spectral hardness can also be used to characterize sources as flat-spectrum and steep-spectrum sources. Furthermore, radio galaxies can be subdivided according to the widths of the optical emission lines into broad- and narrow-line radio galaxies. Correlations between radio-core dominance and γ-ray luminosity supports a scenario where the jet γ-ray flux is greatest along the jet direction [107].

Blazars and radio galaxies can also be classified according to their broadband SED when there is sufficient multiwavelength coverage to reconstruct a spectrum from the radio through the optical and X-ray bands and identify a νF_ν peak frequency $\nu_{pk}^{\rm syn}$ of the lower energy, nonthermal synchrotron component of the spectrum (see Sect. 4). When the peak frequency $\nu_{pk}^{\rm syn}$ of the synchrotron component of the spectrum is $< 10^{14}$ Hz, then a source is called low synchrotron-peaked (LSP), whereas if the SED has $\nu_{pk}^{\rm syn} > 10^{15}$ Hz, then it is referred to as high synchrotron-peaked (HSP). Intermediate synchrotron-peaked (ISP) objects have 10^{14} Hz $< \nu_{pk}^{\rm syn} < 10^{15}$ Hz. SEDs of the bright Fermi LBAS sources are constructed in [86]. Essentially all FSRQs are LSP blazars, whereas BL Lac objects have large numbers in all LSP, ISP, and HSP subclasses.

According to the standard unification scenario for radio-loud AGNs [117], radio galaxies are misaligned blazars, and FR1 and FR2 radio galaxies are the parent populations of BL Lac objects and FSRQs, respectively. To establish this relationship requires a census of the various classes of sources that takes into account the different beaming properties for the Doppler-boosted radiation of blazars. Even if analysis of data of radio galaxies and blazars supports the unification hypothesis, this paradigm still does not explain the reasons for the differences between radio-quiet and radio-loud AGNs, or between BL Lac objects and FSRQs.

New classes of extragalactic Fermi sources found in the 1LAC include starburst galaxies (Sect. 6.1), narrow line radio galaxies (NLRGs), radio-loud narrow-line Seyfert 1s (RL-NLS1s), and radio-quiet AGNs. Five NLRGs are reported in the 1LAC. These objects have narrow emission lines in their optical spectrum, suggesting that they are observed at large angles with respect to the jet direction, with the surrounding dust torus obscuring the broad line region (BLR).

RL-NLS1s have also been recently established as a γ-ray source class [84]. These objects show narrow Hβ lines with FWHM line widths $\lesssim 1500$ km s^{-1}, weak forbidden lines ($[OIII]/H\beta < 3$) and an Fe II bump [111]. By comparison with the $\sim 10^9 M_\odot$ black holes in blazars, the host galaxies of RL-NLS1s are spirals that have nuclear black holes with relatively small ($\sim 10^6$–$10^8 M_\odot$) mass that accrete at a high Eddington ratio. The detection of these objects challenges scenarios (e.g., [114]) where radio-loud AGNs are hosted by elliptical galaxies that form as a consequence of galaxy mergers.

The 1LAC includes 10 associations with radio-quiet AGNs. In 8 of these cases, at least one blazar, radio galaxy, or CRATES source is also found close to the γ-ray

source. In the remaining two cases, the association probabilities are weak. Thus none appear in the 1LAC "clean" sample. In some of these candidate radio-quiet γ-ray sources, such as the Sy 2 galaxies NGC 4945 or NGC 1068 [106], which are also starburst galaxies, the γ rays could be made by cosmic-ray processes rather than from a radio-quiet Sy nucleus. More extensive searches for γ rays from Swift-BAT AGNs has not established, however, that radio-quiet AGNS are GeV γ-ray emitters [94].

3.3 Properties of Fermi AGNs

Various correlations are found by comparing γ-ray properties of Fermi AGNs according to their radio, optical, or SED classification. Probably the most pronounced correlation is between the >100 MeV γ-ray spectral index Γ_γ and optical AGN type. FSRQs have significantly softer spectra than BL Lac objects, with $\langle \Gamma_\gamma \rangle \cong 2.40 \pm 0.17$ for FSRQs and $\langle \Gamma_\gamma \rangle \cong 1.99 \pm 0.22$ for BL Lac objects in the LBAS [36]. The SED classification shows that the mean γ-ray spectral index $\langle \Gamma_\gamma \rangle \cong 2.48, 2.28, 2.13$, and 1.96 when the class varies from FSRQs to LSP-BL Lacs, ISP-BL Lacs, and HSP-BL Lacs, respectively. The progressive hardening from FSRQs to BL Lac objects can be seen in Fig. 7, which also compares with values for radio galaxies and star-forming galaxies [89].

Fermi data reveal complex γ-ray blazar spectra. FSRQs and LSP-BL Lac objects, and most ISP blazars with sufficiently good statistics, show breaks in the ≈ 1–10 GeV range [88]. This was already apparent from the first observations of the bright blazar 3C 454.3 [84], to be discussed in more detail below. The HSP blazars, though, are generally well-described by a flat or rising νF_ν SED in the GeV range, with νF_ν peak frequencies between ≈ 100 GeV–TeV energies implied by VHE data.

Only 121 out of 291 BL Lac objects had measured redshifts at the time of publication of the 1LAC. For sources with measured redshift, BL Lac objects are mostly found at low redshift, $z \lesssim 0.4$, with only a few HSP BL Lac objects at higher redshifts. By contrast, the FSRQs span a wide range from $z \approx 0.2$ to the highest redshift 1LAC blazar with $z = 3.10$.

This significant redshift incompleteness hampers interpretation of AGN properties, in particular, Fig. 7, which can only display sources with known redshifts. For these sources, the hard-spectrum BL Lac objects typically have much lower L_γ than the FSRQs. This divide has been interpreted as a change in the accretion regime at ≈ 1 % of the Eddington luminosity [100]. In addition, the nearby radio galaxies with $z \lesssim 0.1$ inhabit a separate portion of the Γ_γ versus L_γ plane, and are characterized by lower γ-ray luminosities than their parent populations. The two more distant steep spectrum radio sources, 3C 207 ($z = 0.681$) and 3C 380 ($z = 0.692$), and the FR2 radio galaxy PKS 0943-76 ($z = 0.27$) fall, however, within the range of γ-ray luminosities measured from FSRQs. Indeed, steep spectrum radio sources are thought to be slightly misaligned FSRQs.
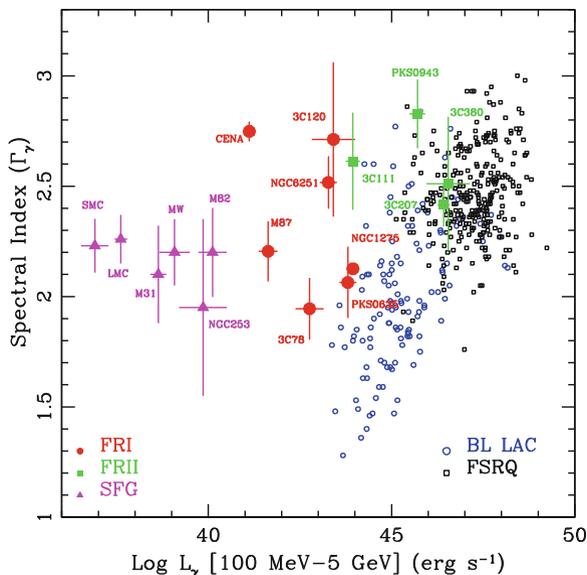
Fig. 7 Gamma-ray spectral slope Γ_γ of BL Lac objects (*open blue circles*), FSRQs (*open black squares*), FR1 radio galaxies (*red circles*), FR2 radio sources (*green squares*), and star-forming galaxies (*magenta diamonds*), are plotted as a function of their 100 MeV–5 GeV γ-ray luminosity L_γ

3C 454.3 and FSRQs

The FSRQ 3C 454.3, at $z = 0.859$, underwent giant flares and became the brightest γ-ray source in the sky for a week in 2009 December and 2010 April [91], and again in 2010 November [9]. The outburst in 2010 April triggered a pointed-mode observation by Fermi. During the December outburst, its daily flux reached $F_{-8} = 2200(\pm100)$, corresponding to an apparent isotropic luminosity of $L_{iso} \approx 3 \times 10^{49}$ erg s^{-1}, making it the most luminous blazar yet detected with Fermi. In its 2010 November outburst, it reached $F_{-8} \cong 6000$ and $L_{iso} \approx 10^{50}$ erg s^{-1}, becoming ≈ 5 times brighter than the Vela pulsar.

Figure 8 shows the light curve of 3C 454.3 [9], which can also be found at the public website of Fermi monitored sources.[12] The fluxes are plotted in durations of one day over the course of the Fermi mission. The flux rises to a plateau level preceding γ-ray outbursts, with the 2008 July outburst showing strong resemblance to those in 2009 August and December, and 2010 April and November. Intense flaring occurs during periods of enhanced activity, which is to say that the flares are not isolated events, but seem to occur during periods of enhanced accretion activity.

As noted already in the initial Fermi report [84], the spectrum of 3C 454.3 breaks strongly by 1.2(±0.3) units at $E_{br} \approx 2$ GeV. Such a break is inconsistent with

---

[12] http://fermi.gsfc.nasa.gov/ssc/data/access/lat/msl_lc/
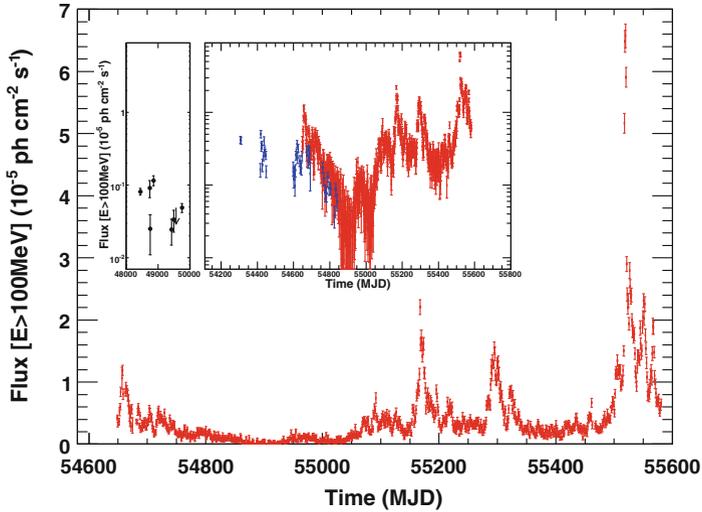
Fig. 8 Fermi LAT daily light curve of 3C 454.3 [9], showing the giant flares in December 2009, April 2010, and November 2010 (MJD 55200 corresponds to 4 January 2010). *Inset* shows the light curve on a logarithmic scale, with *black points* from EGRET [17], and *blue points* from AGILE [115]

simple radiative cooling scenarios, which predict a break by 0.5 units. The more recent analyses of data from the major outbursts of 3C 454.3 [9, 91] confirm the strong spectral break and finds that E_{br} is very weakly dependent on the flux state, even when the flux changes by more than an order of magnitude (see Fig. 9). No consistent pattern expected in acceleration and cooling scenarios [105] is found in the spectral index/flux plane.

The origin of the spectral break in 3C 454.3 bears on several important issues in FSRQs: the location of the γ-ray emission site; the source of soft target photons in Compton-scattering models; and the relation of FSRQs and BL Lac objects in view of the disappearance of such breaks in ISP and HSP blazars. Such a break would be readily understood if the target field was sufficiently strong to attenuate the blazar radiation by $\gamma\gamma$ absorption processes, but the intense Ly α radiation field at 10.2 eV observed with GALEX [98] would make a break at $E_{br} \gtrsim 30$ GeV [113]. Models employing photon attenuation deep within the BLR by He II recombination and Ly α photons with $E > 54.4$ eV [112] have been proposed to explain this break, as have nonthermal electron scattering models with accretion-disk and BLR photon targets [103]. The spectral break could also be due to Klein-Nishina effects in scattering, as has been proposed to explain the SED of PKS 1510-089 [87]. The KN break due to upscattered Ly α radiation occurs at a few GeV, and the observed break energy is insensitive to the Doppler factor. This scattering problem is treated in Sect. 4.
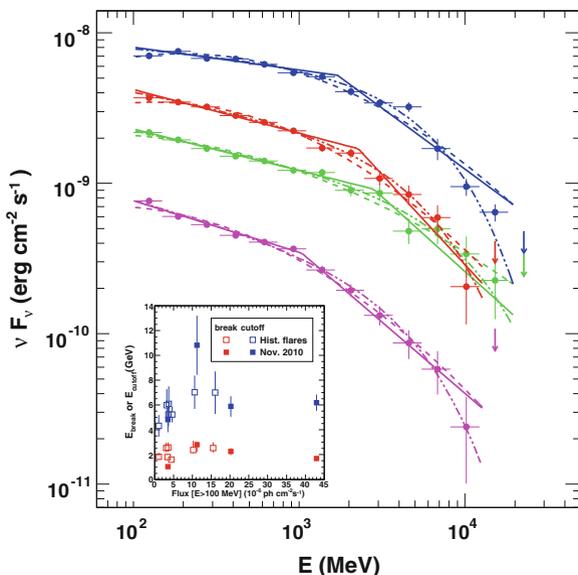
Fig. 9 Fermi LAT νF_ν SEDs for four different time periods corresponding to preflare (*magenta* lower spectrum), plateau (*green* third from top), flare (*blue* top spectrum), and post-flare (*red* second from top) periods, along with the fitted functional forms, including a broken power law (*solid*), log-parabola (*dashed*), and power-law with exponential cutoff (*dashed-triple-dotted*) functions [9]. The *inset* displays E_{break} (*red*) and E_{cutoff} energies (*blue*) as a function of flux for the different periods (*filled symbols*) and for historical flares (*open symbols*)

PKS 2155-304, Mrk 501, and BL Lac Objects

PKS 2155-304, an X-ray selected BL Lac object at $z = 0.116$ and an EGRET source [118], is one of the most prominent representatives of the HSP blazar population. During a period of extraordinary flaring on 2006 July 28, PKS 2155-304 exhibited a succession of γ-ray flares varying on time scales as short as ≈ 5 min with apparent luminosities $> 10^{46}$ erg/s—larger still when EBL absorption is included [95]. This is an extreme, hypervariable source in the sense of Eq. (14).

An 11 day campaign between 2008 August 25 and 2008 September 6 was organized early in the Fermi mission to measure the SED at optical (ATOM; Automatic Telescope for Optical Monitoring), X-ray (RXTE and Swift), and γ-ray (Fermi and HESS) frequencies. It turned out to be in a low state, well fit by a standard one-zone synchrotron/SSC model, with Doppler factor $\delta_D = 32$, magnetic field $B' = 0.018$ G, and comoving size $R' = 1.5 \times 10^{17}$ cm (corresponding to a variability time of 2 days; [96]).[13] When flaring, however, the one-zone synchrotron SSC model for PKS 2155-304 completely fails, or at least requires $\delta_D \gtrsim 100$ [102].

---

[13] The reader is assumed to be familiar with the synchrotron/SSC model; see, e.g., [4, 102, 99, 116].
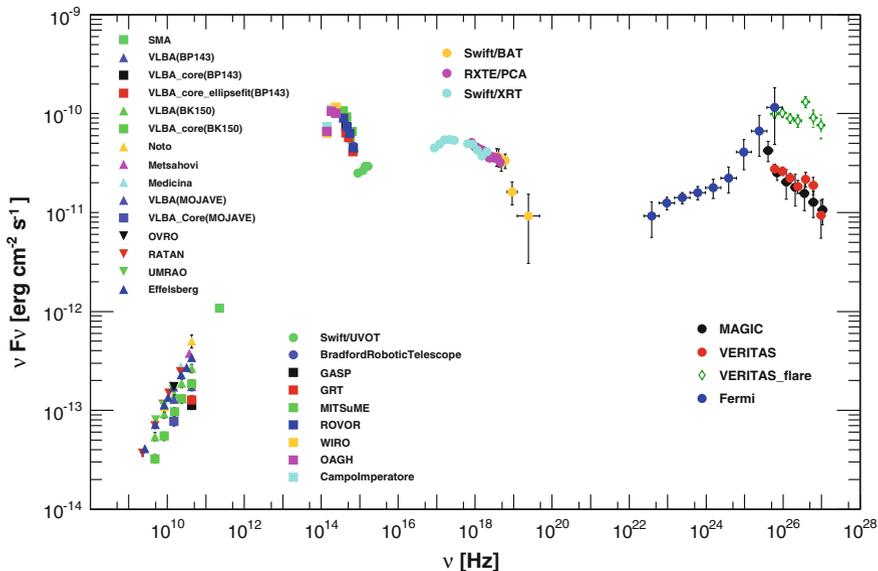
Fig. 10 SED for Mrk 501 averaged over all observations taken during the multifrequency campaign performed between 2009 March 15 (MJD 54905) and 2009 August 1 (MJD 55044) [90]. The legend reports the correspondence between the instruments and the measured fluxes

One-zone synchrotron/SSC models with $\delta_D \gtrsim 10$ give good fits to the long-term average spectra of other HSP BL Lac objects such as Mrk 421 and Mrk 501. The spectacular multiwavelength SED of Mrk 421 shown in Fig. 10 represents one of the most detailed multiwavelength blazar SEDs yet assembled [90]. The optical and X-ray data have been corrected for Galactic extinction, but the host galaxy, which is clearly visible at IR/optical frequencies, has not been subtracted. The MAGIC and VERITAS data have been corrected for EBL absorption (Sect. 7.3). A single-zone synchrotron/SSC model with $12 \lesssim \delta_D \lesssim 22$, $B' \approx 15$–30 mG, and $R' \approx 8$–50 lt-day, gives a good fit to the nonflaring SED shown in Fig. 10. By comparison, radio galaxies are fit by synchrotron/SSC models with Doppler factors of order unity. We return to this point in Sect. 7.

3.4 Second LAT AGN Catalog (2LAC)

The 2LAC [92] is based on the first two years of scientific data taken with the Fermi Gamma ray Space Telescope. It contains 1016 $|b| > 10°$ sources that are associated at high confidence with AGNs (1319 of the 1873 1FGL sources are at $|b| > 10°$). The 2LAC clean sample comprises 885 sources, consisting of 395 BL Lac objects, 310 FSRQs, 156 blazars of unknown type, 8 misaligned AGNs,

4 RL-NLSy1 galaxies, 10 AGNs of other types, and 2 starburst galaxies (NGC 4945 has fallen out of the list); see Table 4. The photon index distribution of the blazars of unknown type suggests that they comprise roughly equal numbers of BL Lacs and FSRQs. Of the 395 BL Lac objects, 220 ($\sim$55 %) lack measured redshifts, and this fraction is roughly the same for LSP, ISP, and HSP BL Lac objects.

As shown in the 2LAC, threshold sensitivity in terms of photon flux is strongly dependent on source spectral index, whereas energy flux is not, and detection of FSRQs and BL Lac objects is complete to an energy flux of $\approx 5 \times 10^{-12}$ erg/cm^2-s. Because more and deeper surveys have taken place in the northern hemisphere, the smaller fraction of southern-hemisphere BL Lac objects indicates that some 60 unassociated sources at negative declination are expected to be BL Lac objects.

Besides validating the results of the 1LAC with additional data and improved analysis techniques, the 2LAC shows that the mean fractional variability of FSRQs is greater than that of BL Lac objects. The duty cycle, as defined by the fraction of time when the average flux exceeds 1.5 standard deviations above the mean flux, is $\approx$5–10 % for bright blazars. The number of BL Lacs and FSRQ have increased by 44 % and 35 % from the 1LAC to 2LAC clean samples, respectively, with the future increases in the number of FSRQs expected to be even more modest. This is due to cosmological effects, as reflected in the flattening in the log N-log S distribution, and K-corrections that pushes the GeV peak of the νF_ν spectra to lower energies where the Fermi-LAT effective area rapidly declines and sensitivity to FSRQs gets worse.

Blazars make up $\gtrsim$97 % of the extragalactic γ-ray sources detected with the Fermi-LAT. The number of misaligned sources has stalled, with the same number of radio galaxies—eleven—in the 2LAC as in the misaligned AGN paper. But these are not the same objects. The radio galaxies 3C 78, 3C 111, and 3C 120 are not found in the 2LAC, evidently because a variable jetted emission component that contributed to the γ-ray emission in the past has gone quiet. Now found in the 2LAC are the FRI radio galaxies Fornax A and Cen B, and the head-tail radio galaxies IC 310 [110], the latter of which is also a MAGIC source [97]. The radio/γ-ray connection [93] and GeV–TeV synergy [85] resulting from the Fermi-LAT for AGN studies, let alone Galactic sources, cannot be adequately reviewed here; see [92] and [70].

4 Relativistic Jet Physics

Here we develop and apply relativistic jet radiation physics [4] to some puzzles arising from the Fermi observations.

4.1 GeV Spectral Break in LSP Blazars

As already described in Sect. 3.3, an important new result in blazar physics not anticipated before the launch of Fermi is the prevalence of a spectral softening in the γ-ray spectra of low synchrotron-peaked (LSP) blazars, including both FSRQs and LSP BL Lac objects. The spectral softening occurs generally between 1 and 10 GeV [36, 45]. For 3C 454.3 [91, 119], the break is $\gtrsim 1$ unit, and the energy of the break is rather insensitive to flux (Fig. 9). A softer spectral break, possibly consistent with radiative cooling, is found at multi-GeV energies in the Fermi/MAGIC spectrum of FSRQ 4C +21.35 [175]. If a GeV spectral softening and significant GeV flux variability accompanied by modest spectral variability is typical of blazars, then some physical process should be able to explain this behaviour.

γ Rays from External Compton Scattering

We treat the problem that blazar γ radiation arises from the scattering of a quasi-isotropic target photon field [172] in the Klein-Nishina regime. Compton scattering takes place in the Thomson limit when $4\gamma\epsilon' \ll 1$ (Eq. (6.123) in [4]), where γ and ϵ' are the electron Lorentz factor and photon energy in the comoving frame.

Consider an external isotropic, monochromatic photon field with photon energy $\epsilon_\star$ (in $m_e c^2$ units). The average energy in the comoving frame is $\langle \epsilon' \rangle \cong \Gamma \epsilon_\star$. The upscattered photon energy in the Thomson regime is $\epsilon'_C \cong (4/3)\gamma^2 \epsilon'$. Hence $\epsilon' \epsilon'_C = (4/3)\gamma^2 \epsilon'^2 = (4\gamma\epsilon')^2/12$. Thus scattering is in the Thomson regime provided $12\epsilon'\epsilon'_C \lesssim 1$, or $\epsilon_\star \lesssim (\delta_D/\Gamma)/[12\epsilon_C(1+z)]$, implying $E_C(\text{GeV}) \lesssim 12/[E_\star(\text{eV})]$ for $\delta_D \cong \Gamma$. Away from the endpoints, therefore, the scattered photon energy marking the transition to the KN regime is very weakly dependent on Doppler factor in highly beamed relativistic jets.

The spectrum of an isotropic nonthermal power-law electron distribution, when transformed to the stationary frame, remains a power-law with the same index as in the comoving frame, but with a strong angular dependence. The endpoints of the distribution are the low-and high-energy electron Lorentz cutoffs, γ'_1 and γ'_2 respectively, boosted by δ_D [147]. Consequently, the shape of the scattered spectrum cannot depend on δ_D provided that $\gamma_1 = \delta_D \gamma'_1 \ll \gamma_{KN} \cong (4\epsilon_*)^{-1} \ll \delta_D \gamma'_2 = \gamma_2$.

If the break energy observed in 3C 454.3 at ≈ 2 GeV is due to the transition to scattering in the KN regime, then the underlying target photon energy $E_* \approx 6$ eV. This is close to the energy of the Ly α photon at 10.2 eV, or Hβ at 2.55 eV. If the origin of the spectral break is due to scattering of nearly mono-energetic line radiation, then this would (1) place the scattering site within the BLR, and (2) explain the near constancy of the spectral break, independent of flux state.

We use the method of Georganopoulos et al. (2001) [147], applied to isotropic radiation fields in the stationary frame of the supermassive black hole and BLR. The

differential Compton-scattered spectrum when isotropic, monoenergetic electrons Compton upscatter isotropic, monochromatic target photons, is given for the full Compton cross section in the head-on approximation by

$$df_\epsilon^C = \frac{3}{4}\frac{c\sigma_T u_0}{d_L^2}\left(\frac{\epsilon_s}{\epsilon_*}\right)^2 \frac{N_e(\gamma, \Omega_s)d\gamma}{\gamma^2} F_C(q, \Gamma_e)H\left[\frac{\gamma}{1 + (1/\Gamma_e)} - \epsilon_s\right] \quad (28)$$

(Eq. (6.129), [4]), with

$$N_e(\gamma, \Omega_s) = \delta_D^3 \frac{N'(\gamma')}{4\pi}, \gamma' = \gamma/\delta_D \quad (29)$$

(Eq. (6.124)),

$$\epsilon_s = (1 + z)\epsilon \equiv \epsilon_z, \quad \Omega_s = \Omega_*, \quad (30)$$

$$F_C(q, \Gamma_e) = 2q\ln q + (1 + 2q)(1 - q) + \frac{1}{2}\frac{(\Gamma_e q)^2}{(1 + \Gamma_e q)}(1 - q), \quad (31)$$

$$q \equiv \frac{\epsilon_s/\gamma}{\Gamma_e(1 - \epsilon_s/\gamma)}, \quad \Gamma_e \equiv 4\gamma\epsilon_*, \quad (32)$$

and q is restricted to the range $(4\gamma^2)^{-1} \le q \le 1$ [134, 154] (Eqs. (6.74)–(6.76), (6.125)–(6.127), [4]). Restriction to the Thomson regime occurs when $\Gamma_e \ll 1$ or $4\gamma\epsilon_* \ll 1$. The final term in Eq. (31) dominates for scattering in the KN regime. (A simpler form for analytic work is the isotropic Thomson cross section,

$$F_{T,iso}(\hat{\epsilon}) = \frac{2}{3}\left(1 - \hat{\epsilon}\right) \quad (33)$$

(Eq. (6.71), [4]), with $\hat{\epsilon} = \epsilon_s/4\gamma^2\epsilon_\star$ (Eq. (6.70)), which assumes isotropic scattering in the electron rest frame.)

Thus an isotropically distributed nonthermal electron distribution in the comoving frame gives a Compton-scattered spectrum

$$f_\epsilon^{C,iso} = \frac{3}{4}\frac{c\sigma_T\epsilon_s^2}{4\pi d_L^2}\delta_D^3 \int_0^\infty d\epsilon_* \frac{u_*(\epsilon_*)}{\epsilon_*^2}\int_{\gamma_{min}}^\infty d\gamma \frac{N_e'(\gamma/\delta_D)}{\gamma^2} F_C(q, \Gamma_e), \quad (34)$$

(Eq. (6.123); considering only upscattering), with

$$\gamma_{min} = \frac{\epsilon_s}{2}\left(1 + \sqrt{1 + \frac{1}{\epsilon_s\epsilon_*}}\right) \quad (35)$$
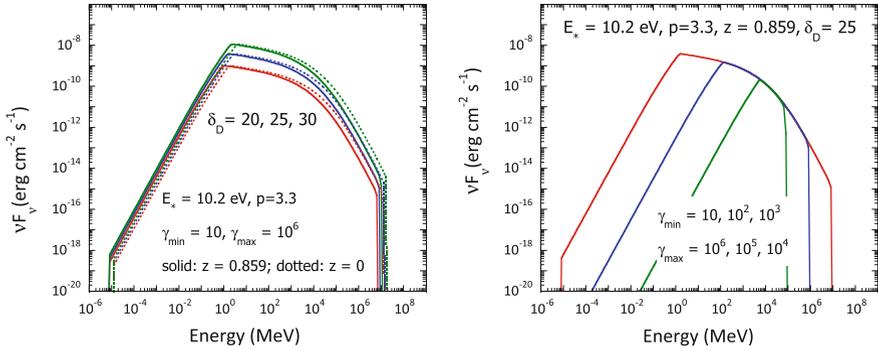
Fig. 11 Spectrum of Compton-scattered line radiation. *Left* Dependence on Doppler factor δ_D for sources at different redshifts. *Right* Dependence on endpoints. Within the endpoints, the spectral shape of the function is independent of δ_D

(Eq. (6.90)). Substituting Eq. (29) for an isotropic power-law electron spectrum in the comoving frame, so that $N'(\gamma') = N_{eo}\gamma'^{-p}H(\gamma'; \gamma'_1, \gamma'_2)$,[14] scattering an external monochromatic line spectrum $u_*(\epsilon_*) = u_*\delta(\epsilon_* - \epsilon_{*o})$ gives

$$f_\epsilon^{C,iso} = \frac{3}{4}\frac{c\sigma_T u_*}{4\pi d_L^2}\left(\frac{\epsilon_s}{\epsilon_*}\right)^2 \delta_D^{3+p} N_{eo} \int_{\min[\gamma_{min},\gamma_1]}^{\gamma_2} d\gamma\,\gamma^{-(p+2)}\,F_C(q,\Gamma_e). \quad (36)$$

This is numerically solved to give the results shown in Fig. 11, which were compared in [91] with the rapidly falling spectrum of 3C 454.3.

Figure 12 shows a model where the γ-ray continuum observed from 3C 454.3 with the Fermi LAT [91] are formed by power-law nonthermal jet electrons, with number index $p = 3.17$, that Compton-scatter Ly α line photons. The model is insensitive to values of lower and upper comoving electron Lorentz factors γ_{min} and γ_{max} provided $\gamma_{min} \lesssim 10^2$ and $\gamma_{max} \gtrsim 10^4$. The Klein-Nishina softening from a power-law electron distribution gives a poor fit to the data, but if one assumes that the underlying electron distribution is curved, as in the case of a log parabola function [160], for example, it may be possible to obtain a good fit to the 3C 454.3 data for this model.

For integrations over blackbody spectra, substitute

$$u_{*,bb}(\epsilon_*) = \frac{2m_ec^2}{\lambda_C^3}\frac{\epsilon_*^3}{\exp(\epsilon_*/\Theta) - 1} \quad (37)$$

into Eq. (34). The dimensionless temperature of the radiation field is $\Theta = k_BT/m_ec^2$. The CMBR can be approximated as an isotropic, monochromatic radiation field

---

[14] The Heaviside functions $H(x - a)$ and $H(x; a, b)$ are defined such that $H(x - a) = 1$ if $x > a$ and $H(x - a) = 0$ otherwise, and $H(x; a, b) = 1$ if $a \le x \le b$ and $H(x; a, b) = 0$ otherwise.

Fig. 12 Model for the
γ-ray spectrum of 3C 454.3
when isotropic power-law
electrons with index $p = 3.17$
in a relativistic jet Compton
scatter Ly α photons [91]
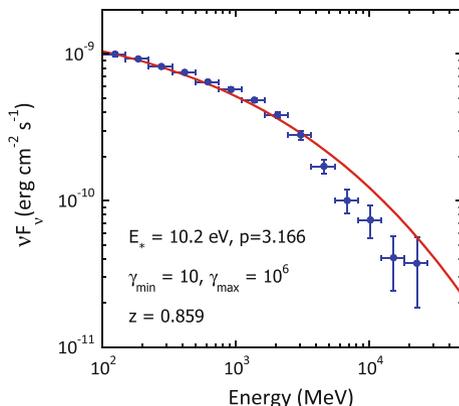
$u_*(\epsilon_*) = u_0 \delta(\epsilon_* - \epsilon_0)$, with $u_0 = 4 \times 10^{-13}(1 + z)^4$ ergs/cm^3 and $\epsilon_0 = 2.70\Theta_{CMBR} \cong 1.24 \times 10^{-9}(1+z)$. The condition $q < 1$ implies $\epsilon_s < \gamma/(1+\Gamma_e^{-1})$.

Compton Emissivity and Electron Energy-Loss Rate

In the calculation of time-dependent blazar spectra with evolving electron distributions, or in pair cascade calculations, it is necessary to consider the full Compton cross section in scattering and electron-energy loss calculations. When $\gamma \gg 1$, the head-on approximation, where the incident photon is assumed to be directed opposite ("head-on" to) the direction of the electron when transformed to the electron rest frame, can be employed.

For general treatments, the time-dependent emissivity $\epsilon_s j_C(\epsilon_s, \Omega_s; \boldsymbol{r}, t)$ is calculated. For external isotropic photon spectra, the angle-dependent emissivity when jetted nonthermal, isostropic electrons with electron spectrum $n_e(\gamma, \Omega_e)$ Compton-scatter photons of an isotropic monochromatic photon field is given, in the head-on approximation, by the expression

$$\epsilon_s j_C(\epsilon_s, \Omega_s) = \frac{3}{4}c\sigma_T u_0 \left(\frac{\epsilon_s}{\epsilon_\star}\right)^2 \int_1^\infty d\gamma \, \frac{n_e(\gamma, \Omega_e)}{\gamma^2} F_C(q, \Gamma_e) H\left(q; \frac{1}{4\gamma^2}, 1\right) \quad (38)$$

(Eq. (6.74)). Here the electron spectrum is written in the stationary frame where the radiation field is isotropic. The νL_ν Compton luminosity for a single electron is given by

$$\epsilon_s L_C(\epsilon_s) = \frac{3}{4}\frac{c\sigma_T u_0}{\gamma^2}\left(\frac{\epsilon_s}{\epsilon_\star}\right)^2 F_C(q, \Gamma_e) \, H\left(\epsilon_s; \frac{\epsilon_\star}{1+\epsilon_\star/\gamma}, \frac{\gamma}{1+\Gamma_e^{-1}}\right). \quad (39)$$

The electron energy loss rate by Compton scattering is given by

$$
-\dot{\gamma}_C = \int_0^\infty d\epsilon_s \, L_C(\epsilon_s) = \frac{3}{4} \frac{c\sigma_T u_0}{m_e c^2 \gamma^2 \epsilon_\star^2} \int_{\epsilon_\star}^{\gamma/(1+\Gamma_e^{-1})} d\epsilon_s \, \epsilon_s \, F_C(q, \Gamma_e)
$$

$$
= \frac{12 c\sigma_T u_0 \gamma^2}{m_e c^2} \int_0^1 dq \, \frac{q}{(1+\Gamma_e q)^3} \, F_C(q, \Gamma_e). \tag{40}
$$

Using the isotropic Thomson kernel, Eq. (33) becomes in the Thomson regime, $\Gamma_e \ll 1$,

$$
\epsilon_s L_C(\epsilon_s) \cong \frac{c\sigma_T u_0}{2\gamma^2} \left(\frac{\epsilon_s}{\epsilon_\star}\right)^2 \left[1 + \frac{3}{4}\frac{x^2}{(1-x)}\right] \left[1 - \frac{x}{\Gamma_e(1-x)}\right] H\left(\epsilon_s; \epsilon_\star, \frac{\gamma}{1+\Gamma_e^{-1}}\right). \tag{41}
$$

Here $x = \epsilon_s/\gamma$ so that $q = x/[\Gamma_e(1-x)]$. To lowest order,

$$
-\dot{\gamma}_T \cong \frac{c\sigma_T u_0}{2\epsilon^2} \int_0^{\Gamma_e} dx \, x \left(1 - \frac{x}{\Gamma_e}\right) = \frac{c\sigma_T u_0 \Gamma_e^2}{12\epsilon^2} = \frac{4}{3} c\sigma_T u_0 \gamma^2, \tag{42}
$$

recovering the familiar result in the limit $\gamma \gg 1$, $\epsilon_\star/\gamma \ll 1$. Expanding the energy-loss rate expression, again using the isotropic Thomson kernel, gives the expansion

$$
-\dot{\gamma}_{C,i=1} \rightarrow \frac{4}{3} c\sigma_T u_0 \gamma^2 \left(1 - \frac{3}{2}\Gamma_e + \frac{163}{40}\Gamma_e^2 + \mathcal{O}[\Gamma_e^3]\right) \tag{43}
$$

correct to $\mathcal{O}(\Gamma_e^2)$.

4.2 Leptonic Jet Models

A photon flux of $F_{-8} = 2200$ from 3C 454.3 at $z = 0.859$ implies an apparent $>100\,\text{MeV}$ isotropic γ-ray luminosity of $L_\gamma = 3.3 \times 10^{49} \, \text{erg s}^{-1}$ [91]. The flux reached a factor ≈ 6 larger in the November 2010 outburst [9], corresponding to an apparent luminosity of $\approx 2 \times 10^{50}$ erg/s, which is the record-holder for all blazars, including PKS 1622-297 from the EGRET era [108].

Using the measured flux and a one-day variability timescale at the time that the most energetic photon with $E \approx 20\,\text{GeV}$ was detected implies a minimum Doppler factor of $\delta_{D,\text{min}} \approx 13$, as we show. Assuming that the outflow Lorentz factor $\Gamma \approx 20$, consistent with the inferred value of $\delta_{D,\text{min}}$ and with radio observations at a different epoch [155], then simple arguments suggests a location $R \lesssim c\Gamma^2 t_{\text{var}}/(1+z) \approx 0.2(\Gamma/20)^2(t_{\text{var}}/\text{day})$ pc, which is at the outer boundary of the BLR. Flux variability on time scales as short as 3 h was measured at another bright flux state [9], which suggests that the γ-ray emission site would be even deeper in the BLR. This stands in contrast to inferences based on coherent optical polarization changes over timescales

of weeks prior to a γ-ray flare in 3C 279 [121], which places the emission site at much larger distances. The MAGIC detection of VHE emission from PKS 1222+21 (4C +21.35) [128] is even stronger evidence for γ-ray production at the pc scale or farther from the central engine.

Jet Doppler Factor

Emission from bulk magnetized plasma in relativistic motion provides the best explanation for the large apparent isotropic luminosities L_{iso}, energy releases $\mathcal{E}_{iso}$, and short variability times t_{var} from cosmological γ-ray sources. The comoving emission-region size scale R' is related to t_{var} through the expression $R' \lesssim c\Gamma t_{var}/(1+z)$. Combined Fermi LAT and GBM observations give the best GRB data from which to determine the minimum bulk outflow Lorentz factor Γ_{min} through $\gamma\gamma$ opacity arguments. For blazars, multiwavelength X-ray measurements combined with Fermi-LAT data are important for inferring the minimum Doppler factor.

It is simple to derive Γ_{min} in a blast-wave formulation, noting that the internal photon energy density

$$u'_\gamma \approx \frac{4\pi d_L^2 \Phi}{\Gamma^2 4\pi R^2 c} \approx \frac{(1+z)^2 d_L^2 \Phi}{\Gamma^6 c^3 t_{var}^2}, \tag{44}$$

using $R \approx \Gamma^2 c t_{var}/(1+z)$. The optical depth for $\gamma\gamma \to e^+ e^-$ processes is $\tau_{\gamma\gamma}(\epsilon'_1) \cong R'\sigma_T(\epsilon'_1/2)u'_\gamma(2/\epsilon'_1)/(m_e c^2)$, where $R' = R/\Gamma$ and $\epsilon' = 2/\epsilon'_1$ from the threshold condition. The condition $\tau_{\gamma\gamma}(\epsilon'_1) < 1$ with the relation $\Gamma\epsilon'_1/(1+z) = \epsilon_1$ implies

$$\Gamma \gtrsim \Gamma_{min} \cong \left[\frac{\sigma_T d_L^2 (1+z)^2 f_{\hat\epsilon}\epsilon_1}{4t_{var}m_e c^4} \right]^{1/6}, \quad \hat\epsilon = \frac{2\Gamma^2}{(1+z)^2\epsilon_1} \tag{45}$$

[124]. Here f_ϵ is the νF_ν flux at photon energy $m_e c^2 \epsilon$, which is evaluated at $\epsilon = \hat\epsilon$ due to the peaking of the $\gamma\gamma$ cross section near threshold.

The minimum Doppler factor δ_{min} ($\delta_D = [\Gamma(1 - \beta\cos\theta)]^{-1}$) defined by the condition $\tau_{\gamma\gamma}(\epsilon_1) = 1$, can be estimated to $\approx 10\%$ accuracy compared to results of more detailed numerical calculations through the expression

$$\delta_{min} \cong \left[\frac{\sigma_T d_L^2 (1+z)^2 f_{\hat\epsilon}\epsilon_1}{4t_v m_e c^4} \right]^{1/6}, \tag{46}$$

where f_ϵ is the νF_ν spectrum measured at frequency $\nu = m_e c^2 \epsilon/h$, t_v is the variability time, and $E_1 = m_e c^2 \epsilon_1$ is the maximum photon energy during the period in which f_ϵ is measured. The νF_ν flux f_ϵ in Eq. (46) is evaluated at $\epsilon = \hat\epsilon = 2\delta_D^2/(1+z)^2\epsilon_1$ from the pair-production threshold condition.

For 3C 454.3, writing $f_\epsilon = 10^{-10} f_{-10}$ erg cm^{-2} s^{-1} in Eq. (46) gives $\delta_{min} \approx$ 16.5 $\left[f_{-10} E_1 (10 \, \text{GeV}) / t_v (6 \, \text{h}) \right]^{1/6}$ and $\hat{E}(\text{keV}) \cong 3.4 (\delta_{min}/15)^2 / E_1 (10 \, \text{GeV})$. Contemporaneous Swift XRT observations give the νF_ν flux between 0.2 and 10 keV. From Swift public data [98], $\nu F_\nu (4 \, \text{keV}) \approx 6 \times 10^{-11}$ erg/cm^2-s, so that $\delta_{min} \approx 15$. Because $\delta_D > \delta_{min}$, we write $\delta = 20 \delta_{20}$ and $\delta_{20} \approx 1$. This value compares favorably with $\delta_D = 24.6 \pm 4.5$, bulk Lorentz factor $\Gamma = 15.6 \pm 2.2$, and observing angle $\theta(°) = 0.8 \pm 0.2$ obtained from superluminal observations [155], derived from measurements made at a different time.

Variability Time Scale

In the case of 3C 454.3, the emission region for one of its major flares is constrained in a colliding shell geometry to be at distance $R \lesssim 2c\Gamma^2 t_v / (1 + z) \approx 0.1 \Gamma_{20}^2 t_v (6 \, \text{h})$ pc, assuming $\Gamma_{20} \equiv \Gamma / 20 \cong 1$. This is within the BLR [176]. Ways to avoid this conclusion are to assume that $\Gamma \gg 20$, which would be associated with large jet powers, or for only a small portion of the emission region to be active, which would lower the jet's radiative efficiency given that only a small fraction of the emitting surface is radiating. This might be implausible in view of the already large apparent isotropic γ-ray luminosity of 3C 454.3. Recollimation shocks (e.g., [135]) at the pc scale might reduce the characteristic size of the emission region, though the contrast ($c\delta t_v / [(1 + z) R_{pc} (\text{pc})] \sim 2 \times 10^{-3} t_v \text{pc} / R_{pc}$) between the size of the emission region and location $R = R_{pc}$ pc seems unexpectedly small even for a narrow jet. Alternately, flaring episodes with short variability times might take place within the BLR, whereas the more slowly varying emissions could be radiated by jet plasma at larger distances, but then we would expect large spectral variations due to the different target photon sources.

The uses of variability studies to infer properties of blazars, GRBs, or other sources is an ongoing concern, because flow properties are inferred from the variability timescale. The auto-correlation function can be used to infer a characteristic variability timescale, as can breaks in the power density spectrum of sources [122]. Other techniques for variability analysis of the extensive compilation of blazar light curves are currently in development.

Equipartition Field and Jet Power

Here we perform an equipartition power analysis for the giant outbursts from 3C 454.3 [9, 91]. In the blob framework, where a spherical emitting region with radius R' in the comoving frame moves with Lorentz factor Γ at an angle θ to the line of sight, the absolute jet power is [136, 148] (Sect. 5.5, [4]), including the photon power,

$$L_j = 2\pi r'^2 \beta c \frac{\Gamma^2}{\delta_D^2} \frac{(\delta_D B_{eq})^2}{8\pi} \left(\chi^2 + \frac{4}{3\chi^{3/2}} \right) + 4\pi d_L^2 \frac{\Gamma^2}{2\delta_D^4} \Phi. \qquad (47)$$

Here $\Phi = 10^{-9}\Phi_{-9}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$ is the measured bolometric photon energy flux ($4\pi d_L^2\Phi = 3.6 \times 10^{48}\,\mathrm{erg\,s^{-1}}$ for $\Phi_{-9} = 1$; $\Phi_{-9} \cong 10$ and 50 during the December 2009 and November 2010 outbursts, respectively), and the factor $\Gamma^2/2\delta_D^4$ in the last term of this expression recovers the increased energy requirements due to debeaming when $\theta \gg 1/\Gamma$, or $\delta_D \ll \Gamma$, in the limit that the jet opening angle is $\ll 1$. The factor χ is the deviation from equipartition, so that $\delta_D B = \chi(\delta_D B_{eq})$ and

$$\delta_D B_{eq} = B_{cr}\left[\frac{9\pi d_L^2 f_{\epsilon_{pk}}m_e c^2(1+\zeta_{pe})\ln(\epsilon_2/\epsilon_1)}{4\sqrt{(1+z)\epsilon_{pk}}U_{cr}^2 c\sigma_T V_b'}\right]^{2/7} \tag{48}$$

(Eq. (7.80), [4]). Here $f_{\epsilon_{pk}}$ is the νF_ν peak of the synchrotron spectrum, $U_{cr} = B_{cr}^2/8\pi$, $B_{cr} = 4.414 \times 10^{13}$ G, and ζ_{pe} is the relative energy in protons to nonthermal electrons which emit synchrotron radiation with a spectrum $\alpha = 0.5$ between observed photon energies ϵ_1 and ϵ_2.

Letting $(1 + \zeta_{pe})\ln(\epsilon_2/\epsilon_1) = 10^2\Lambda_2$, the peak synchrotron frequency $\nu_{pk} = 10^{13}\nu_{13}$ Hz, and $f_{\epsilon_{pk}} = 10^{-10}f_{-10}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$, substitution of parameters for 3C 454.3 gives

$$B_{eq}(\mathrm{G}) = 3.25\frac{(f_{-10}\Lambda_2)^{2/7}}{\delta_{20}^{13/7}\nu_{13}^{1/7}t_v^{6/7}(\mathrm{days})} \tag{49}$$

i.e., an equipartition field of a few G. The corresponding jet power is

$$L_j(\mathrm{erg\,s^{-1}}) \cong \frac{\Gamma^2}{\delta_D^2}\left[2.5 \times 10^{46}\frac{(f_{-10}\Lambda_2)^{2/7}\delta_{20}^{2/7}t_v^{2/7}(\mathrm{days})}{\nu_{13}^{1/7}}\left(\chi^2 + \frac{4}{3\chi^{3/2}}\right)\right. \tag{50}$$
$$\left.+ 5 \times 10^{45}\frac{\Phi_{-9}}{\delta_{20}^2}\right].$$

The Eddington luminosity for a $10^9 M_9$ Solar mass black hole is $1.26 \times 10^{47}\,\mathrm{erg\,s^{-1}}$. Estimates for the black-hole mass in 3C 454.3 range from $M_9 \approx 0.5$ [98] to $M_9 \approx 4$ [150]. Thus we see that the jet power from 3C 454.3 would be super-Eddington if χ departs from its equipartition value by more than a factor ≈ 4. Moreover, the system cannot be severely debeamed, though this is already unlikely from other arguments, e.g., core dominance parameter and superluminal motion observations [155].

By taking the ratio of the synchrotron and Compton νF_ν peaks, using the relations $\epsilon_{pk,syn} \cong (3/2)\gamma_{pk}^2(B/B_{cr})\delta/(1+z)$ and $\epsilon_{pk,C} \cong (4/3)\gamma_{pk}^2\Gamma\epsilon_\star\delta/(1+z)$, we have $\epsilon_\star \cong (\epsilon_{pk,C}/\Gamma\epsilon_{pk,syn})(B/B_{cr})$ or

$$E_\star(\mathrm{eV}) \cong 4.6\frac{E_{pk,C}(100\,\mathrm{MeV})B(3\,\mathrm{G})}{\nu_{13}\Gamma_{20}}. \tag{51}$$

This suggests that the soft photon energy scattered to make the GeV emission is a few eV, which would correspond to either atomic line radiation or scattered Shakura-Sunyaev accretion-disk radiation.

The energy density of the external radiation field can also be estimated from the 3C 454.3 observations. Let A_C represent the ratio of the energy fluxes in the Compton and synchrotron components. If $u_\star$ denotes the energy density of the target radiation field in the stationary frame, and u_B denotes the comoving magnetic-field energy density, then for $\Gamma \gg 1$, $4u_\star \Gamma^2/3 \approx A_C U'_B$, or $u_\star (\mathrm{erg\,cm}^{-3}) \cong 0.007(A_C/10)B^2(3\,\mathrm{G})/\Gamma_{20}^2$. This can be compared to energy densities in the BLR using relations between the BLR radius and line luminosity [98]. The implied energy densities are an order of magnitude larger, suggesting that either $B > 3\,\mathrm{G}$ or $\Gamma < 20$.

4.3 Hadronic Jet Models

Measurements of air showers induced by cosmic rays impacting the atmosphere give the best evidence for the existence of powerful accelerators of UHECRs with $E \gtrsim 10^{18}\,\mathrm{eV}$, with Larmor radii so large that they almost certainly originate from beyond the Galaxy.

Adiabatic Losses, and Photopair and Photopion Losses on the CMBR

The radial scale factor

$$R = \frac{R_0}{1+z} \tag{52}$$

at redshift z. In an adiabatic expansion process, $dQ = 0 = dU + p\,dV$ and the energy content $U = uV$, where u is the energy density and V is the volume. Thus $u\,dV + p\,dV = -V\,du$, or $(u+p)dV = -V\,du$, so

$$-\frac{dV}{V} = \frac{dx}{x} + \frac{dy}{y} + \frac{dz}{z} = 3\frac{dR}{R} = \frac{du}{u+p}. \tag{53}$$

For a relativistic gas with $p = u/3$,

$$-\frac{3dR}{R} = \frac{du}{\left(\frac{4u}{3}\right)} = -3d\ln R = \frac{3d\ln u}{4}, \tag{54}$$

which implies $u \propto R^{-4}$. Because the energy density $U = uV \propto R^{-1}$, $\gamma \propto R^{-1}$, so $-\dot{\gamma} \propto -(1/R^2)dR/dt_\star$ (see Sect. 9.4 in [4]).

With R given by Eq. (52),

$$- \dot{\gamma} = (1+z)^{-1} \left| \frac{dz}{dt_*} \right| \gamma = H_0 \sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}, \tag{55}$$

for a flat cold dark matter cosmology with cosmological constant Λ. Thus

$$- \frac{d \ln \gamma}{dt} = \frac{1}{(1+z)} \left| \frac{dz}{dt_*} \right| = H_0, \tag{56}$$

The mean-free path for energy losses due to photopair production is given in Sect. 9.3.2 of [4] by the expression

$$r_{\phi e}(E_{20}) = \frac{cE}{|dE/dt|} = \frac{c}{|d \ln \gamma / dt|}$$

$$\cong \frac{1.0[E_{18}(1+z)]^{1/2}}{(1+z)^3 \mathcal{F}_{\phi e}[E_{18}(1+z)]} \, \text{Gpc} \equiv \frac{\sqrt{y}}{(1+z)^3 \mathcal{F}_{\phi e}(y)} \, \text{Gpc}, \tag{57}$$

where

$$\mathcal{F}_{\phi e}(y) = 0.74 + 1.78 \ln(y/2) + \frac{0.69 \ln(y/2)}{(y/2)^{3/2}}, \tag{58}$$

taking the asymptote $y \gg 1$ or $E_{18} \gg (1+z)^{-1}$. If the ankle in the spectrum is a consequence of photopair losses of ultra-high energy protons, as advocated by Berezinskii and colleagues [132], then the energy of the ankle would reflect the redshift epoch of greatest source production of the UHECRs, noting that $r_{\phi e} \propto (1+z)^{-3}$ at the maximum energy loss of protons due to photopair losses on the CMB.

Approximately, then

$$\left[\frac{d \ln E}{dt} \right]_{\phi e} = - \frac{(1+z)^3 [0.74 + 1.78 \ln(y/2)]}{\sqrt{y} \, (\text{Gpc/c})}, \tag{59}$$

when $y = E_{18}(1+z) \gg 1$. The conversion $(\text{Gpc/c}) = 1.03 \times 10^{17} \, \text{s} \cong 3.2 \times 10^9$ years.

The mean-free-path for a proton to lose energy through photopion losses on the CMBR is given by Ref. [4], Sect. 9.2.4, by

$$r_{\phi\pi}(E_{20}) = \frac{c}{|d \ln E / dt|_{\phi\pi}} \cong \frac{13.7 \exp[4.0/E_{20}(1+z)]}{(1+z)^3 [1 + 4.0/E_{20}(1+z)]} \, \text{Mpc}, \tag{60}$$

which, though derived in the limit $E_{20} \ll 4/(1+z)$, gives a good approximation even at higher energies.

Photopion Production Efficiency

From the relations between time elements in the stationary frame (starred), the comoving frame (primed), and the observer frame (unscripted), $R = \beta c t_* = \beta c \Gamma t'$, so $t' \cong R/\beta \Gamma c$ for relativistic flows, and $t \cong (1+z)R/\beta \Gamma^2 c$, since $t \cong (1+z)t'/\Gamma$. The comoving dynamical time scale is therefore given by $t'_{dyn} = R'/c = R/\Gamma c$.

The energy flux

$$\Phi = \frac{d\mathcal{E}}{dAdt} = \frac{L_*}{4\pi d_L^2}, \quad L_* = \frac{d\mathcal{E}_*}{dt_*} = \Gamma^2 \frac{d\mathcal{E}'}{dt'}, \tag{61}$$

noting $\Gamma \mathcal{E}' = \mathcal{E}_*$ and $dt_* = dt'/\Gamma$. Therefore $\epsilon L(\epsilon) = 4\pi d_L^2 f_\epsilon = \Gamma^2 \epsilon' L(\epsilon')$, and

$$n(\epsilon) = \frac{\epsilon L(\epsilon)}{4\pi R^2 c m_e c^2 \epsilon}, \text{ so } n'(\epsilon') = \frac{\epsilon' L'(\epsilon')}{4\pi \Gamma^2 R^2 m_e c^3 \epsilon}. \tag{62}$$

The causality constraint for rapidly variable emissions from a relativistic blast wave is

$$\Delta r' \cong \frac{\Delta r}{\Gamma} \lesssim \frac{\Delta t'}{c} = \frac{\Delta t_*}{\Gamma c} \cong \frac{\Gamma \Delta t}{(1+z)c}, \tag{63}$$

and $t_{var} \cong (1+z)R/\Gamma^2 c$.

We write the target νF_ν photon spectrum as

$$f_\epsilon = f_{\epsilon_{pk}} \left[\left(\frac{\epsilon}{\epsilon_{pk}} \right)^a H(\epsilon_{pk} - \epsilon) + \left(\frac{\epsilon}{\epsilon_{pk}} \right)^b H(\epsilon - \epsilon_{pk}) \right]. \tag{64}$$

The threshold for photopion ($\phi\pi$) processes is $\gamma'_p \epsilon' \gtrsim m_\pi/m_e \cong 400$, using the approximation of [131] that the product of the photopion inelasticity and cross section is $K_{p\gamma}\sigma_{\phi\pi} = 70\,\mu b$ above comoving photon energy $\epsilon' \cong 400$. Because $\gamma'_p = E_p/\Gamma m_p c^2$ and $\epsilon' = (1+z)\epsilon/\Gamma$, the threshold *escaping* proton energy is

$$E_p^{thr} \cong \frac{400 m_p c^2 \Gamma^2}{(1+z)\epsilon} \text{ and } E_p^{pk} \cong \frac{400 m_p c^2 \Gamma^2}{(1+z)\epsilon_{pk}} \cong \frac{3.7 \times 10^{17} \Gamma_3^2}{(1+z)\epsilon_{pk}} \text{ eV} \tag{65}$$

where

$$\Gamma_3 \equiv \frac{\Gamma}{1000}.$$

The timescale for energy loss due to photohadronic processes is

$$t'^{-1}_{p\gamma}(E_p) \cong c(K_{p\gamma}\sigma_{p\gamma}) \int_{\epsilon'_{thr}}^{\infty} d\epsilon' \, n'(\epsilon'). \tag{66}$$

Thus

$$t_{p\gamma}^{\prime -1}(E_p) \cong \frac{c(K_{p\gamma}\sigma_{p\gamma})d_L^2}{R^2 m_e c^3 \Gamma^2} \int_{\epsilon'_{thr}}^{\infty} d\epsilon' \frac{f_\epsilon}{\epsilon'^2}. \tag{67}$$

Using Eq. (64) gives

$$t_{p\gamma}^{\prime -1}\left(E_p^{pk}\right) \cong \frac{(K_{p\gamma}\sigma_{p\gamma})d_L^2 f_{\epsilon pk}}{R^2 m_e c^2 \Gamma \epsilon_{pk}(1-b)(1+z)}, \tag{68}$$

The photopion production efficiency

$$\eta_{p\gamma}(E_p) = \frac{t'_{dyn}}{t'_{p\gamma}(E_p)}. \tag{69}$$

At $E_p = E_p^{pk}$,

$$\eta_{p\gamma}^{pk} = \eta_{p\gamma}\left(E_p^{pk}\right) = \frac{(K_{p\gamma}\sigma_{p\gamma})d_L^2 f_{\epsilon pk}}{\Gamma^4 m_e c^4 t_{var}\epsilon_{pk}(1-b)}. \tag{70}$$

Thus

$$\eta_{p\gamma}(E_p) \cong \eta_{p\gamma}^{pk} \begin{cases} \left(\frac{E_p}{E_p^{pk}}\right)^{1-b}, & E_p < E_p^{pk} \\ \begin{cases} \left(\frac{E_p}{E_p^{pk}}\right)^{1-a}, & a < 1 \\ 1, & a > 1 \end{cases}, & E_p^{pk} < E_p \end{cases}. \tag{71}$$

To illustrate these results, we apply them to Fermi observations of GRBs. For GRB090510 at $z = 0.903$ and $d_L \cong 1.8 \times 10^{28}$ cm, $E_p^{pk} \cong 2 \times 10^{17}(\Gamma_3^2/\epsilon_{pk})$ eV and $\eta_{p\gamma}^{pk} \cong 0.03 f_{-5}/\Gamma_3^4 t_{-2}\epsilon_{pk}(1-b)$, and is thus at the $\sim$1–10 % level. For GRB 080916C at $z = 4.35$ and $d_L \cong 1.25 \times 10^{29}$ cm [120], $E_p^{pk} \cong 7 \times 10^{16}(\Gamma_3^2/\epsilon_{pk})$ eV and $\eta_{p\gamma}^{pk} \cong 0.0015 f_{-6}/\Gamma_3^4 t_{var}(\text{s})\epsilon_{pk}(1-b)$.

For GRB 090926A at $z = 2.1062$ and $d_L = 16.54$ Gpc, the estimate above gives $E_p^{pk} \cong 10^{17}(\Gamma_3^2/\epsilon_{pk})$ eV and $\eta_{p\gamma}^{pk} \cong 0.01(f_{-6}/6)/\Gamma_3^4 \hat{t}_{var}\epsilon_{pk}(1-b)$. The spectral parameters to derive the photopion efficiency are given for the Band function in Table 5 [125]. Here the variability time scale t_{var} is scaled to $0.15\hat{t}_{var}$ s, which is the characteristic FWHM time of the well-resolved pulse. One percent efficiency

Table 5 Band function fits to GRB 090926A

Time interval	f_{-6} (10^{-6} erg cm^{-2} s^{-1})	a	b	ϵ_{pk}
(a): 0.0–3.3 s	3.5	1.58 ± 0.03	$-0.64^{+0.07}_{-0.09}$	0.66
(b): 3.3–9.7 s	6	1.45	-0.46	0.56
(c): 9.7–10.5 s	5	1.41	-1.69	0.41
(d): 10.5–21.6 s	1	0.30	-0.80	0.36

Intervals (c) and (d) are best fit with additional cut-off power-law/power-law component

276 C. D. Dermer

can easily become 100 % if $\Gamma_3 \approx 0.3$, and more for higher energy protons, making escape difficult. If long GRBs accelerate UHECRs, then the blast-wave Lorentz factor must be close to that given in the simple $\gamma\gamma$ opacity limit, otherwise the GRB would be highly radiative by hadronic processes. With the strong dependence on Γ, photohadronic interactions become 100 % efficient when $\Gamma \approx 300$, so if GRBs accelerate UHECRs, an accompanying photohadronic γ-ray spectral component is predicted [140, 141, 162, 169].

Proton Models

The comoving synchrotron cooling timescale of an ion with atomic mass A and charge Z is given by (Eqs. (7.49)–(7.50), [4])

$$t'_{syn} = \frac{A^3}{Z^4} \left(\frac{m_p}{m_e}\right)^3 \frac{6\pi m_e c}{\sigma_T B'^2 \gamma'}.$$ (72)

Balancing the synchrotron energy-loss timescale with the shortest energy-gain timescale allowed in standard first- or second-order Fermi acceleration processes gives the relation

$$B'\gamma'^2 = \frac{A^2}{Z^3} \left(\frac{m_p}{m_e}\right)^2 \frac{6\pi m_e c}{\sigma_T \phi},$$ (73)

where $\phi \gtrsim 1$ is the number of radians over which a particle gains $\sim e$ of its original energy. The observed peak synchrotron photon energy is

$$\begin{aligned}
\epsilon_{syn} &\cong \frac{\Gamma \epsilon'_{syn}}{(1+z)} \cong \frac{3\Gamma}{2(1+z)} \frac{Z}{A} \left(\frac{m_e}{m_p}\right) \frac{B'}{B_{cr}} \gamma'^2 \\
&\cong \frac{\Gamma}{(1+z)\phi} \frac{27}{8\alpha_f} \frac{m_p}{m_e} \approx \frac{10^8 \Gamma_3}{(1+z)(\phi/10)} \frac{A}{Z^2}.
\end{aligned}$$ (74)

Overlooking other limitations on particle acceleration, Eq. (74) shows that proton synchrotron radiation from GRB jets and blazars can reach ≈ 50 TeV ($\Gamma \approx 10^3$) and ≈ 5 TeV ($\Gamma \approx 100$), respectively. Proton synchrotron models have been developed in Refs. [127, 161] for γ-ray blazars, and in [171] for GRBs. We consider application of this type of model to these two source classes.

GRBs

Eliminating γ' from these equations using the observer time $t_{syn} = (1+z)t'_{syn}/\Gamma$ for the emission to be radiated at measured energy $m_e c^2 \epsilon_{syn} = \Gamma m_e c^2 \epsilon'_{syn}/(1+z) = 100 E_{100}$ MeV implies a comoving magnetic field [171, 179]

$$B'(\text{G}) = \frac{3}{2^{5/3}} \left(\frac{A^{5/3}}{Z^{7/3}}\right) \left(\frac{1+z}{\Gamma B_{cr}\epsilon_{syn}}\right)^{1/3} \left(\frac{8\pi m_e c}{\sigma_T t_{syn}}\right)^{2/3}$$

$$\cong \frac{1.3 \times 10^5}{t_{syn}^{2/3}(\text{s})} \left(\frac{A^{5/3}}{Z^{7/3}}\right) \left(\frac{1+z}{\Gamma_3 E_{100}}\right)^{1/3}, \tag{75}$$

and an isotropic jet power, dominated by magnetic-field energy, of

$$L_* > L_{*,B} \cong \frac{1}{2} R^2 c \Gamma^2 B'^2 \cong \frac{1}{2} \frac{c^3 \Gamma^6 t_{var}^2 B'^2}{(1+z)^2} = \frac{3^2}{2^{13/3}} \frac{A^{10/3}}{Z^{14/3}} \frac{c^3}{(1+z)^{4/3}} \left(\frac{m_p}{m_e}\right)^{10/3}$$

$$\times \frac{\Gamma^{16/3} t_{var}^2}{B_{cr}^{2/3} \epsilon_{syn}^{2/3}} \left(\frac{8\pi m_e c}{t_{syn}\sigma_T}\right)^{4/3} \cong 2.5 \times 10^{59} \frac{A^{10/3}}{Z^{14/3}} \frac{\Gamma_3^{16/3} t_{var}^2(\text{s})}{[(1+z)t_{syn}(\text{s})]^{4/3} E_{100}^{2/3}} \text{ erg/s}, \tag{76}$$

letting the blast-wave radius $R \cong \Gamma^2 c t_{var}/(1+z)$. For Fe ($A = 56$, $Z = 26$), the power requirements are reduced by a factor ≈ 0.17.

Equation (76) shows that $L_B \propto \Gamma^{16/3}$. The absolute energy release is

$$\mathcal{E}_{*,abs} = \frac{\Delta t}{1+z} f_b L_* \gtrsim \frac{2.5 \times 10^{59}}{(1+z)^{7/3}} \frac{A^{10/3}}{Z^{14/3}} \frac{\Gamma_3^{16/3} t_{var}^2(\text{s})}{E_{100}^{2/3} t_{syn}^{4/3}(\text{s})} \Delta t(\text{s}) f_b \text{ erg}, \tag{77}$$

where f_b is the beaming factor (compare Eq. (11)). Using a 100 kG magnetic field as a fiducial, so $B = 10^5 B_5$ G, then the absolute GRB energy release is $\mathcal{E}_{*,abs} \cong 1.3 \times 10^{59} B_5^2 \Delta t(\text{s}) t_{var}^2(\text{s}) f_b \Gamma_3^6/(1+z)^2$ erg. Such large energy requirements pose a problem for highly magnetized GRB models.

In the model of Razzaque et al. (2010) [171], taking a lower limit $\Gamma_3 \cong 0.5$ implied by naive $\gamma\gamma$ opacity arguments, gives the absolute energy requirements for GRB 080916C [120] of

$$\mathcal{E}_{*,abs} \cong 2 \times 10^{52} E_{100}^{-2/3} (\Gamma_3/0.5)^{16/3} (\theta_j/1 \text{ deg})^2 t_{syn}^{5/3}(\text{s}) \text{ erg},$$

after multiplying by a two-sided jet beaming factor $f_b = (1 - \cos\theta_j) \cong 1.5 \times 10^{-4}(\theta_j/1 \text{ deg})^2$ in the limit $\theta_j \ll 1$, from Eq. (11). The precise values of t_{var}, t_{syn} and Δt depend on model interpretation. For the external shock model [139] applied to GRB 080916C, the rapid variability may be made by irregularities in the complex surrounding medium, so $t_{var} \cong 0.1$ s, the proton synchrotron cooling timescale to make the delayed onset is $t_{syn} \cong 4$ s, and the GRB duration $\Delta t \cong 10$–50 s.

From Eq. (77), strong proton/ion synchrotron radiations can be emitted in the Fermi range near 1 GeV or even 100 MeV due to strong cooling in an intense magnetic field, but the absolute energies are large without a very small beaming factor. The jet break time with apparent isotropic energy release $\approx 2 \times 10^{57}$ ergs is $t_{br} \cong 0.3(\theta_j/1 \text{ deg})^{16/3} n_0^{-1/3}$ days. For such a narrow jet, the jet break would have taken place before Swift slewed to GRB 080916C at $\approx T_0 + 17.0$ h [170].

More complicated geometries might, however, relax the bulk Lorentz factor requirement further [149]. If the inner engine makes the prompt MeV radiation and residual shell collisions at larger radii make LAT γ-ray photons, then Γ could be as low as $\sim$300 [159]. In this case, the absolute energy release from GRB 080916C could easily be $\lesssim 10^{52}$ erg.

For a proton synchrotron model of GRBs to be viable [171], a narrow jet opening angle of order 1° along with a value of $\Gamma \lesssim 0.5$ gives $\mathcal{E}_{*,abs} \cong$ few $\times 10^{52}$ ergs, within ranges implied by interpretation of radio and γ-ray observations and beaming [137]. The detection of distinct components in GRB spectra suggests that a cascading interpretation be more carefully considered [130]. The local rate density of Type Ib/c supernova progenitors of long GRBs like GRB 080916C or low-luminosity GRBs like GRB 980425 or GRB 060218 cannot exceed $\approx$300 Gpc^{-3} year^{-1} [173], and the local beaming-corrected rate density of long duration GRBs is $\approx$10–50 Gpc^{-3} year^{-1} [151, 158]. The local Type 1b/c rate is $9^{+3}_{-5} \times 10^3$ Gpc^{-3} year^{-1} [173]. This agrees with the star-forming galaxy SN Ib/c rate of $\approx$0.28 century^{-1} per 300 Mpc^3 per L_* galaxy $\approx 10^4$ Gpc^{-3} year^{-1}. These imply a beaming rate of $\approx$(10–50)/9000 $\approx 10^{-3}$–0.006. With the $\approx$1° opening angle required to explain GRB 080916C, then all Type 1b/c SNe would have to make GRBs, presenting another challenge for strong magnetic-field jet models of GRBs.

Blazars

Proton synchrotron and photohadronic models for blazars have been developed, as noted, in [127, 161]. The protons and ions make a γ-ray component that originates from pion-induced cascades, and the synchrotron radiations of pions and muons make additional γ-ray emissions. Primary electron synchrotron radiation is still usually required to make the nonthermal radio/optical/X-ray synchrotron blazar emission.

Anita Reimer gives [123] a detailed fit to the famous HBL Mrk 421 at $z = 0.031$ at luminosity distance $d_L = 134$ Mpc. Mrk 421 is one of the first Whipple TeV blazars and the first one with sub-hour [146] variability detected. For the composite SED of Mrk 421 averaged between 2009 January 9 (MJD 54850) to 2009 June 1, Reimer makes a proton fit with the parameters shown in Table 6.

The Hillas criterion [152] for a single-zone blob model requires the comoving Larmor radius to be smaller than the comoving blob radius, that is,

$$r'_L < r'_b. \tag{78}$$

Table 6 Proton blazar model for Mrk 421 [123]

Parameter	Value
B'	50 G
$\delta_D \approx \Gamma$	12
r'_b	4×10^{14} cm
E_{max}	$7.2 \times 10^{19} Z \Gamma_{12} B'_{50}$ eV[a]

[a] $\Gamma_{12} = \Gamma/12$, $B'_{50} = B'/50$ G, $\delta_{12} = \delta_D/12$

With Eq. (16) giving the maximum escaping particle energy $E_{max} \approx \Gamma E'_{max} \approx Z e B' r'_b$, we see that this model blazar can accelerate super-GZK ($E \gtrsim 5 \times 10^{19}$ eV) particles if charged ions, but this is only possible for protons with slightly larger magnetic fields. These protons have maximum comoving Lorentz factors

$$\gamma'_{p,max} \approx 6.4 \times 10^9 B'_{50}.$$

For particles well-trapped by the magnetic field of the plasma, $\gamma'_{p,max} = 10^9 \gamma'_9$ with $\gamma'_9 \ll 1$.

For the 2-sided jet power as defined in Eq. (47), taking $\delta_D \approx \Gamma$ gives

$$L_j = 2\pi r_b'^2 \beta c \Gamma^2 \sum_i u'_i, \quad i = p, e, B, \tag{79}$$

and the total Poynting power dominated by magnetic field energy is

$$L_{j,B} = \frac{1}{4} r_b'^2 \beta c \Gamma^2 B'^2 \cong 4.3 \times 10^{44} \Gamma_{12}^2 B_{50}'^2 \text{ erg/s}, \tag{80}$$

which is not excessive, considering that the Eddington luminosity is $L_{Edd} \approx 10^{47} M_9$ erg/s, and $M_9 \approx 0.2$–0.9 for the supermassive black hole power Mrk 421.

We still need to calculate the particle power and the efficiency for hadronic production of γ-ray emission. In the case of proton synchrotron radiation amounting to total energy flux Φ (erg/cm^2-s), the particle power is

$$L_{j,p} = 2\pi r_b'^2 \beta c \Gamma^2 \left(\frac{N_{p0} m_p c^2 \gamma'_{p,max}}{V'_b} \right) = \frac{3}{2} \frac{m_p c^3 N_{p,0} \gamma'_{p,max} \Gamma^2}{r'_b}, \tag{81}$$

for a spherical comoving emission region.

According to our standard approach [4],

$$\nu F_\nu = f_\epsilon \cong \delta_D^4 \frac{L'}{4\pi d_L^2} \sim \Phi.$$

The proton-synchrotron energy loss rate is

$$-\left(\frac{dE'}{dt'} \right)_{p,syn} = \frac{4}{3} \left(\frac{m_e}{m_p} \right)^2 c \sigma_T U'_{B'} \gamma'^2, \tag{82}$$

where $U'_{B'} = B'^2/8\pi$.

For a large range of proton spectra, the γ-ray power from photohadronic processes can be approximated by the number $N_{p,0}$ of protons with Lorentz factor $\gamma' \approx \gamma'_{p,max}$, from which follows that

$$N_{p,0} = \left(\frac{m_p}{m_e}\right)^2 \frac{3\pi d_L^2 \Phi}{\delta_D^4 c \sigma_T U'_{B'} \gamma'^2_{p,max}} \tag{83}$$

and

$$L_{j,p} = \left(\frac{m_p}{m_e}\right)^2 \frac{9\pi m_p c^3 d_L^2 \Gamma^2 \Phi}{2 r'_b \delta_D^4 c \sigma_T U'_{B'} \gamma'_{p,max}} \approx 3.2 \times 10^{44} \frac{\Gamma_{12}^2 \Phi_{-10}}{\delta_D^4 B'^2_{50} \gamma'_9} \text{ erg/s}, \tag{84}$$

where the bolometric γ-ray energy flux supplied by the proton synchrotron process is $10^{-10} \Phi_{-10}$ erg/cm^2-s. The particle power is reasonable, even if $\gamma'_9 \to 0.1$. The proton synchrotron frequency, from Eq. (74),

$$\epsilon_{p,syn} \cong \frac{3\delta_D}{2(1+z)} \frac{m_e}{m_p} \frac{B'}{B_{cr}} \gamma'^2_{p,max} \cong 1.1 \times 10^4 \gamma'^2_9 B'_{50} \delta_{12}, \tag{85}$$

so

$$E_{p,syn} \cong 5.6 \gamma'^2_9 B'_{50} \delta_{12} \text{ GeV}. \tag{86}$$

The equipartition field is, from Eq. (49), or Eq. (7.81) [4]

$$B_{eq}(\text{G}) \cong 130 \frac{d_{28}^{4/7} (1+z)^{5/7} f_{-10}^{2/7} \Lambda^{2/7}}{t_d^{6/7} \delta_D^{13/7} \nu_{13}^{1/7}}. \tag{87}$$

In the case of Mrk 421, $\nu_{13} = 10^4 \nu_{17}$, $\Lambda = 10 \Lambda_{10}$, $t_d = (1+z) r'_b / c\delta_D = 1200/\delta_{12}$ s, implying that

$$B_{eq}(\text{G}) \cong 4.5 \frac{f_{-10}^{2/7} \Lambda_{10}^{2/7}}{\delta_{12} \nu_{17}^{1/7}}, \tag{88}$$

so that this model is magnetically dominated with a magnetic field $\approx 10 \times B_{eq}$.

The minimum bulk Lorentz factor from $\gamma\gamma$ constraints, using Eq. (45), is $\Gamma_{min} = 20.3 f_{-10}[E_1/(10\,\text{TeV})]^{1/6} = 13.8 f_{-10}[E_1/(\text{TeV})]^{1/6}$, for a maximum measured photon energy E_1 and using a variability timescale of $\approx 10^3$ s. Note that f_{-10} represents the νF_ν flux for the target photons and, though both have the same units, is a differential quantity compared to the bolometric energy flux Φ. The condition for γ-ray transparency is somewhat inconsistent with the detection of multi-TeV photons from Mrk 421.

The cooling regime for ultra-high energy protons through synchrotron cooling is determined by the ratio $\rho \equiv t'_{syn}/t'_{dyn}$, with strong synchrotron cooling when $\rho \ll 1$ and weak cooling when $\rho \gg 1$. From Eq. (72) we have

$$\rho = \frac{3(m_p/m_e)^3}{4\sigma_T U'_{B'} \gamma'_p r'_b} \to \frac{140}{\gamma'_9}. \tag{89}$$

Thus this proton synchrotron model for Mrk 421 is in the weak cooling regime. Alternately, we can use the expression [171] for the saturation Lorentz factor

$$\gamma'_{sat} = \left(\frac{m_p}{m_e}\right)\sqrt{\frac{B_{cr}}{B'}}\sqrt{\frac{9}{4\alpha_f}} = \frac{3 \times 10^{10}}{\sqrt{B'_{50}}}, \tag{90}$$

for which particles with larger Lorentz factors are in the strong cooling regime (here $\alpha_f = e^2/\hbar c \approx 1/137$ is the fine structure constant).

We can use Eq. (71) to determine the photopion production efficiency for this model. For $\epsilon_{pk} = 10^{17}\nu_{17}/1.24 \times 10^{20} = 8.1 \times 10^{-4}\nu_{17}$, $a \cong 0$ and $b \cong -1$, we obtain

$$\eta^{pk}_{p\gamma} = \frac{5.6 \times 10^{-7} f_{-10}}{t(days)(\Gamma/12)^4\nu_{17}} \rightarrow 4.3 \times 10^{-5}\frac{f_{-10}}{\nu_{17}}, \tag{91}$$

taking $t_v = (1 + z)r'_b/c\delta_D = 0.013/\delta_{12}$ days. The comoving Lorentz factor of protons interacting with target photons with energy ϵ'_{pk} is $\gamma'_{pk} \cong 6 \times 10^6$. Protons with larger Lorentz factors interact primarily with the $a \cong 0$ portion of the νF_ν spectrum with frequencies less than ν_{pk}. The efficiency of proton energy loss with $\gamma' \gg \gamma'_{pk}$ approaches

$$\eta_{p\gamma}(E_p) \approx \eta^{pk}_{p\gamma}\left(\frac{E_p}{E_{pk}}\right) \cong 7.4 \times 10^{-3}\frac{f_{-10}\gamma'_9}{\nu_{17}}. \tag{92}$$

For photopion losses (as for proton synchrotron losses), the energy loss is most efficient for the highest energy particles, reaching $\sim 1\%$ for $\gamma'_9 \sim 1$. Somewhat larger efficiencies are allowed—provided these allowances are consistent with the variability data—if the protons are trapped on long times compared with the crossing time, and the emission region is slow to expand.

In conclusion, hadronic models for blazars and GRBs face no insurmountable objections based on power or energetics.

4.4 Cascade Halos and the Intergalactic Magnetic Field (IGMF)

The magnetic field B_{IGM} in the IGM is bounded by a primordial field generated by quantum fluctuations in the early universe or decoupling transitions of the fundamental forces [144]. Dynamo processes amplify the seed fields. Gamma ray astronomy provides a method for measuring the intergalactic magnetic field (IGMF) by exploiting the effects of the $\gamma\gamma \rightarrow e^+e^-$ pair-production process on γ rays from extragalactic TeV sources interacting with soft photons of the EBL.

When VHE γ rays interact with ambient radiation fields and the EBL, then blazars, and therefore radio galaxies (i.e., misaligned blazars) are surrounded by anisotropic

Table 7 Derived limits on B_{IGM} for the source 1ES 0229+200 assuming persistent emission

1ES 0229+200	θ_j (rad)	B_{IGM}(G)
Reference [166]	π	$\gtrsim 3 \times 10^{-16}$
Reference [177]	0.1	$\gtrsim 5 \times 10^{-15}$
Reference [178]	0.03	$\gtrsim 2 \times 10^{-15}$
Reference [143]	0.1	$\gtrsim 5 \times 10^{-15}$

jets of relativistic leptons made when the γ rays materialize into energetic e^+e^- pairs. These leptons spiral in the ambient magnetic field to make extended synchrotron radiation halos [126], and GeV γ rays by Compton-scattering photons of the CMBR [138]. Arrival time information in pair cascades generated by impulsive sources of high-energy, multi-TeV γ radiation can also be used to infer the strength of the IGMF [168, 164].

Because the emission of 1–10 TeV photons from a source at redshift $z \ll 1$ is attenuated by, primarily, the IR EBL, then the cascade spectrum can be calculated for a given spectral model of the EBL and properties of B_{IGM}. For sufficiently weak magnetic fields, the pairs travel rectilinearly while Comptonizing CMB photons to γ ray energies. The absence of this cascade signature in joint Fermi-HESS observations of candidate TeV blazars implies a lower limit on B_{IGM} [138] under the assumption that the blazar persistently emits high-energy radiation for arbitrarily long time. Neronov & Vovk [166] and Tavecchio et al. [177] argued that nondetection of the TeV blazars 1ES 1101-232 ($z = 0.186$), 1ES 0229+200 ($z = 0.14$), 1ES 0347-121 ($z = 0.188$), and H 2356-309 ($z = 0.165$) by Fermi implies a lower limit $B_{IGM} \gtrsim 3 \times 10^{-16}$ G (Table 7).

The high-energy electrons and positrons also undergo deflections in the ambient magnetic field, so the emissions arriving latest generally come from leptons that have been most severely deflected. This will cause steady sources to be surrounded by an extended angular halo formed by leptons deflected back into the line of sight. Ando & Kusenko [129] claim that $\sim 30'$ halos are found in stacked Fermi data for 170 AGNs that are bright at 10–100 GeV. By fitting the angular distribution of the halo, they deduce that $B_{IGM} \approx 10^{-15}(\lambda_{coh}/\text{kpc})^{-1/2}$, for magnetic-field correlation length $\lambda_{coh} \sim 10$–100 kpc. This claim is disputed in [167], in part because of the use of P6_v3 response functions which are known to have inaccurate PSF from on-orbit calibration data [8].

More realistic limits on B_{IGMF} based on evidence that the blazar was operating only during the time frame over which it was observed, was proposed in [142], and independently, in [143].

Cascade Radiation from EBL Attenuation of TeV Photons

Consider a source and observer separated by distance d, as shown in Fig. 13. Photons with dimensionless energy $\epsilon_1 = h\nu_1/m_ec^2$ emitted at angle θ_1 to the line of sight between the source and observer, travel an average distance $\lambda_{\gamma\gamma} = \lambda_{\gamma\gamma}(\epsilon_1, z)$ before
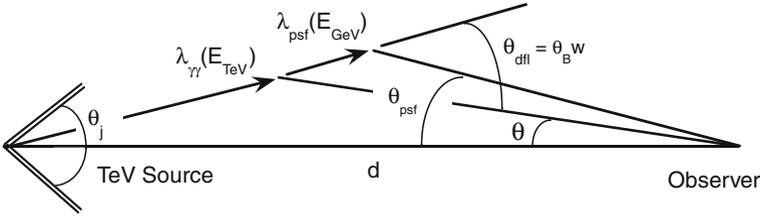
Fig. 13 Sketch of the geometry of the event. A photon with energy E_{TeV} TeV emitted at angle $\theta_1 \leq \theta_j$ to the line of sight, where θ_j is the jet half-opening angle, interacts with an EBL photon to create an electron-positron pair. The electron is deflected by angle θ_{dfl} and scatters a CMB photon to energy E_{GeV} GeV and angle θ, which is detected as a source photon by the Fermi LAT when θ falls within the angular point spread function $\theta_{psf}(\mathrm{GeV})$ from the source direction

materializing into an electron-positron pair via $\gamma\gamma$ absorption with photons of the EBL. After production, the pairs cool by scattering CMB radiation, which is detected at an angle θ to the line of sight to the source when the secondary electrons and positrons (hereafter referred to as electrons) are deflected by an angle θ_{dfl}. Thus $\theta_1 = \theta_{\mathrm{dfl}} - \theta$. For the purposes of the argument, we neglect redshift effects for the TeV blazar sources under consideration, which limits the treatment to sources at $z \lesssim 0.2$ (see [165] for redshift corrections).

The time delay Δt between the reception of photons directed towards the observer and those which undergo the process described above is given through the expression

$$c\Delta t = \lambda_{\gamma\gamma} + x - d = \lambda_{\gamma\gamma} + \frac{d \sin(\theta_{\mathrm{dfl}} - \theta)}{\sin \theta_{\mathrm{dfl}}} - d$$
$$= \lambda_{\gamma\gamma}(1 - \cos \theta_{\mathrm{dfl}}) - d(1 - \cos \theta), \tag{93}$$

noting that $x = d \sin \theta_1 / \sin \theta_{\mathrm{dfl}}$ and $\lambda_{\gamma\gamma} = d \sin \theta / \sin \theta_{\mathrm{dfl}}$. In the limit of small observing and deflection angles, Eq. (93) implies

$$\Delta t \cong \frac{\lambda_{\gamma\gamma}}{2c} \theta_{\mathrm{dfl}}^2 \left(1 - \frac{\lambda_{\gamma\gamma}}{d}\right). \tag{94}$$

Moreover, this time delay is observed by photons received at

$$\theta \cong \theta_{\mathrm{dfl}} \frac{\lambda_{\gamma\gamma}}{d} \tag{95}$$

to the line of sight.

The deflection angle depends on the Lorentz factor $\gamma = 10^6 \gamma_6$ of the produced electrons, and can be written in terms of the received photon energy $E = E_{\mathrm{GeV}}$ GeV. The average photon energy of the CMB at low redshift is $\epsilon_0 \approx 1.24 \times 10^{-9}$ in $m_e c^2$ units, so that the mean Thomson-scattered photon energy is $\epsilon_T \approx (4/3)\epsilon_0 \gamma^2$. Thus, an electron with Lorentz factor γ scatters CMB radiation to photon energy E when

$\gamma_6 \cong 1.1\sqrt{E_{\text{GeV}}}$. The characteristic length scale for energy losses due to Thomson scattering is

$$\lambda_T = \frac{3m_e c^2}{4\sigma_T u_{\text{CMB}}\gamma} = \left(\frac{0.75}{\gamma_6}\right) \text{Mpc}, \tag{96}$$

where $u_{\text{CMB}} \cong 4 \times 10^{-13} \text{erg cm}^{-3}$ is the CMB energy density at low redshifts. While losing energy, the electron is deflected by an angle $\theta_B \cong \lambda_T/r_L$ in a uniform magnetic field of strength $B_{\text{IGM}} = 10^{-15} B_{-15}$ G oriented perpendicular to the direction of motion of the electron, where the Larmor radius $r_L = m_e c^2\gamma/eB \cong 0.55(\gamma_6/B_{-15})$ Mpc. Thus the deflection angle for an electron losing energy by scattering CMB photons to energy E in a uniform field is $\theta_B = \lambda_T/r_L \cong 1.1 B_{-15}/E_{\text{GeV}}$. Introducing a coherence length λ_{coh} that characterizes the typical distance over which the magnetic field direction changes by $\approx\pi/2$, then the deflection angle

$$\theta_{\text{dfl}} \cong \theta_B \begin{cases} 1, & if \ \lambda_T < \lambda_{coh} \\ \sqrt{\frac{\lambda_{coh}}{\lambda_T}}, & if \ \lambda_T > \lambda_{coh}. \end{cases} \equiv w\theta_B, \tag{97}$$

with $w \equiv H(\lambda_{coh} - \lambda_T) + \sqrt{\lambda_{coh}/\lambda_T} H(\lambda_T - \lambda_{coh})$.

The EBL model of [145] for sources at $z = 0.14$ gives $\lambda_{\gamma\gamma}(E) \cong 200$ Mpc, 125 Mpc, and 70 Mpc at $E = 1, 3$, and 10 TeV, respectively. A low EBL based on galaxy counts [156] gives values of $\lambda_{\gamma\gamma}(E) \cong 280$ Mpc, 150 Mpc, and 85 Mpc at $E = 1, 3$, and 10 TeV, respectively. Thus we write $\lambda_{\gamma\gamma} = 100\lambda_{100}$ Mpc, so that $\lambda_{100} \approx 1$ gives a minimum value of Δt and θ for the reprocessed TeV radiation.

For a source at distance $d = d_{\text{Gpc}}$ Gpc, with $d_{\text{Gpc}} \sim 1$ corresponding to $z \sim 0.2$, the time delay for emission observed at angle

$$\theta \cong 0.01 \frac{\lambda_{100}}{d_{\text{Gpc}}} \left(\frac{B_{-15}w}{E/10\,\text{GeV}}\right) \tag{98}$$

from the line of sight is given from Eq. (94) by

$$\Delta t\,(\text{year}) \cong 2 \times 10^6 \lambda_{100} \left(\frac{B_{-15}w}{E/10\,\text{GeV}}\right)^2 \left(1 - 0.1\frac{\lambda_{100}}{d_{\text{Gpc}}}\right) \tag{99}$$

Equation (99) shows that small time delays are implied when $\lambda_{\gamma\gamma}$ is small or $\lambda_{\gamma\gamma} \approx d$. When $\lambda_{\gamma\gamma} \lesssim \lambda_T$, an additional delay $\approx\lambda_T\theta_{\text{dfl}}^2/c$ arises during the time that the electrons are losing energy and being deflected by the IGMF [153, 169]. Such small values of $\lambda_{\gamma\gamma} \sim 1$ Mpc are only relevant at low redshifts to $\gtrsim 100$ TeV photons pair producing within ≈ 1 Mpc of their source which, however, may be in the galaxy cluster environments where the magnetic field is not representative of the dominant volume of the voids. Thus we can dismiss such an origin of a short time delay without assuming special properties of the TeV sources. The case $\lambda_{\gamma\gamma} \sim d$ formally gives short time delays, but this corresponds to the case when the mean free path for $\gamma\gamma$ pair production is about equal to the source distance, which occurs

for γ rays with energies of several hundred GeV when $d \sim 1\,\text{Gpc}$. In this case, the secondary electrons which take half the energy of the high-energy photon scatter CMB photons to $\ll 100\,\text{MeV}$. Even so, the attenuation of the high-energy photons takes place over the entire distance d, so that the upscattered photons with short delay time comprise only a very small fraction of the incident flux attenuated close to the observer.

The remaining alternative to avoid assuming that TeV blazars are steady on timescales of millions of years is to suppose that either $B_{IGM} \ll 10^{-15}\,\text{G}$ or that $\lambda_{coh} \ll 1\,\text{Mpc}$. If $\lambda_{coh} \sim 1\,\text{Mpc}$, this contradicts the claim that the IGMF has been measured to be $B_{IGM} \gtrsim 10^{-15}\,\text{G}$. If $\lambda_{coh} \ll 1\,\text{Mpc}$, then the field must be even larger in order that the electrons are deflected away from the direction of the photon source. Thus TeV blazars must be assumed to be steady emitters on long timescales. Here we relax this assumption, and suppose most cautiously that the blazar has been operative over the last few years that the they have been observed. This reduces the implied lower limit by several orders of magnitude, as will be seen. Here we treat the blazar as a point source; see [142] when the Fermi-LAT PSF constraint is included (Fig. 13).

Derivation of the Cascade Spectrum

Consider a source that emits TeV-scale photons from within a photon beam of half-angle θ_j, and an apparent isotropic spectral luminosity $\epsilon_* L_*(\epsilon_*)$ within the beam ("flat-topped" jet). If the TeV photons can escape from the nuclear environments without $\gamma\gamma$ absorption, then they may still be attenuated by interactions with EBL photons. After $\gamma\gamma$ attenuation, e$^+$-e$^-$ pairs lose energy, primarily by Compton scattering the CMBR. This upscattered radiation will be detected by an observer until the cooling pairs are deflected out of the beam.

We use the notation of Ref. [4]. From Eq. (7), denote the νF_ν spectrum by

$$f_\epsilon = \nu F_\nu = \frac{\epsilon_* L_*(\epsilon_*)}{4\pi d_L^2}, \ \epsilon = h\nu/m_e c^2, \ \epsilon_* = (1+z)\epsilon.$$

The correction due to EBL attenuation is

$$f_\epsilon = \frac{\epsilon_* L_*(\epsilon_*)}{4\pi d_L^2} \exp[-\tau_{\gamma\gamma}(\epsilon; z)] = \frac{m_e c^2 \epsilon^2 \dot{N}(\epsilon)}{4\pi d_L^2} \exp[-\tau_{\gamma\gamma}(\epsilon, z)]. \qquad (100)$$

At low redshifts, $z \ll 1$, $\epsilon_* \approx \epsilon$, and $\epsilon L(\epsilon) = m_e c^2 \epsilon^2 \dot{N}(\epsilon)$, where $\dot{N}(\epsilon)$ is the photon injection function. The number of photons surviving to the observer is $\dot{N}(\epsilon) \exp[-\tau_{\gamma\gamma}(\epsilon; z)]$, so the number of absorbed photons is $\dot{N}_{abs}(\epsilon) = \dot{N}(\epsilon)[1 - \exp[-\tau_{\gamma\gamma}(\epsilon; z)]]$, and this also represents the electron injection function $\dot{N}_{inj}(\gamma_i)$, with normalization

$$\int_1^\infty d\gamma_i \ \dot{N}_{inj}(\gamma_i) = 2 \int_0^\infty d\epsilon \ \dot{N}_{abs}(\epsilon), \tag{101}$$

because each photon makes two leptons with $\gamma_i \cong \epsilon/2$. So $\dot{N}_{inj}(\gamma_i) = 4\dot{N}_{abs}(\epsilon) = 4\dot{N}(\epsilon)\{\exp[\tau_{\gamma\gamma}(\epsilon, z) - 1]\}$. Using Eq. (100) gives the injection function

$$\dot{N}_{inj}(\gamma_i) = \frac{16\pi d_L^2 f_\epsilon}{m_e c^2 \epsilon^2} \ \{\exp[\tau_{\gamma\gamma}(\epsilon, z)] - 1\}, \ \epsilon = 2\gamma_i. \tag{102}$$

This injection source of e^+-e^- loses energy by Compton-scattering photons of the CMBR to GeV energies, and this cascade GeV component is considerably dimmed when the leptons are deflected out of the beam.

Photons with energy ≈ 1 TeV make leptons with $\gamma \approx 10^6$,[15] which scatter the CMBR to $\gamma^2 \epsilon_0 \sim 10^3$, or to photon energies ~ 500 MeV. Scattering is in the Thomson regime for $4\gamma\epsilon_0 \ll 1$ or $\gamma \ll 2 \times 10^8$, that is, electrons with energies $\ll 100$ TeV.

The Thomson energy-loss rate

$$-\dot{\gamma}_T = -\frac{d\gamma}{dt}|_T = \frac{4}{3} c\sigma_T \frac{u_0}{m_e c^2} \gamma^2 \equiv \nu_T \gamma^2, \tag{103}$$

with $u_0/m_e c^2 = 4.9 \times 10^{-7}$ cm^{-3}. The solution to the steady-state electron continuity equation is

$$N(\gamma) = \frac{1}{\nu_T \gamma^2} \int_\gamma^\infty d\gamma' \ \dot{N}(\gamma'). \tag{104}$$

Limiting $\gamma > \gamma_{dfl}$, where γ_{dfl} is the deflection Lorentz factor where the lepton is deflected out of the jet opening angle by the IGMF. Taking the luminosity spectrum from Compton scattering [4], Eqs. (102) and (28) give

$$\epsilon_s L_C(\epsilon_s) = 4\pi d_L^2 f_{\epsilon_s}$$
$$= \frac{12\pi d^2 c\sigma_T u_0}{\nu_T m_e c^2} \left(\frac{\epsilon_s}{\epsilon_0}\right)^2 \int_{\gamma_{low}}^\infty d\gamma \ \frac{F_C(q, \Gamma_e)}{\gamma^4} \int_\gamma^\infty d\gamma_i \ \frac{f_\epsilon \ \{\exp[\tau_{\gamma\gamma}(\epsilon, z)] - 1\}}{\epsilon^2}, \tag{105}$$

with $\epsilon = 2\gamma_i$. The interior integral is the injection function from high-energy γ rays absorbed by photons of the EBL, and depends on a model of the optical-depth function $\tau_{\gamma\gamma}(\epsilon, z)$ for a photon with detected energy ϵ that was emitted by a source at z. In general, one uses the expression Eq. (31) for the Compton kernel $F_C(q, \Gamma_e)$, though the isotropic Thomson kernel, Eq. (33), is sufficiently accurate our purposes here.

---

[15] At high energies, most of the energy is taken by one of the leptons. See, e.g., [157].

The lower limit γ_{low} in Eq. (105) is the maximum of various constraints given by kinematic factors, engine duration, and γ_{dfl}. The kinematic Lorentz factor γ_{knm} given by Eq. (35). In the Thomson regime, $\gamma_{knm} \rightarrow (1/2)\sqrt{\epsilon_s/\epsilon_0}$.

The deflection Lorentz factor is determined by the condition that the energy-loss timescale is equal to the timescale for e^+-e^- pairs to be deflected out of the beam. The Thomson energy loss timescale $-\dot{\gamma}_T = \nu_T \gamma^2$ implies $t_T = 1/\nu_T \gamma$. The deflection timescale

$$t_{dfl} = \frac{\theta_j r_L}{c} = \theta_j \left(\frac{m_e c \gamma}{eB} \right), \quad \text{when } \lambda_T < \lambda_{coh}, \tag{106}$$

and $t_{dfl} = \theta_j (m_e c \gamma / eB) \sqrt{\lambda_T / \lambda_{coh}}$ when $\lambda_T > \lambda_{coh}$. Solving gives

$$\gamma_{dfl} = \begin{cases} \sqrt{\dfrac{eB}{\theta_j m_e c \nu_T}}, & \gamma_{dfl} > \dfrac{c}{\nu_T \lambda_{coh}} \\ \left(\dfrac{c \lambda_{coh}}{\nu_T} \right)^{1/3} \left(\dfrac{eB}{m_e c^2 \theta_j} \right)^{2/3}, & \gamma_{dfl} < \dfrac{c}{\nu_T \lambda_{coh}} \end{cases}. \tag{107}$$

This constraint implies $\gamma_{dfl} = 3.7 \times 10^6 \sqrt{B_{-15}/\theta_{-1}}$ for $\gamma > 7.5 \times 10^5 / \lambda_{coh}(\text{Mpc})$, and $\gamma_{dfl} = 6.2 \times 10^6 \lambda_{coh}^{1/3}(\text{Mpc})(B_{-15}/\theta_{-1})^{2/3}$ for $\gamma > 7.5 \times 10^5 / \lambda_{coh}(\text{Mpc})$.

To avoid solving a time-dependent electron continuity equation, we introduce electron Lorentz factor limits $\gamma(\Delta t)$ to define the time the engine was operating. Following Eq. (34), noting that $f_{\epsilon_s} = \epsilon_s L(\epsilon_s)/4\pi d^2$, gives

$$f_{\epsilon_s} = \frac{3}{2} \left(\frac{\epsilon_s}{\epsilon_0} \right)^2 \int_{\max \left[\sqrt{\frac{\epsilon_s}{4\epsilon_0}}, \gamma_{dfl}, \gamma(\Delta t) \right]}^{\infty} d\gamma \, \gamma^{-4} \left(1 - \frac{\epsilon_s}{4\gamma^2 \epsilon_0} \right),$$

$$\times \int_{\gamma}^{\infty} d\gamma_i \, \frac{f_\epsilon \{\exp[\tau_{\gamma\gamma}(\epsilon, z)] - 1\}}{\epsilon^2}, \tag{108}$$

with $\epsilon = 2\gamma_i$.

The minimum Lorentz factor γ related to period Δt of activity of central engine is determined by equating the time delay with the extra pathlength followed by photons. Thus

$$\Delta t \cong \frac{\lambda_{\gamma\gamma}(\epsilon_1) + \lambda_T(\gamma)}{2c} \, \theta_{dfl}^2, \tag{109}$$

and $\lambda_{tot} = \lambda_{\gamma\gamma}(\epsilon_1) + \lambda_T(\gamma) = 100\lambda_{100}$ Mpc, $\theta_{dfl} = w\theta_B$, $\theta_B = \lambda_T/r_L$. From this we derive

$$\gamma(\Delta t)$$

$$= \begin{cases} \sqrt{\dfrac{eB}{m_e c \nu_T}} \left(\dfrac{\lambda_{tot}}{2c\Delta t} \right)^{1/4} \cong \dfrac{9.9 \times 10^9 \lambda_{100}^{1/4} B_{-15}^{1/2}}{[\Delta t(s)]^{1/4}}, & \dfrac{7.5 \times 10^5}{\lambda_{coh}(\text{Mpc})} < \gamma \\ \left(\dfrac{eB}{m_e c^2} \right)^{2/3} \left(\dfrac{\lambda_{tot}\lambda_{coh}}{2\nu_T \Delta t} \right)^{1/3} \cong \dfrac{2.3 \times 10^{11} \lambda_{100}^{1/3} B_{-15}^{2/3} \lambda_{coh}^{1/3}(\text{Mpc})}{[\Delta t(s)]^{1/3}}, & \gamma < \dfrac{7.5 \times 10^5}{\lambda_{coh}(\text{Mpc})} \end{cases}. \tag{110}$$

Note that there is no separate constraint that the electrons have cooled long enough to scatter significant emission in the given band. In the stationary frame, one can define the cooling electron Lorentz factor γ_{cool} giving the characteristic Lorentz factor of electrons that have cooled in time Δt, expressed in terms of the GeV photon energy E_{GeV} to which CMBR photons are Compton-scattered CMB photons. At low redshifts,

$$t_{\mathrm{T}} = \Delta t = \frac{3}{4c\sigma_{\mathrm{T}}u_0\gamma_{cool}(\Delta t)} = \frac{\sqrt{3}}{2c\sigma_{\mathrm{T}}u_0}\sqrt{\frac{\epsilon_0}{\epsilon}} \simeq \frac{7.7 \times 10^{19}\mathrm{s}}{\gamma} \simeq \frac{2.2\,\mathrm{Myr}}{\sqrt{E_{\mathrm{GeV}}}}, \quad (111)$$

we see that it takes nearly a million years for the generated electrons and positrons to cool and begin to make their strongest emissions in the GeV–10 GeV band of Fermi. But this applies for pair halo emissions made at large angles to the jet axis. The cooling timescale compared to the engine starting time, as measured by an observer within the jetted emission cone, is instead given by Eq. (109), which we write as

$$\Delta t\,(\mathrm{s}) \cong 60\,\lambda_{100}\,\frac{B_{-22}^2}{E_{\mathrm{GeV}}^2}w^2. \quad (112)$$

The lack of distinct GeV echoes from impulsive or flaring high-energy sources can limit parts of the B_{IGM}-λ_{coh} parameter space. Better yet, the discovery of such echoes in GRB light curves, as originally proposed by Plaga (1995) [168], would finally reveal the primordial magnetic field. From Eq. (111), the Thomson cooling length of relativistic electrons scattering the CMBR is, at redshift z, given by

$$\lambda_{\mathrm{T}} = \frac{2.3 \times 10^{30}\mathrm{cm}}{\gamma(1+z)^4} \simeq \frac{0.7\,\mathrm{Mpc}}{\sqrt{E_{\mathrm{GeV}}}}, \quad (113)$$

where the last expression applies at low redshifts.

Figure 14 shows calculations with this semi-analytic model [142]. Here we use a photon-energy dependent expression for $\lambda_{\gamma\gamma}$ and perform an integration over the CMBR spectrum. If the jet is persistent on long time scales, then the jet opening angle is limited to be $\lesssim 0.4$, as can be seen in the middle panels of this figure. Restricting TeV activity of 1ES 0229+200 to $\approx$3–4 years, during which the source has been observed, leads to a more robust lower limit of $B_{\mathrm{IGM}} \gtrsim 10^{-18}$ G, which can be larger by an order of magnitude if the intrinsic source flux above $\approx$5–10 TeV from 1ES 0229+200 is strong.

If there were no intergalactic magnetic field at all, then the pairs made from a source at distance d, minus the distance $\lambda_{\gamma\gamma}$ over which they are made, will have cooled to electron Lorentz factors γ given by $d - \lambda_{\gamma\gamma} = \lambda_{\mathrm{T}} \cong 750\,\mathrm{Mpc}/(\gamma/10^3)$. It is interesting to think that the pair injection process will have seeded the voids of intergalactic space with an ultra-relativistic nonthermal electron-positron component that has already cooled to low energies. Given how active TeV sources are in our
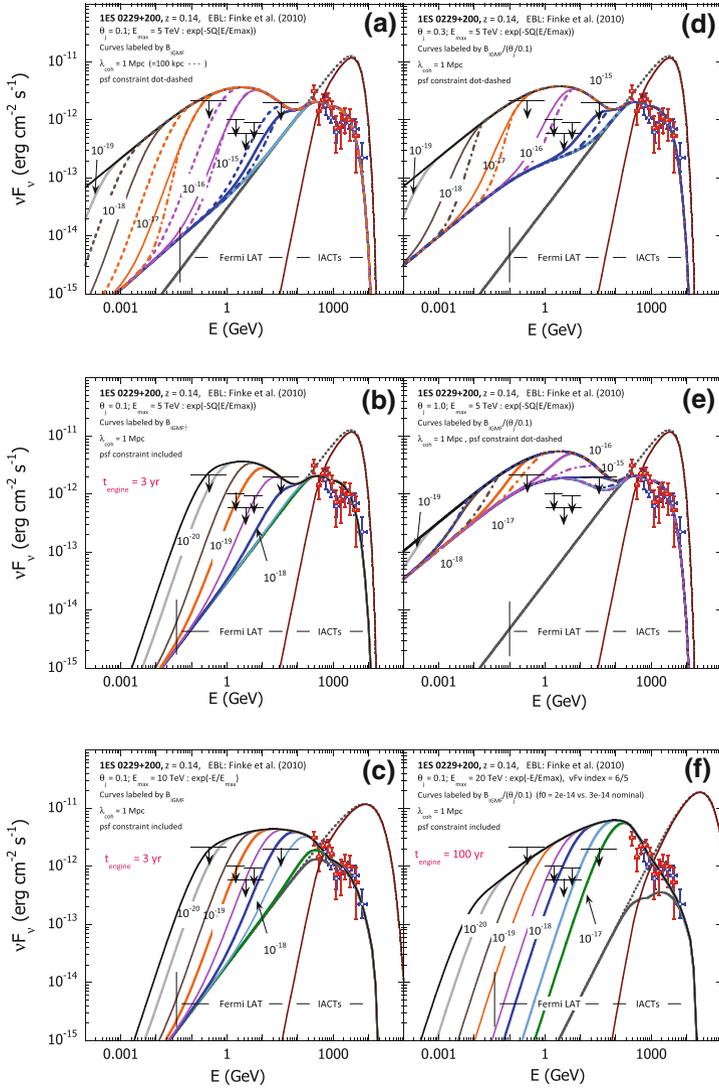
Fig. 14 Model of cascade radiation spectrum, Eq. (100), applied to HESS, VERITAS, and Fermi observations of 1ES 0229+200, using model spectra (*solid curves*) and EBL model of [145] to give attenuated source spectra. **a** Cascade spectra for 1ES 0229+200 assuming persistent TeV emission at the level observed with HESS and VERITAS, for different values of B_{IGM} and $\lambda_{coh} = 1\,\mathrm{Mpc}$ (*solid*) or $\lambda_{coh} = 100\,\mathrm{kpc}$ (*dot-dashed*) for a jet opening angle $\theta_j = 0.1$. The intrinsic TeV source spectrum is given by power-law with νF_ν index $= 4/5$ with super-exponential cutoff $\propto \exp[-(E/5\,\mathrm{TeV})^2]$. The PSF constraint for the $\lambda_{coh} = 1\,\mathrm{Mpc}$ case is shown by the *dot-dashed curves*. **d** Same as for **a**, except that $\theta_j = 0.3$. **e** Same as **a**, except that $\theta_j = 1.0$. **b, c, d** Here the TeV engine operates for 3 years, 3 years, and 100 years, respectively, but the intrinsic TeV source spectrum differs. In **b**, it is the same as **a**. In **c**, it has an exponential cutoff $\propto \exp(-(E/10\,\mathrm{TeV}))$. In **f**, the intrinsic source spectrum use parameters of Dolag et al. (2011) [143]. In (b), (c), and (d), $\lambda_{coh} = 1\,\mathrm{Mpc}$ and cascade spectra are calculated for different values of B_{IGM}, as labeled

present declining phase of the universe (in terms of star formation activity), TeV sources must have been vigorous operating in the early universe.

Without making severe assumptions about the γ-ray duty cycle and radiative behavior of blazars like 1ES 1101-232 and 1ES 0229+200 on long timescales, the best limit to the strength of the IGMF is $\gtrsim 10^{-18}$ G for $\lambda_{coh} \gtrsim 1$ Mpc [142, 143]. Constraints on the value of $B_{\mathrm{IGM}} \lesssim 3 \times 10^{-19}$ G can be obtained from a search for pair echos in the analysis of GRB data [163, 174], so if the larger field is correct, then no GeV echo radiation is predicted from GRBs.

5 γ Rays from Cosmic Rays in the Galaxy

Cosmic rays are the most energetic particles in the universe, and sources of

- the light elements Li, Be, B;
- the Galactic radio emission;
- the Galactic γ-ray emission;
- Galactic pressure;
- collisional excitation of atoms and molecules;
- terrestrial ^{14}C and ^{10}Be, with half lifes of ≈ 5700 years and ≈ 1.5 Myr, respectively; and
- astrobiological effects.

Cosmic rays are composed mainly of protons and ions, but also include energetic electrons, positrons, and antiprotons.[16] They make up an important particle background in the space radiation environment and contribute to the space weather. Cosmic-ray electrons emit radio and X-ray synchrotron radiation, X-ray and γ-ray bremsstrahlung radiation, and X-rays and γ rays from ("inverse") Compton scattering [207]. It is believed with good reason that Galactic GeV–PeV cosmic rays are accelerated by supernova remnants [221, 225]. The origin of the UHECRs, and its relation to the origin of the cosmic rays, is an open question; the hypothesis that their origin involves rotating black holes is developed in [4].

The difficulty to solve the problem of cosmic-ray origin for the $\sim$GeV cosmic-ray protons and ions that carry the bulk of the cosmic-ray energy content is that, being charged particles, they do not point back to their sources as a consequence of intervening magnetic fields that deflect them in transit. Sites of high-energy particle interactions are identified by γ rays, but ascertaining whether the emission is made by hadronic cosmic rays is complicated by the possible leptonic origin of most of the γ rays. GeV–PeV neutrinos, by comparison with charged cosmic rays, unambiguously point to the sources of the cosmic-ray hadronic interactions, but are faint and difficult to detect.

Cosmic-ray models use strong nuclear interaction cross sections and acceleration, loss and transport physics to derive the cosmic ray intensity and spectrum, whether in

---

[16] High-energy neutrinos could also, depending on definition, be included.

supernova remnants, diffuse and extended clouds of gas, galaxy clusters, the Galactic Center region, or wherever there is a significant column of gas with an illuminating cosmic-ray flux. The Sun and Solar flares are especially instructive for study of particle acceleration, transport, and radiation physics from γ-ray observations.

5.1 γ Rays from Solar System Objects

Besides direct observations, the effects of cosmic rays in the Solar system are traced via the cosmic-ray induced γ-ray flux of the Sun, moon, and Earth. In the Solar cavity, the cosmic-ray intensity is modulated by the outflowing Solar wind plasma, making an anti-correlated decrease and lag in the $\lesssim 10\,\mathrm{GeV}$/nucleon cosmic-ray flux reaching in the Solar cavity with a period of the 11-year sunspot cycle (one-half the 22 year Solar cycle). Interaction of these cosmic rays with Solar system objects make GeV γ-ray flux that varies on these timescales. With 3 years of data taken with Fermi, this only amounts to $\lesssim 30\,\%$ of the sunspot cycle, the first 2.5 years of which were taken with the Sun in a deep and extended Solar minimum.

Due to its proximity, the Earth is the strongest γ-ray source for the Fermi-LAT, which is why zenith-angle cuts are made on source spectral reconstruction to eliminate the interfering effects of the cosmic-ray induced γ-ray "albedo" emission (recall footnote 5). The γ-ray spectrum of the Earth albedo depends on angle to the nadir. At Fermi-LAT's $h = 565\,\mathrm{km}$ orbit, it views the horizon at $\theta_{ndr} \approx \arcsin(1 + h/r_{\mathrm{E}})^{-1} \cong 66.5°$; the Earth's radius is $r_{\mathrm{E}} = 6378.1\,\mathrm{km}$. Fermi is therefore exposed to $(1/2)(1 - \cos(\pi - \theta_{ndr})) \approx 70\,\%$ of the full sky, with the Earth occulting $\approx 30\,\%$ of the sky.

Observations of Earth albedo flux reveals a number of interesting things. One is the exposure bias toward the North Ecliptic Pole due to favored rocking to the North. Recall that Fermi rocks about zenith, early in the mission by 39°, which was increased to 50° later. Because of the intensity of the Earth albedo, a standard analysis cut is to accept photons only within 105° of zenith. For studies of the albedo, then, an acceptance of $\theta_{ndr} \lesssim 80°$ gathers mostly cosmic-ray induced terrestrial γ-ray emission, i.e., albedo. Knowing the albedo spectrum gives, empirically, a better characterization of the γ-ray background, and is useful for Fermi-LAT calibration. Deconvolving the γ-ray spectrum with an Earth atmosphere model and knowledge of the interaction cross sections would give the primary cosmic-ray spectrum and information about the deflection of primary cosmic rays by the geomagnetic field.

In [180], two data sets are gathered: one during the first 90 days of the Fermi mission, and a second for a two-orbit pointing at the Earth's limb. The analysis covers a total of 6.4×10^6 events giving the albedo spectrum in the 100 MeV–TeV range, with $218\,\mathrm{ph}(>100\,\mathrm{GeV})$, and $16\,\mathrm{ph}(>500\,\mathrm{GeV})$.

From two-dimensional intensity maps with increasing energy range, a beautiful high-energy ring forms above 3–10 GeV due not to Fermi's energy-dependent PSF but to beaming of the emission during formation (see Fig. 15). The bright limb at
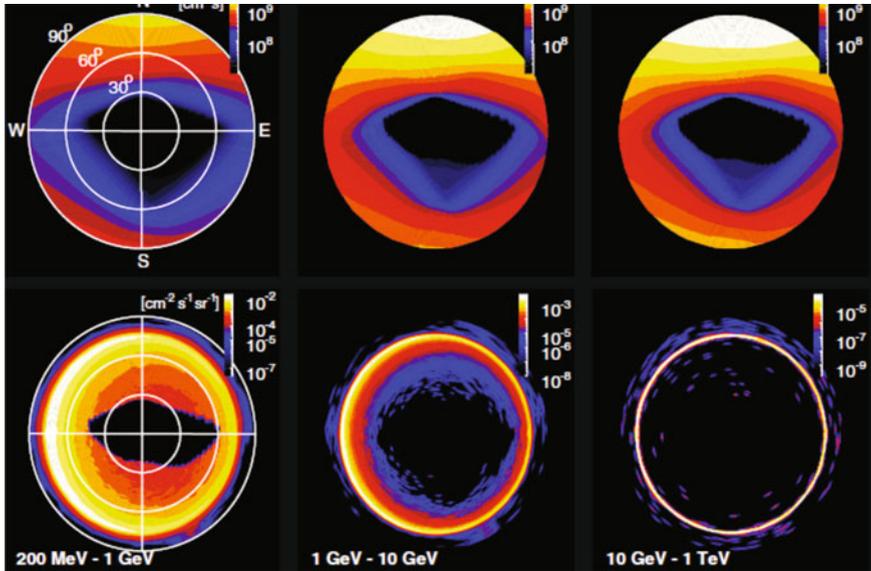
Fig. 15 Exposure maps (*top*) and intensity maps (*bottom*) of terrestrial γ-ray albedo emission [180]

the Earth's horizon are γ-ray light showers made by grazing incidence cosmic rays coming directly towards the LAT.

The next interesting feature of albedo emission is the angle-dependent spectrum. At the nadir, the spectrum is soft, and generated by γ-rays that are backscattered by large angles, which is considerably less likely due to threshold and beaming effects at $\gtrsim 1$–10 GeV/nucleon, where the particle Lorentz factor $\gamma \gtrsim$ a few. The deflection of cosmic rays by the geomagnetic field makes an east-west effect that fades out at high energies. The spectral intensity rises and hardens until the Earth's limb is reached. See [180] for details.

For the moon, without an atmosphere, a steep secondary nuclear production spectrum with flux $F_{-8} = 110 \pm 20$ [220] is made as cosmic-ray GeV protons and ions impact the surface of the moon, confirming the EGRET detection [256]. A search for γ-ray emission from asteroid populations and other Solar system rocks and dust [238, 239] is currently in progress.

The γ-ray emission from the Sun consists of two components: cosmic rays impacting the surface of the Sun to make an albedo-type emission, and cosmic-ray electrons Compton-scattering solar photons to γ-ray energies [237, 240, 241]. Both components of emission are sensitive to the phase of the Solar cycle, and both components have been detected. In analysis of 18 months of data [189], the solar disk emission is found at the level of $F_{-8} \approx 46$. The measured integral flux of the extended non-disk emission from a region of $20°$ radius centered on the Sun is $F_{-8} \approx 70$. So the Sun and moon are really bright γ-ray sources, and one has to be alert to data contamination and spurious variations when the Sun or moon drift past.
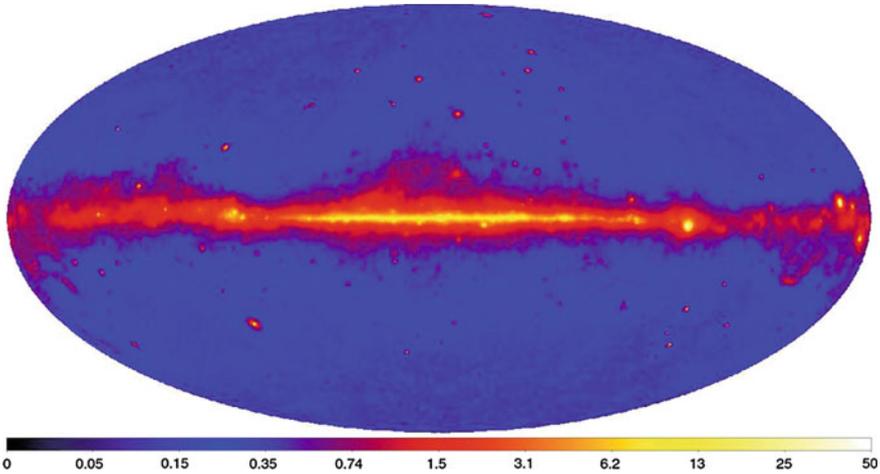
Fig. 16 Fermi-LAT 2FGL 100 MeV–10 GeV all-sky map, using 2 years of sky survey data (2008 Aug 4–2010 Aug 1) [70]. Scale in units of 10^{-7} erg/cm^2-s-sr. The 2FGL consists of 1873 sources, with 1170 associations and 127 identifications, of which 1319 are at $|b| > 10°$

The Fermi-LAT measurement of the Solar emission confirms Seckel's model [247], but at a much higher (by a factor ≈ 7) flux. The disk emission is practically flat in a νF_ν spectrum up to ≈ 8 GeV, where it begins to fall off. The observed spectrum and angular profile is in good agreement with theoretical predictions for the quiet Sun emission [237].

The first Fermi-LAT as well as nuclear γ-ray line flare in Solar Cycle 24 is the 2010 June 12 M3 flare [195]. The LAT emission lasted for only ≈ 50 s, compared to past long-duration Solar γ-ray flares. Combined Fermi GBM and LAT analysis reveals a rich γ-ray line spectrum superimposed on a continuum with a hard-energy tail that is consistent with either a separate nonthermal bremsstrahlung component, or pion production in energetic nuclear events. The hard bremsstrahlung X-rays and the ≈ 300 MeV γ rays flare up within 3 s of each other, placing strong requirements on acceleration theory and target properties. As we enter the active portion of the Solar cycle, a wealth of new Fermi data on particle acceleration physics of large Solar flares is anticipated.

Also interesting and timely is the use of Fermi for the study of terrestrial γ-ray flashes [216] that accompany thunderstorms found mainly in the sub-tropics and tropics. Experiments in nadir-pointing modes to increase sensitivity to TGFs are currently underway. Interesting analysis effects having to do with the shortness of the pulse duration and the strong soft X-ray emission affecting the ACD, which is also important for Solar flares, reveals *Fermi's* capability for this science.

5.2 GeV Photons from Cosmic Rays

The cosmic-ray induced γ-ray glow of the Milky Way is the most pronounced and distinctive feature of the Fermi sky (Fig. 16). A dominant fraction of the Galactic γ-ray emission is believed to be truly diffuse, and made by cosmic-ray bombardment of gas and dust in the interstellar medium (ISM). The most important hadronic process for γ-ray production is secondary nuclear production from the collisions of cosmic-ray protons and ions with ISM particles. The most abundant secondaries are pions (others are kaons and heavier baryons and baryonic resonances). The pions decay according to the scheme

$$\pi^0 \to 2\gamma, \ \pi^+ \to \mu^+ + \nu_\mu, \ \pi^- \to \mu^- + \bar{\nu}_\mu, \tag{114}$$

$$\mu^+ \to e^+ + \bar{\nu}_\mu + \nu_e, \ \mu^- \to e^- + \nu_\mu + \bar{\nu}_e. \tag{115}$$

For the purposes of Galactic cosmic-ray physics, the decays are essentially instantaneous, and result in γ-ray, electron, and neutrino injection emissivities (as defined below, Eq. (118)) proportional to the cosmic-ray intensity and gas density at that location. The secondary electrons and positrons, as well as those accelerated at cosmic-ray sources, lose energy during propagation mainly by synchrotron and Compton losses at high energies, bremsstrahlung losses at intermediate energies, and ionization and Coulomb losses at low energies, and can also in principle be reaccelerated.

The diffuse Galactic γ-ray glow is the superposition of all the radiations made by π^0-decay γ rays, γ rays from cosmic-ray electrons that Compton-scattered the available radiation fields, and electron bremsstrahlung γ rays. Electron synchrotron radiation would only contribute to the Galactic γ-ray emission from localized sources, such as pulsar wind nebula.

A simple expression for the demodulated cosmic-ray proton intensity in the local interstellar space inferred from measurements of the near-Earth cosmic-ray intensity is

$$J_p(E_p, \Omega_p) = 2.2 E_p^{-2.75} \text{ CR p/cm}^2\text{- s-GeV-sr} \tag{116}$$

[211], so that the GeV cosmic-ray flux represents tens of cosmic-ray protons per cm^2 per s. The cosmic-ray kinetic-energy density is dominated by the kinetic energy $T_p (= E_p - m_p c^2)$ of $\sim$GeV protons, and is given by

$$u_{CRKE} = \frac{4\pi}{c} \int_0^\infty dT_p \ T_p \ J_p(E_p, \Omega_p) \cong 0.7 \, \text{eV/cm}^3, \tag{117}$$

to which ions contribute another factor $\sim$30–50 %. The dominant elementary hadronic process is $p + p \to \pi^{\pm,0}$, so astrophysical studies of secondary nuclear production focus on this process. For more detail on radiative processes and cosmic-ray propagation, see [205, 219, 230, 251].

Background Modeling

Reliability of source detection is improved if the intensity of the diffuse background radiation is known, because a clumpy gaseous structure illuminated by cosmic-ray induced γ-ray fluorescence could be mistaken for a point source. Moreover, knowledge of the background is required to recognize dim sources, because the background dominates for all but the brightest sources. The background model for Fermi analysis takes into account diffuse Galactic γ rays from interactions of cosmic rays with material found in the various phases of the ISM, including the neutral hydrogen, HI, molecular hydrogen, H_2, ionized hydrogen, HII or H^+, and the dark gas phase [224]. The distribution of neutral atomic hydrogen (HI) is traced by 21 cm line surveys. The molecular hydrogen (H_2) distribution is derived indirectly, most commonly by using 2.6 mm line observations of carbon monoxide (CO). (The ratio of H_2 to CO—the so-called 'X'-factor, $X = N(H_2)/W_{CO}$[17] is derived from γ-ray observations.) The total atomic and molecular gas column density can also be traced indirectly from extinction and reddening by dust, depending on the relative fraction of dust and gas. Cosmic rays also interact with ionized hydrogen. The low-density ionized gas can be inferred from dispersion measures of pulsar signals in the radio band.

Cosmic rays move in large-scale galactic magnetic field and diffuse by scattering off magnetic turbulence. In the thick disk of the Galaxy, where the bulk of the diffuse γ-ray emission is made at $|b| \lesssim 5°$–$10°$, the cosmic-ray intensity changes with galactocentric distance since the source distribution is peaked at star-forming arms at 4–6 kpc. The cosmic-ray intensity also changes with the distance from the Galactic plane because of escape, and because the diffusion coefficient is energy dependent, with high-energy particles diffusing faster through the Galaxy. In addition to the CMBR, the different distributions of background optical and IR fields, and location-dependent Galactic magnetic fields, means that cosmic-ray electrons suffer position-dependent energy losses.

For the electronic component, gradients in the particle distribution can be severe and contribute a Galactic background contribution to the γ-ray emission. The GALPROP (GALactic cosmic ray PROPagation) model [242, 248–250], started in 1996 and developed independently of Fermi, determines cosmic-ray diffusion coefficients from fits to cosmic-ray data. The spatial and momentum diffusion equations for cosmic-ray transport are solved, taking into account source injection, energy and fragmentation losses, and energy changes for cosmic-ray protons, ions, and electrons. The stellar optical field, assorted IR and PAH lines in the $10\,\mu$ valley, and a FIR dust peak at $\sim 100\,\mu$ [242] provide target photons to be Compton scattered by relativistic electrons. The GALPROP model is constrained by the energy dependence of the B/C and ^{9}Be/^{10}Be ratio, from which predictions for the e^+, e^-, $\bar{p}$, and γ-ray spectra and intensity can be made.

Template modeling of γ-ray emission from Gould belt clouds in Cassiopeia and Cepheus [184] shows a weak Galactocentric gradient from the Gould belt to the

---

[17] $N(H_2)$ is the column density of molecular hydrogen, and W_{CO} is the brightness temperature of CO integrated over velocity [233].

Perseus arm, and an increase in the X-factor from $\cong 0.87 \times 10^{20}\,\mathrm{cm}^{-2}(\mathrm{K\,km\,s^{-1}})^{-1}$ in the Gould belt clouds to $\cong 1.9 \times 10^{20}\,\mathrm{cm}^{-2}(\mathrm{K\,km\,s^{-1}})^{-1}$ in the Perseus arm. The dark gas represents $\approx 50\,\%$ of the mass traced by CO.

Diffuse Galactic γ Rays from Cosmic Rays

There can be little doubt that cosmic-ray interactions make a large fraction of the γ rays observed with Fermi. This is established most clearly by Fermi-LAT observations [182] towards a region in the third quadrant between Galactic longitude $200°$–$260°$ and latitude $22°$–$60°$ that contains no known molecular clouds. After subtracting point sources and Compton emission, the residual 100 MeV–10 GeV γ-ray intensity exhibits a very strong linear correlation with atomic gas column density.

According to the model of Cordes & Lazio (2002) [209], the N(HII) column density is only $(1–2) \times 10^{20}\,\mathrm{cm}^{-2}$ and fairly smooth in the third quadrant regions analyzed in Ref. [182]. The contribution from ionized gas has a small effect on the measured emissivity, as the N(HI) column density ranges from $(1–13) \times 10^{20}\,\mathrm{cm}^{-2}$. The measured integral γ-ray emissivity measured with the Fermi LAT is $1.63 \pm 0.05 \times 10^{-26}$ photons/s-sr-H-atom and $0.66 \pm 0.02 \times 10^{-26}$ photons/s-sr-H-atom above 100 MeV and above 300 MeV, respectively, with an additional systematic error of $\sim 10\,\%$.

These numbers are explained in the first approximation if cosmic rays pervade the gaseous disk of the Milky Way with the same intensity, Eq. (116), as observed locally. The pion production rate per unit volume, or differential photon emissivity, is

$$\dot{n}_{pH\to\pi^0}(T_\pi) = 4\pi n_H \int_0^\infty dT_p\, J_p(T_p, \Omega_p)\, \frac{d\sigma_{pH\to\pi^0}(T_p)}{dT_\pi}, \qquad (118)$$

and the bolometric π^0 γ-ray production rate per H atom is, therefore

$$q_\gamma \equiv \frac{dN_\gamma}{dt dV d\Omega}/n_H \cong 2\zeta \int_{T_{p,thr}}^\infty dT_p\, J_p(T_p, \Omega_p)\sigma_{pp\to\pi^0}(T_p), \qquad (119)$$

including a factor for two photons per π^0 and a metallicity correction $\zeta \approx 1.5$. Between $E_p \approx 1.3\,\mathrm{GeV}$, just above threshold, and $E_p \approx 10\,\mathrm{GeV}$, the inclusive π^0 production cross section is approximately linear, and can be written as

$$\sigma_{pp\to\pi^0}(\mathrm{mb}) \cong 4\,\mathrm{mb}(E_p/1.3\,\mathrm{GeV}),$$

from which, with Eq. (116), we find that

$$q_\gamma \cong 2.2 \times 10^{-26} \left(\frac{\zeta}{1.5}\right)\ [\mathrm{s\text{-}sr\text{-}H\text{-}atom}]^{-1}, \qquad (120)$$

roughly agreeing with the Fermi-LAT measurements [182] quoted above.
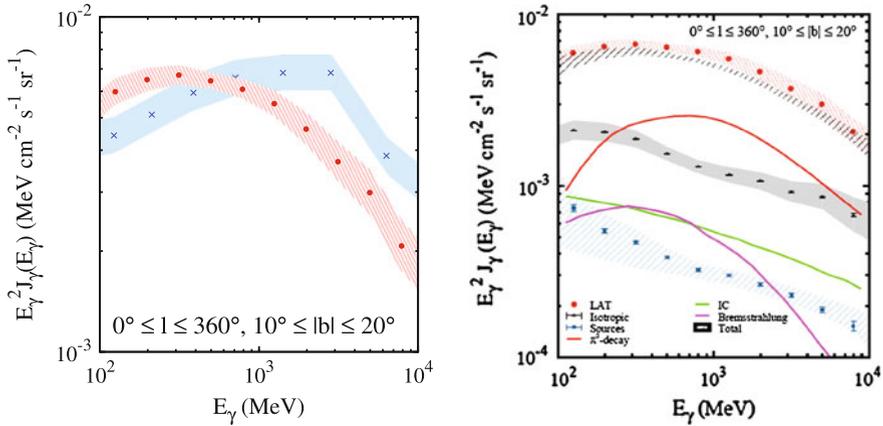
Fig. 17 Galactic diffuse emission intensity as measured in the latitude range $10° \leq |b| \leq 20°$ [42]. (*left*) LAT Galactic diffuse γ-ray intensity data are given by *red dots* and *red cross-hatched error regions*; EGRET data are given by the *blue crosses* and *blue shaded regions*. (*right*) LAT data compared to a model consisting of π^0 decay γ rays (*red*); bremsstrahlung from primary and secondary electrons and positrons (*magenta*); Compton-scattered soft photons to γ-ray energies by cosmic-ray electrons (*green*); unidentified background consisting of point sources, isotropic diffuse Galactic, and the isotropic extragalactic γ-ray background

The intermediate latitude, $10° \leq |b| \leq 20°$, Galactic diffuse γ-ray emission intensity spectra as measured with the Fermi-LAT and EGRET are shown in the left panel of Fig. 17 [42]. The prominence of the π^0 decay feature is apparent, as also are the large systematic differences at both low, $\lesssim 300$ MeV, and high, $\gtrsim 300$ MeV energies. In the right panel of Fig. 17, the LAT spectrum shown in the left panel is compared with spectra from an *a priori* model (based only on local cosmic-ray data) for the diffuse Galactic γ-ray emission, updated from GALPROP. As the Fermi LAT accumulates data, and analysis becomes even more accurate below 100 MeV for the diffuse class, we can foresee using Fermi data to determine the best interstellar cosmic-ray proton and He ion spectra rather than the other way around.

Other Fermi LAT Cosmic Ray Results

CR e spectrum: The Fermi LAT can be used as a cosmic-ray electron detector by using its anticoincidence dome to identify incoming charged particles, and distinguishing between leptons and ions from the tracks in the tracker and the calorimeter. The Fermi telescope has an acceptance of >2 m^2-sr for combined cosmic ray e$^+$ and e$^-$ (CR e), since it cannot distinguish charge directly.

The extraordinary statistics of the Fermi enabled a precise measurement of the CR e spectrum, showing a featureless spectrum consistent with a power law of number index $\cong -3.04$ between ≈ 25 and 900 GeV [183]. The CR e spectrum is

harder than a GALPROP model prediction where the diffusion coefficients tuned to the low-energy electron spectra are extrapolated to higher energies. The measurement disagrees with the ATIC report [208]. The PAMELA measurement of CR e$^+$ [197] and the cosmic-ray electron spectrum inferred at TeV energies with HESS data [201], in addition to the Fermi measurement of the cosmic-ray electron spectrum, place constraints on cosmic-ray electron models that require either (or a combination of) modifications to the propagation characteristics; local sources, most likely pulsars [259]; or contributions from dark matter (see lectures by Prof. Bergström and [243]).

The 2010 Fermi analysis gives the CR e spectrum from $\approx 7\,$GeV to 1 TeV, with a slight hardening above 100 GeV and a softening above 500 GeV [192].

EGRET excess: This term refers to EGRET measurements of the diffuse Galactic γ-ray intensity that were found, irrespective of direction, to be in excess of that predicted using the local demodulated cosmic ray spectrum and measured target gas mass [227]. Possible explanations included an

1. unusual location and local cosmic-ray spectrum measured here at Earth;
2. nuclear physics wrong;
3. addition of γ ray signal from annihilating dark matter; or
4. EGRET miscalibration.

With the launch of the Fermi Gamma ray Space Telescope, measurements [42] of the diffuse Galactic emission have been found to favor the latter hypothesis, namely, that EGRET was poorly calibrated above $\approx 5\,$GeV. The differences in the EGRET and Fermi LAT diffuse Galactic γ-ray emission are shown in the left panel of Fig. 17.

The GALPROP model decomposes the Galactic plane emission into pionic, electron bremsstrahlung and Compton fluxes, point sources, and an isotropic diffuse background. The $\pi^0 \rightarrow 2\gamma$ signature is clearly seen in the spectrum of the diffuse Galactic γ radiation [42]. Discrepancies between EGRET and Fermi spectra still remain at $\lesssim 100\,$MeV, where systematics effects become severe for both the EGRET spark chamber and the LAT tracker; see Fig. 17. The latest analysis of diffuse Galactic γ radiation in the Third Galactic Quadrant finds weak evidence at best for a Galactocentric gradient in the cosmic-ray intensity [193].

Fermi has also measured the positron flux and fraction using the Earth as a magnetic field [196].

5.3 Fermi Bubbles

It is not possible to cover, or even mention, all Fermi-LAT diffuse studies in this short review, e.g., the Galactic center region, the Cygnus Loop [228], and γ-ray emission from massive star clusters and OB associations like the cosmic-ray filled cocoon region in Cygnus [194], but we would be remiss not to mention the Fermi bubbles that D. Finkbeiner and colleagues find [212, 254]. In their analysis, backgrounds determined by template fitting are subtracted from the γ-ray data, leaving (apparently) well-defined structures—the Fermi bubbles—at ≈ 1–$10\,$GeV, and which probably

extend to much higher energies. They are symmetrically arranged north and south of the Galactic Center with a width of $\approx 40°$, and extend some $\approx 50°$ in height. They are correlated with the WMAP haze, which is excess 20–40 GHz emission found after correcting for the dipole anisotropy and subtracting out emission traced by Hα and soft synchrotron radiation extrapolated from the 408 MHz Haslam survey [215]. The 1–100 GeV luminosity of both bubbles, if emanating from the Galactic Center region, is $\approx 2 \times 10^{37}$ erg/s. With a minimum lifetime of ≈ 10 kpc/c, its minimum energy content is $\gtrsim 4 \times 10^{49}$ erg. The bubble luminosity represents $\sim 5\%$ of the 100 MeV–100 GeV luminosity of the Galaxy [252].

The haze itself, if not an artifact of template fitting or mis-extrapolation of the radio emission [234], could be formed by dust, spinning dust, or dark matter, but a synchrotron origin seems favored, especially given the Fermi bubbles. Indeed, a leptonic model for the bolometric SED of the bubbles, with synchrotron radiation for the WMAP haze and Compton-scattered CMB γ rays for the Fermi-LAT emission, can be arranged by tailoring the electron distribution [254]. A detailed model including inverse Compton scattering off the CMB, FIR, and optical/UV radiation fields fits the average LAT γ-ray spectrum [235], which is practically flat between 1 and 100 GeV. There is furthermore no energetics problem in a leptonic model, but it requires reacceleration and a low-energy cutoff in the electron spectrum. The sharp edges reported for the Fermi bubbles, to be confirmed in ongoing Fermi-LAT analyses, stand in contrast to the WMAP haze. A specific hadronic model [210] suffers from weak radiative efficiency in the dilute wind from the Galactic center, and lateral diffusion would smear the edges of the bubbles.

Of even greater interest is what the existence of the Fermi bubble says about our Galaxy. Was it a giant Galactic explosion, the residuum of a more dynamic period in the history of our Galaxy, a superwind from a past starburst episode, or something else?

5.4 γ-Ray Supernova Remnants

The cosmic-ray power P_{CR} can be estimated as

$$P_{CR} \sim \left(\frac{1 \text{ eV/cm}^3}{t_{esc}} \right) V_{MW} \sim \frac{1.6 \times 10^{-12} \text{ erg/cm}^3 \times 4 \times 10^{66} \text{cm}^3}{20 \text{ Myr}}$$

$$\sim 10^{40} \text{ erg/cm}^3, \tag{121}$$

using a residence time of ≈ 20 Myr from analysis of ^{10}Be abundance in cosmic rays [219], and the Milky Way volume $V_{MW} \sim \pi(200 \text{ pc})(15 \text{ kpc})^2 \sim 4 \times 10^{66} \text{ cm}^3$. Thus the sources of the Galactic cosmic rays need to supply $\gtrsim 10^{40}$ erg/s in the form of nonthermal particle power throughout the Galaxy. One Galactic supernova every thirty years supplying $\approx 10^{51}$ erg in outflowing kinetic energy will inject with

10 % efficiency $\approx 10^{41}$ erg/s of cosmic-ray power, so significant inefficiency can be accommodated if SNRs accelerate the cosmic rays. For the cosmic-ray power implied by Fermi-LAT γ-ray observations of the Milky Way for leptonic and hadronic cosmic-ray models, see [252].

The substantial time-averaged kinetic powers of SNRs make them the favored candidate source population for the hadronic cosmic rays in our Galaxy. This power is greater than the time-averaged power available from, for example, pulsars, galactic X-ray binaries, stellar winds, or novae. In the electromagnetic window, the confirming signature of enhanced cosmic-ray activity is the π^0 γ-ray bump [221, 225] peaking at $m_{\pi^0}/2 \approx 70$ MeV in a number spectrum, and at several hundred MeV (depending on the spectrum) in a νF_ν representation. The confirming signature of π^0 production at $\lesssim 200$ MeV competes with systematics and inaccuracies in the background model which rise due to the large LAT PSF in the lowest decade of its energy range (Table 1).

Nearly 300 SNRs are known, mostly through radio detections. The Chandra X-ray catalog contains nearly 100 SNRs,[18] and ROSAT Galactic SNRs must number several dozen. The poor EGRET PSF made it impossible to identify specific SNRs with sites of enhanced >100 MeV emission, though several associations could be made, including IC 443, W28, and W44 [214, 253, 257].

In the 1FGL, 41 SNRs are associated with LAT γ-ray sources, and 3 are, through morphological features, identified with SNRs (Table 2), namely W51C, W44, and IC 443. By the time of 2FGL, 62 associations and 9 (morphological) identifications of SNRs and PWN with Fermi LAT and other wavelengths have been made (Table 2). The identifications represent 6 SNRs, adding also W28, W30, and the Cygnus Loop [228], and three PWNe, namely J0835.3-4510 (Vela X), J1509.6-5850 (MSH 15-52; possibly the remnant of a supernova in 185 CE), and J1826.1-1256 (HESS J1825-137). A fourth PWN, MSH 15-52, was found in a search of off-pulse emission of PSR B1509-58 [188].

Some important questions to be answered are

- Are cosmic-ray protons and electrons accelerated at SNR shocks, and with what relative efficiencies?
- With what efficiency is shock kinetic energy converted to cosmic-ray energy?
- What is the spectrum of accelerated particles escaping from the shock acceleration site?
- What is the maximum particle energy, and do SNRs accelerate CRs up to or beyond the knee of the cosmic-ray spectrum?
- Does magnetic field amplification take place in the vicinity of SNRs?

Fermi has now detected historical/young ($\lesssim 3000$ years; see Table 8) SNRs, namely Cas A (1680 CE), Tycho (1572 CE), and RX J1713.7-3946 (393 CE), intermediate age ($\approx 10^4$ years) SNRs, e.g., IC 443 and W51C, and middle-aged ($\gtrsim 2 \times 10^4$ years) SNRs, including W44, W28, the Cygnus Loop, and G349.7+0.2. All SNRs detected with Fermi, other than Cas A, show evidence for molecular cloud interactions.

---

[18] http://hea-www.harvard.edu/ChandraSNR/

Table 8 Historical supernovae [223]

SN	Type[a]	Distance (kpc)	Size[b]	γ-ray detection[c] GeV/VHE	Pulsar[d]
1987A	II	55	0.8″		
1680; Cas A[e]	II	3.4	5.9′ × 5.5′	Y/y	
1604; Kepler	Ia	5	4.5′ × 3.7′		
1572; Tycho	Ia	2.4	8.7′ × 8.6′	Y/y	
1181; 3C 58 (?)	II	2.6	6′ × 3′		J0205+6449
1054; Crab	II	2	2.3′ × 2.1′	y/Y	Crab
1006	Ia	1.6	32′ × 26′	?/Y	
393; RX J1713.7-3946	II (?)	1	≈40′ × 40′	y/Y	
386?					
369?					
185; MSH 15-52 (?)	II			y/Y	1509-58

[a][231, 232]
[b]Chandra X-ray sizes[16], except for SN 1987A (HST), RX J1713.7-3946 [200], and 3C 58 (radio; [244])
[c]Flux measured at GeV or VHE energies; Y: larger νF_ν flux, y: smaller νF_ν flux
[d]See Table 3
[e]Detection by Flamsteed disfavored [223]

Detecting the π^0 decay feature in Fermi-LAT SNR spectra depends on good background modeling to expose the low-energy ($\lesssim 200\,$MeV) flux. Shock compression of pre-existing cosmic rays could be argued to account for apparent cosmic-ray production at SNR shocks, but this seems less plausible now that we have measured GeV γ-ray luminosities ranging from $\approx 10^{33}$ erg/s for the Cygnus loop, seen as an $\approx 1°$ radial loop and the largest resolved γ-ray SNR, to $\approx 10^{34}$–10^{35} erg/s for young SNRs, and $\approx 10^{35}$–few $\times 10^{36}$ erg/s for SNRs showing strong molecular cloud interactions [258]. This implies cosmic-ray kinetic energy on the order of $\mathcal{E}_{CR}(\text{erg}) \sim 10^{36} L_{36} t_{pp} \cong 10^{51} L_{36}/n(\text{cm}^{-3})$, in accord with the SNR hypothesis for cosmic-ray origin. This estimate is not, however, so strong, given that distance uncertainties can easily lead to factor of 4 or more uncertainty in L.

For Cas A, one of the youngest SNRs detected with Fermi [187], combined VHE (MAGIC, VERITAS, and HEGRA) data show a νF_ν peak at a few GeV, with ambiguous evidence for a low-energy cutoff, but clear evidence for a TeV steepening (see Fig. 18). Cas A, with an extent of $\approx 6'$, appears almost as a point source for the LAT. The somewhat older SNR RX J1713.7-3946 is famous for being the first SNR for which a VHE (HESS) map was made [199]. The TeV emission is extremely well-correlated with X-ray maps, e.g., HESS and Suzaku [255]. The Fermi-LAT map [190] displays an asymmetry towards the north, in the direction of a molecular cloud where the X-ray and TeV emission is enhanced. The GeV spectrum itself is remarkably hard, with spectral index $= 1.5 \pm 0.1$. The νF_ν spectrum of this remnant peaks between $\approx 200\,$GeV and 1 or 2 TeV, with a rapid cutoff at higher energies. Nevertheless, the spectrum of RX J1713.7-3946 extends to $\approx 100\,$TeV, making it the source of the most energetic γ rays yet detected. Its hard spectrum cannot be made
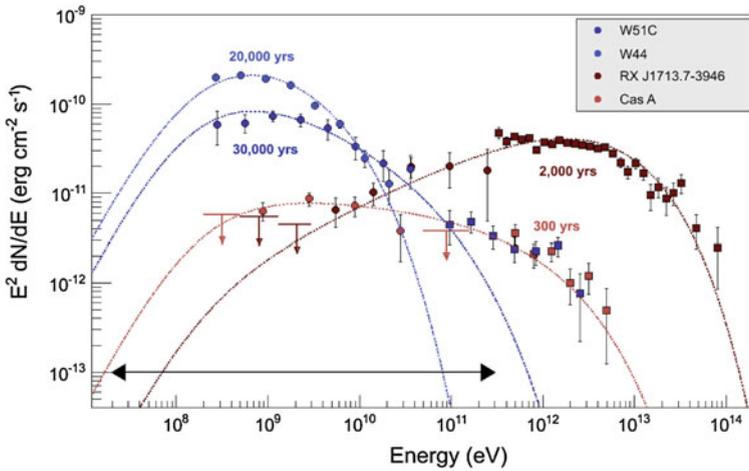
Fig. 18 Combined Fermi-LAT and TeV spectra of selected SNRs [218]. The *double arrow* is the Fermi energy range

by conventional hadronic models. One-zone leptonic models can fit the X-ray/γ-ray spectra of RX J1713.7-3946, as can two-zone models.

The prototypical Type Ia Tycho SNR has also been recently reported as a GeV source with a GeV–TeV γ-ray spectrum implying, if hadronic, a proton spectrum with a -2.3 number index [222]. Claims have been made that the combined LAT/VERITAS spectrum requires hadrons [236], but a two-zone leptonic model with both bremsstrahlung and Compton processes can fit the spectrum as well [202]. Because of the weakness of the GeV/TeV γ-ray flux, the νF_ν peak of Tycho can be constrained to be $\lesssim$ few GeV, and the spectrum is not well-resolved in the crucial energy range $\lesssim$300 MeV where the π^0 bump might be evident. For this purpose, a better SNR might be IC 443 [185], a 3–4 kyr old core collapse SN (like Cas A and RX J1713) surrounded by clouds of gas. With an 0.75° radio radius, it is resolvable with Fermi. The SED of IC 443 is a flat νF_ν spectrum between $\approx$300 MeV and 5 GeV with a peak at 1 or 2 GeV. The spectrum extends to VERITAS/MAGIC VHE energies like a power-law with νF_ν slope equal to ≈ -0.6.

The middle-aged shell SNR W44 [186] has an age of $\approx$20 kyr. The γ-ray shape closely following the 4.5 μ Spitzer IR image which traces shocked H_2. Its spectrum peaks at 1 GeV or less, and it falls off sharply at higher energies, so is not detected with the VHE telescopes. Its physical extent precludes the γ-ray emission from originating from a pulsar, though a pulsar, B1853+01 with age $\approx$20 kyr, is found in this SNR and could be associated with W44. If so, W44 is probably the result of a core-collapse SNR, and therefore Type II.

As a final example of Fermi-LAT observations of SNRs, consider W51C, the first resolved GeV SNR at GeV energies. It is an $\approx$10 kyr old remnant with extended GeV emission compatible with the location and shape of the ROSAT emission [181].

The spectrum of W51C is similar to that of W44, peaking below ≈ 1 or $2\,\text{GeV}$, and falling off steeply at higher energies. HESS weakly detects emission at about $1\,\text{TeV}$ from W51C, (whereas W44 is not detected at VHE energies), implying a νF_ν spectral index ≈ -0.5 between $10\,\text{GeV}$ and $1\,\text{TeV}$.

In terms of cosmic-ray origin theory, a few remarks can be made. First, almost all Fermi-LAT SNRs, other than Cas A, show evidence for interactions with molecular clouds, in particular, those with OH ($1720\,\text{MHz}$) maser emission from the OH hydroxyl molecule tracing dense, shocked gas [217, 226]. Out of 24 known maser SNRs reported in [226], 10 have GeV and/or TeV associations, and 6 have both. Detection of illuminated molecular cloud complexes in front of the SNR shock could reveal the existence of runaway cosmic rays that more likely would be protons than electrons.

Second, nonthermal cosmic-ray protons lose energy on the secondary nuclear production timescale $t_{pp} = (n\sigma_{pp}c)^{-1} \approx 35\,\text{Myr}/n(\text{cm}^{-3})$, whereas nonthermal electrons lose energy through bremsstrahlung on the free-free timescale $t_{ff} = (n\sigma_{ff}c)^{-1} \approx 6 \times 35\,\text{Myr}/n(\text{cm}^{-3})$, where $\sigma_{ff} \approx \alpha_f \sigma_T \cong \sigma_{pp}/6$. In either case, the process is radiatively inefficient unless the target density $n \gg 1\,\text{cm}^{-3}$, but the nonthermal electron bremsstrahlung is nearly as efficient as nuclear production processes. The nonthermal bremsstrahlung or secondary nuclear production model face energetics problems if they requires $\gtrsim 10^{50}\,\text{erg}$ in cosmic rays.

Third, ionization/Coulomb losses can harden a nonthermal electron spectrum, making γ rays from a leptonic bremsstrahlung model masquerade as a π^0 feature. A full spectral model for multiwavelength production from SNRs requires, most generally, particle acceleration at the forward and reverse shocks, zones of different magnetic field strength at which particle acceleration can occur, leptonic bremsstrahlung and Compton scattering, and secondary nuclear production.

Perhaps there is a trend in the SEDs of SNRs that can reveal the likely γ-ray production mechanism (Fig. 18). For leptonic models, both (i) nonthermal bremsstrahlung and (ii) Compton scattering, are potentially feasible. For hadronic models, only (iii) secondary nuclear/particle process is effective in the SNR environment. Confrontation of these three nonthermal processes with multiwavelength SNR data imply power and spectral and morphological constraints that can in principle identify the dominant radiation process, though in practice, this has not been so simple. But establishing a trend from hard to soft γ-ray spectra with age, even scaled to the Sedov age, is oversimplified, as is evident from the rather soft spectrum of Tycho.

For useful studies, a plot of the VHE ($100\,\text{GeV}$–$10\,\text{TeV}$) flux divided by the GeV ($300\,\text{MeV}$–$30\,\text{GeV}$) flux versus SNR age (to avoid uncertainties in distance measurements) for Type Ia and Type II SNe might provide some insight on the evolution of particle acceleration with remnant age. At present, it still seems premature to claim that the problem of the origin of the Galactic cosmic rays is solved.

Fig. 19 Shock geometry in rest frame of shock. In this figure, δ is the shock width
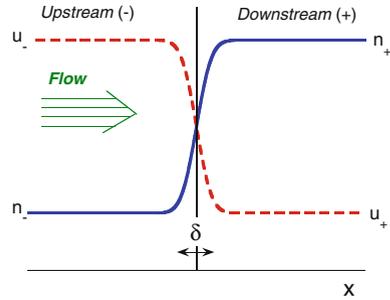
5.5 Nonrelativistic Shock Acceleration of Electrons

First-order Fermi acceleration is highly developed as a mechanism to accelerate cosmic rays [206, 213]. Here we illustrate this process by calculating the nonthermal spectrum of test particles accelerated at a shock-discontinuity in density and velocity. We use a continuity equation approach,[19] which complements the approach to shock acceleration using the convection-diffusion equation and probability arguments [4].

The geometry we consider is particle acceleration at a discontinuity in velocity and density, as illustrated in Fig. 19. The upstream unshocked material has speed u in a frame where the supersonic flow intercepts gas at rest. For a strong nonrelativistic shock with compression ratio $\chi = 4$, the upstream $(-)$ flow approaches with speed $u_- = \beta_- c = 4\beta c/3$ and the downstream $(+)$ flow recedes with speed $u_+ = \beta_+ c = \beta c/3$ in the comoving primed frame stationary with respect to the shock. Consequently $u = u_- - u_+ = \beta c$. Here we treat the acceleration of relativistic nonthermal particles with Lorentz factor $\gamma \gg 1$ and speed $\beta_{par} c \approx c$.

In the general case, the compression ratio

$$\chi = \frac{u_-}{u_+}. \tag{122}$$

For a lossless adiabatic shock wave,

$$\chi = \frac{\hat{\gamma} + 1}{\hat{\gamma} - 1 + 2/\mathcal{M}_1^2} \quad \overset{\rightarrow}{\hat{\gamma} = 5/3} \quad \frac{4}{1 + 3/\mathcal{M}_1^2} \tag{123}$$

(Eq. (13.8), [4]), where $\hat{\gamma}$ is the adiabatic index, and $\hat{\gamma} = 5/3$ for a nonrelativistic monatomic gas. When $\mathcal{M}_1 \gg 1$ and $\hat{\gamma} = 5/3$, $\chi \to 4$. It follows from Eq. (122) and the condition $u = u_- - u_+$ that

---

[19] The treatment of [191] assumes energy-independent escape timescale representing some second-order Fermi acceleration scenarios.

$$u_- = \frac{\chi u}{\chi - 1}, \quad u_+ = \frac{u}{\chi - 1}, \tag{124}$$

so that in the limit $\mathcal{M} \gg 1$, $u_- \to 4u/3$ and $u_+ \to u/3$ in the frame comoving with the shock.

The change of energy of a relativistic particle with dimensionless momentum $p = \beta\gamma = \sqrt{\gamma^2 - 1} \to 1$, when $\gamma \gg 1$, in a complete cycle of Fermi acceleration is, from Eq. (12.10) [4], given by

$$\left(\frac{\Delta\gamma}{\gamma}\right)_{F1} \cong \frac{4}{3}\beta, \tag{125}$$

where $u = \beta c = u_- - u_+$. For second-order Fermi acceleration,

$$\left(\frac{\Delta\gamma}{\gamma}\right)_{F2} \cong \frac{4}{3}\beta^2, \tag{126}$$

and both expressions receive an added boost $\propto \Gamma^2$, at least in the first cycle of acceleration, for shock acceleration by a relativistic flow with speed $\sqrt{1 - 1/\Gamma^2}c$.

The rate of energy gain, or acceleration rate, in nonrelativistic first-order Fermi acceleration is given for particles

$$\dot{\gamma}_{\mathrm{FI}} \cong \frac{\Delta\gamma}{t_{cyc}}, \tag{127}$$

where the cycle time t_{cyc} is given by

$$t_{cyc} = \frac{4}{v}\left(\frac{\kappa_-}{u_-} + \frac{\kappa_+}{u_+}\right) = \frac{4}{vu_-}(\kappa_- + \chi\kappa_+)$$
$$= \frac{mc^2}{QB_-}\frac{4p}{3u}\frac{\chi - 1}{\chi}\left(\eta_- + \chi\eta_+\frac{B_-}{B_+}\right) \cong \frac{mc^2}{QB_-}\frac{4\gamma}{3u}\frac{\chi - 1}{\chi} \propto \gamma. \tag{128}$$

[229], where $v \cong c$ is the particle speed, $\chi = u_-/u_+ = \rho_+/\rho_-$ is the compression ratio, and $\rho = mn$ is the mass density. The spatial diffusion coefficient

$$\kappa_\pm = \frac{1}{3}\lambda_\pm v = \frac{1}{3}\eta_\pm r_{\mathrm{L}\pm}v = \frac{1}{3}\eta_\pm r_{\mathrm{L}\pm}^o pv,$$

here writing the diffusion coefficients in terms of the parameters $\eta_\pm$ that give the particle mean-free-paths scaled to the values in the local magnetic field. In the Bohm diffusion approximation, the diffusion mean-free path is set equal to the Larmor radius, so $\eta_\pm = 1$. Thus the Bohm approximation is

$$\kappa_{B\pm} = \frac{1}{3}r_{\mathrm{L}\pm}v,$$

where the Larmor radius

$$r_{L\pm} = r^o_{L\pm} p = \frac{mc^2}{QB_\pm}\, p. \tag{129}$$

The acceleration rate of nonthermal relativistic particles by first-order shock acceleration can therefore be expressed as

$$\dot{\gamma}_{acc} = \frac{4\beta}{3t_{cyc}}\gamma = \frac{4}{3}\frac{u}{ct_{cyc}}\gamma. \tag{130}$$

The acceleration rate is independent of p or γ because $t_{cyc} \propto \gamma$ from Eq. (128).

In nonrelativistic shock acceleration, the escape probability per cycle is

$$\mathcal{P}_{esc} = 4\beta_+ = 4\frac{u_+}{c} = \frac{4u}{c(\chi-1)}. \tag{131}$$

[203, 204]. The escape time is therefore

$$t_{esc} = \frac{t_{cyc}}{\mathcal{P}_{esc}} = \frac{t_{cyc}}{4\beta_+}. \tag{132}$$

The steady-state particle continuity equation with gains, losses, and escape takes the form (Eq. (C10), [4])

$$\frac{\partial}{\partial\gamma}[\dot{\gamma}\, n(\gamma)] + \frac{n(\gamma)}{t_{esc}(\gamma)} = \dot{n}(\gamma). \tag{133}$$

Equation (133) has solution, for an energy-gain process $\dot{\gamma} > 0$, given by (Eq. (C11), [4])

$$n(\gamma) = \dot{\gamma}^{-1} \int_1^\gamma d\gamma'\, \dot{n}(\gamma') \exp\left[-\int_{\gamma'}^\gamma \frac{d\gamma''}{t_{esc}(\gamma'')\dot{\gamma}(\gamma'')}\right]. \tag{134}$$

Equations (130) and (132) imply

$$t_{esc}(\gamma)\dot{\gamma}_{acc}(\gamma) = \frac{\beta}{3\beta_+}\gamma = \frac{\chi-1}{3}\gamma, \tag{135}$$

so

$$n(\gamma) = \frac{3\dot{n}_0}{4\beta\gamma} t_{cyc} \left(\frac{\gamma}{\gamma_0}\right)^{-3/(\chi-1)}. \tag{136}$$

noting that this represents the differential number density spectrum of particle accelerated at the shock discontinuity.

Taking a δ-function source injection $\dot{n}(\gamma) = \dot{n}_0\delta(\gamma'-\gamma_0)$, the injection spectrum downstream where the particles are no longer subject to acceleration at the shock is

$$\dot{n}_{ds}(\gamma) = \frac{n(\gamma)}{t_{esc}} = \frac{4\beta_+}{t_{cyc}} n(\gamma) = \frac{3\dot{n}_0}{(\chi - 1)\gamma_0} \left(\frac{\gamma}{\gamma_0}\right)^{-A_{tp}}, \qquad (137)$$

where

$$A_{tp} = \frac{u_- + 2u_+}{u_- - u_+} = \frac{2 + \chi}{\chi - 1} \qquad (138)$$

is the test-particle number index well known in studies of nonrelativistic shock acceleration [206].

Now consider shock acceleration with the addition of Thomson/synchrotron radiative losses, so

$$\dot{\gamma} = \dot{\gamma}_{acc} + \dot{\gamma}_{rad} = \frac{4}{3}\frac{\beta}{t_{cyc}}\gamma - \nu\gamma^2. \qquad (139)$$

Following the approach above gives

$$n(\gamma) = \frac{\dot{n}_0}{\left(\frac{4}{3}\frac{\beta\gamma}{t_{cyc}} - \nu\gamma^2\right)} \exp\left[-\frac{3}{\chi - 1}\int_{\gamma_0}^{\gamma}\frac{dx}{x(1 - kx^2)}\right], \qquad (140)$$

where

$$k = \frac{3\nu T}{4\beta}, \text{ and } T = \frac{m_e c^2}{Q B_-}\frac{4}{3u}\left(\frac{\chi - 1}{\chi}\right)\left(\eta_- + \chi\eta_+ \frac{B_-}{B_+}\right). \qquad (141)$$

Therefore

$$\dot{n}_{ds}(\gamma) = \frac{n(\gamma)}{t_{esc}} = \frac{3\dot{n}_0}{(\chi - 1)(1 - k\gamma^2)^{3/2}}\left(\frac{\gamma}{\sqrt{1 - k\gamma^2}}\right)^{-A_{tp}}. \qquad (142)$$

This is a pileup spectrum for $\chi > 5/2$, that is, when A_{tp} is harder than 3. The maximum electron energy is given by $1 - k\gamma^2 = 0$, or

$$\gamma_{max} = \frac{1}{\sqrt{k}} = \sqrt{\frac{4\beta}{3\nu T}}. \qquad (143)$$

Even though the spectrum piles up, the number of particles and the energy of the accelerated particles is convergent.

The synchrotron emission made by this electron distribution has a maximum value at dimensionless photon energy

$$\epsilon_{syn,max} = \frac{3}{2}\frac{B}{B_{cr}}\gamma^2 = \frac{9\pi e}{B_{cr}\sigma_T}\left(\frac{\chi}{\chi - 1}\right)\beta^2\frac{(B_- B_+/B^2)}{\left(1 + \frac{u_{ph}}{u_B}\right)(\eta_- + \chi\eta_+\frac{B_-}{B_+})} \qquad (144)$$

(cf. [198]). Thus

$$\epsilon_{syn,max} = \frac{27}{8\alpha_f} \left(\frac{\chi}{\chi - 1} \right) \beta^2 \frac{B_-/B_+}{\left(1 + \frac{u_{ph}}{u_B} \right) \left(\eta_- + \chi\eta_+ \frac{B_-}{B_+} \right)}. \tag{145}$$

The apparent divergence due to the term $\chi - 1$ in the denominator of Eq. (145) can be seen not to arise, considering that $\chi u^2/(\chi - 1) = u_- u$. The leading term, $\epsilon_{syn,max} \cong 27/8\alpha_f \cong 462$, or $E_{syn,max} \cong 236$ MeV, represents a bound for nonrelativistic shock acceleration, remarkably close in value to the cutoff energy of the Crab pulsar wind nebula [48]. This indicates that the pulsar wind nebula of the Crab is formed by a wind termination shock moving out at mildly relativistic velocities. Small changes in the bulk speed from a knot in the Crab pulsar wind, as imaged with Chandra, could produce the Fermi and AGILE flares from the Crab [52]. Better imaging at ≈ 100 MeV would localize the emission source, but is hardly possible with Fermi.

Addition of a diffusion term will produce smoothed, realistic pile-up electron injection distributions formed in first-order shock acceleration. The addition of a diffusion term is under study in work with P. Becker. Note that these pileup functions differ from the Schlickeiser pile-ups [245, 246] where escape is always independent of particle energy. At relativistic energies and with relativistic flows, the maximum dimensionless synchrotron energy is $\epsilon_{syn,max} \cong (27/8\alpha_f)\Gamma/(1 + z)$, which is relevant when interpreting the maximum photon energies in GRBs as a consequence of synchrotron emission formed by particle acceleration at an external shock.

6 γ Rays from Star-Forming Galaxies and Clusters of Galaxies, and the Diffuse Extragalactic γ-Ray Background

6.1 γ Rays from Star-Forming Galaxies

Galaxies with ongoing star formation, most notably the Milky Way in which we live, are illuminated at γ-ray energies by secondary products of cosmic-ray interactions with gas and dust. The characteristic $\gtrsim 100$ MeV γ-ray luminosities of normal star-forming galaxies are $\sim 10^{38}$–10^{40} erg s^{-1}, some factors of $\sim 10^3$–$10^{10} \times$ smaller than those of active galaxies. Yet the star-forming galaxies vastly outnumber the AGNs. For example, the space density of a typical L$_*$ spiral galaxy like the Milky Way is $\approx 3 \times 10^6$–10^7 Gpc^{-3}, by comparison with the space density of FR II radio galaxies, which is ≈ 2000 Gpc^{-3} [278]. (The volume of 1 Gpc^3 extends to about $z = 0.15$ from the present epoch.)

Besides the Milky Way, the Large Magellanic Cloud, detected earlier with EGRET with flux $F_{-8} \cong 19$, is now measured with Fermi LAT at the level of $F_{-8} \cong 26 \pm 2$ [260]. At a distance of ≈ 50 kpc, the LMC has about 10 % of the mass and a supernova rate ≈ 20 % of the Milky Way. The LAT resolves the LMC, and finds that 30 Doradus, its major star-forming region, is a bright source of γ rays that does not consist of significant point source contributions. The γ-ray spectrum is consistent with an

origin in cosmic-ray production. Its γ-ray emission correlates well with massive star forming regions and an ionized H^+ template, but more poorly with neutral or molecular gas distribution. The γ-ray emission morphology is surprisingly compact, and indicates that cosmic rays are accelerated in star-forming regions that are not very diffusive, thus accounting for the bright compact emission centered around 30 Doradus.

Our other notable dwarf companion galaxy, the Small Magellanic Cloud, is also for the first time detected in γ rays. The LAT measures a flux of $F_{-8} = 3.7 \pm 0.7$ [261] from an extended, $\sim 3°$ region. Unlike the LMC, the γ-ray emission from the SMC is not clearly correlated with the distribution of massive stars or supernova remnants, though the emission may trace supergiant shells.

Though not unexpected [284], the discovery of starburst galaxies at GeV [262] and TeV energies is important to test predictions of cosmic-ray origin. The two nearest starburst galaxies, M82 and NGC 253, each at a distance of ≈ 4 Mpc, were detected at a level of $F_{-8} \cong 1.6$ and $F_{-8} \cong 0.6$, respectively [262], with the LAT, soon after being reported as VERITAS [266] and HESS [267] sources, respectively. Properties of the brightest star-forming galaxies at γ-ray energies are shown in Table 9. Although correlations with mass, SN rate, and their product can be extracted, the situation, as exhibited by the distribution of γ-ray emission in the LMC, is far more complex. As can be seen from Fig. 7, starburst galaxies define a separate track in the spectral index versus luminosity plane.

Now that the LAT sensitivity for star-forming galaxies has reached beyond our neighboring galaxies, we can expect the number to grow. The Andromeda galaxy, M31, has been detected at a level of $F_{-8} \cong 0.9$ [265]. Other starburst and infrared luminous galaxy detections from sources like Arp 220 [287] or Mrk 273 are keenly anticipated as the LAT photon statistics accumulate.

Table 9 Properties of star-forming galaxies

Galaxy	d (kpc)	R_{SN} (century^{-1})	M_{Gas} ($10^9 M_\odot$)	F_{-8}[a]	$4\pi d^2 F_\gamma$ (10^{41} ph/s)	L_γ^b (10^{39} erg/s)	Index
MW	–	2.0 ± 1.0	6.5 ± 2.0	–	11.8 ± 3.4^c	1.2 ± 0.5	2.2 ± 0.15
LMC	52 ± 2	0.5 ± 0.2	0.67 ± 0.08	26.3 ± 2.0	0.78 ± 0.08	0.041 ± 0.007	2.26 ± 0.11
SMC	61 ± 3	$\approx 0.12^d$	≈ 0.45	3.7 ± 0.7	0.16 ± 0.04	0.008 ± 0.003	2.23 ± 0.12
M31	780 ± 30	1.1 ± 0.2	7.7 ± 2.3	0.9 ± 0.2	6.6 ± 1.4	0.43 ± 0.09	2.1 ± 0.22
M82	3600 ± 300	20 ± 10	2.5 ± 0.7	1.6 ± 0.5	250 ± 90	13 ± 5	2.2 ± 0.2
N253	3900 ± 400	20 ± 10	2.5 ± 0.6	0.6 ± 0.4	110 ± 70	7.2 ± 4.7	1.95 ± 0.4

References [252, 261, 262]

[a] $F_\gamma = 10^{-8} F_{-8}$ ph (>100 MeV)/cm^2-s

[b] Fluxes and luminosities in 100 MeV–5 GeV range

[c] Value is strongly dependent on assumed size of Galactic halo used to model MW; [252] find $0.6 \lesssim F_{-8} \lesssim 1.0$

[d] Reference [286]

6.2 γ Rays from Clusters of Galaxies

Clusters and superclusters of galaxies are the largest manifestations of the ongoing structure formation. Regions collapse by gravitation, driven primarily by dark matter fluctuations. The magnitude of the density fluctuation determines the formation time, insofar as larger structures form by accreting smaller clumps in a hierarchical merging scenario. The result, as seen in numerical simulation codes, is lumpy, continuous accretion forming filaments and webs of enhanced density and magnetic field.

Clusters of galaxies are the most energetic events in the universe since the big bang, releasing as much as $GM_1 M_2/R \sim 10^{63}$–10^{64} erg, here taking $M_1 \sim 10^{15} M_\odot \sim 10 M_2$ and $R \sim 1$ Mpc. Yet they are hardly the most luminous, since the energy is released over a Hubble timescale $H_0^{-1} \sim 4.3 \times 10^{17}$ s ~ 13.6 Gyr, implying a power of $\sim 2 \times 10^{45}$–2×10^{46} erg/s. This energy, carried primarily in the gravitational potential of the dark matter, goes into heating, turbulence, and gas motions. In the events of cluster mergers, turbulence is generated in wakes, and shocks are formed in the collision. Nonthermal particles are accelerated by shocks and turbulence.

Evidence for nonthermal particle acceleration in merging clusters of galaxies is already known from the existence of radio halos and relics. The relativistic electrons and positrons making this radio emission are either accelerated directly or produced as secondary pairs by cosmic-ray protons and ions colliding with particles in the intracluster medium. The unpolarized central *radio halos*, $\sim$ Mpc size, with a morphology similar to the X-ray bremsstrahlung, reflect particle acceleration at the merger shocks between two merging clusters, intermittent AGN activity, and dynamical friction of galaxies in the hot, $\sim$keV, thermal plasma. By contrast, *radio relics*, also about $\sim$1 Mpc in size, which lie on the cluster outskirts, display elongated morphologies, and are up to $\sim$50 % polarized, may reflect a different origin from the radio halos. Most likely, they represent ongoing accretion of pristine gas from the big bang.

The halo radio emission of, e.g., the Coma cluster of galaxies, is very soft above $\approx$1 GHz, and can be modeled in a merging cluster framework if the system is observed soon after the cluster and subcluster have merged [270]. By scaling nonthermal cosmic-ray proton and ion energy to the electron energy required to make the observed radiation, predictions for the γ-ray emission can be made. The identification of a nonthermal hard X-ray feature as Compton-scattered CMBR would mean a larger nonthermal electron total energy, giving more optimistic predictions for γ-ray emission from clusters of galaxies. In any case, clusters of galaxies act as storage volumes for the cosmic-ray protons [271], which have radiative lifetimes exceeding the Hubble time. Because of the long, $\sim$10 Gyr, crossing time of electrons, versus a radiative lifetime of $\sim$0.1 Myr, in-situ acceleration of pre-existing relativistic electrons by turbulence [272, 273] is now favored to explain radio halos and relics.

The Fermi LAT has many high-priority cluster candidates, including not only Coma and Abell 2256, but Hydra, Centaurus, and Fornax. The Perseus cluster is dominated by the AGN 3C 84 in the central elliptical galaxy NGC 1275, and is

detected with the Fermi LAT [264]. But this γ-ray emission is from a variable blazar core (though there may be steady emission at the level of $F_{-8} \lesssim 4$). Radiation mechanisms involving the energetic leptons together with the decay of neutral pions produced by hadronic interactions have the potential to produce abundant GeV photons. Using data from 2008 August to 2010 February, upper limits of 33 galaxy clusters, selected according to their proximity, mass, X-ray flux, temperature, and non-thermal activity were reported by the Fermi collaboration [268]. The flux upper limits, in the 0.2–100 GeV range, are typically at the level of $F_{-8} \cong 0.1$–5. These results limit the relativistic-hadron-to-thermal energy density ratio to be $\lesssim 5$–10 % in several clusters [274, 281].

As disappointing as it has been that γ rays have not yet been detected from galaxy clusters, we are still learning about the injection conditions in the low Mach number merger shocks, the high Mach number accretion shocks, and the injection conditions of turbulence models. A purely secondary production model where all the radio emission results from secondary electrons formed in secondary nuclear production is likely ruled out [268, 275], and dark matter annihilation cross sections are further constrained.

6.3 Extragalactic γ-Ray Background and Populations

Figure 20 shows the spectrum of the diffuse extragalactic γ-ray background (EGB) obtained in an analysis of the first 10 months of Fermi-LAT science data in the range 200 MeV–100 GeV based on the GALPROP model for the Galactic emission [7], alongside the EGRET EGB [26] and the EGRET EGB based on a GALPROP analysis of the Galactic diffuse EGRET emission [28]. The Fermi-LAT intensity extrapolated to 100 MeV based on the power-law fit $I(>100\,\mathrm{MeV}) = (1.03 \pm 0.17) \times 10^{-5}$/cm^2-s-sr is significantly lower than that obtained from EGRET data, namely $I_{\mathrm{EGRET}}(>100\,\mathrm{MeV}) = (1.45 \pm 0.05) \times 10^{-5}$/cm^2-s-sr [26]. The GALPROP analysis of the EGRET data [28] agrees, however, with the Fermi flux extrapolated to 100 MeV. Furthermore, the Fermi-LAT spectrum is compatible with a featureless power law with index -2.41 ± 0.05 [7]. This is significantly softer than the EGRET spectrum, with index -2.13 ± 0.03 [26]. To check that the different spectra are not due to the instrumental point-source sensitivities, a threshold flux $F_{-8}^{thr} = 10$, comparable to the average EGRET sensitivity, and an isotropic γ-ray intensity like that quoted above is measured. *Therefore, the discrepancy cannot be attributed to a lower threshold for resolving point sources.* Most likely the relatively hard spectral slope for the EGRET EGB was due to underestimation of the EGRET effective area above $\approx$ few GeV [27], or due to lack of a model for Compton γ rays, which can make a significant contribution to the EGB at high latitudes.

The origin of the EGB is an interesting open question. Because of the preponderance of blazars in the Third EGRET catalog [17], a large fraction of this emission was expected to be contributed by blazars, though some pre-Fermi studies [12, 280] found that blazars made only a small fraction of this emission. The detection
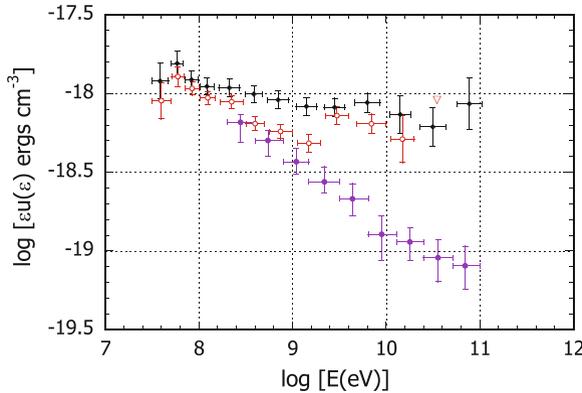
Fig. 20 Fermi measurement of the diffuse extragalactic gamma-ray background (EGB) [7] (*purple filled circles*). Also plotted is the EGRET result [26] (*filled black circles*) and background intensity from GALPROP analysis of EGRET data [28] (*open red circles*)

of ∼1000 high Galactic latitude sources [70], most of which are blazars [92], has allowed a precise characterization of the contribution of blazars to the EGB. Such studies have shown that unresolved blazars contribute ≲30 % of the EGB emission [263, 269].

If not blazars, then other source classes must be invoked to account for the EGB. Possibilities include emission from star-forming galaxies, misaligned blazars (i.e., radio galaxies), γ rays made by particles accelerated by structure-formation shocks in clusters of galaxies, [277, 282, 283], and dark matter; see references in [263]. Star-forming galaxies such as our own Milky Way are known to be γ-ray emitters due to the interaction of cosmic rays with interstellar gas and radiation fields [262]. The superposition of the numerous but individually γ-ray weak star-forming galaxies can contribute a greater fraction of the EGB than the rare, individually bright γ-ray blazars [276, 285]. Fermi detection of several star-forming galaxies, including NGC 253, M82, and the LMC [262], supports this possibility, though the exact percentage remains highly uncertain.

Included in the EGB are pulsar contributions, including millisecond pulsars. Because millisecond pulsars (MSPs) are also found in early-type galaxies, they might contribute a significant fraction of the EGB. A way to determine their integrated diffuse emission would be through the integration of their flux (or luminosity) distribution in γ-rays. The typical spectrum of MSPs is hard (e.g., photon index of ∼1.5) and shows an exponential cut-off around a few GeV [50]. Thus the diffuse emission arising from MSPs should show a bump around a few GeV, similar to a feature found in the Galactic diffuse emission that has been ascribed to annihilating dark matter [279]. The left-hand panel of Fig. 21 illustrates a possible decomposition of the EGB into different source classes, including a contribution from WIMP dark-matter annihilation.
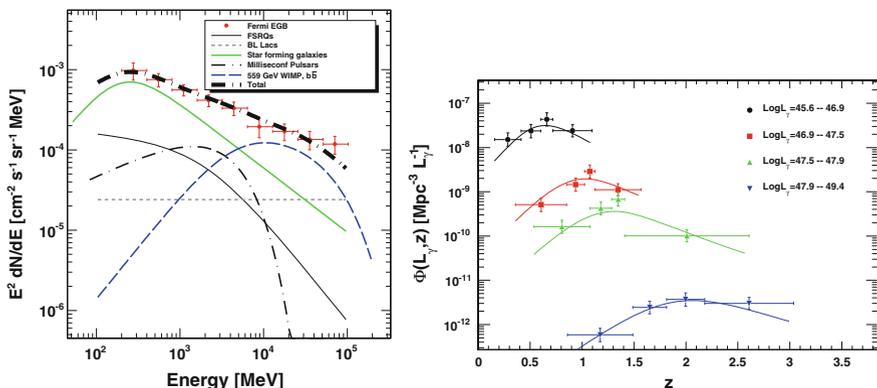
Fig. 21 *Left panel* Fermi-LAT measurement of the extragalactic diffuse background spectrum (data points; see [7]), and model contributions from different source classes. The *curves* represent the integrated contributions of: (1) FSRQs (*thin solid line*), (2) BL Lac objects (*thin short dashed line*), (3) star-forming galaxies, (4) MSPs (*dot-long dashed*), (5) WIMP $b\bar{b}$ annihilation (*long dashed*) and (6) sum of all the previous (*hatched*). *Right panel* Luminosity function of γ-ray selected FSRQs [269]

A preliminary γ-ray luminosity function for FSRQs detected with the Fermi-LAT is shown in the right-hand panel of Fig. 21 [269]. The luminosity function shows the change in the space density of FSRQs as a function of redshift and for different luminosity classes. What is apparent from Fig. 21 is that the space density of powerful FSRQs peaked at redshifts ≈ 2, and then declined to the space density now observed. The peak of maximum growth occurs at different epochs for different luminosity classes, with the more luminous sources reaching their maximum space density earlier in the history of the universe, while the bulk of the population (the lower luminosity blazars) are more abundant at present times. This cosmic downsizing behavior cannot persist to very large ($z \gg 1$) redshifts and early times, as may be reflected by the sudden turnoff of high-z γ-ray blazars. The largest redshift γ-ray blazar in both the 1LAC and 2LAC is at $z = 3.1$, despite many candidate higher-redshift blazars.

7 Microquasars, Radio Galaxies, and the EBL

7.1 γ-Ray Binaries

The literature on the subject of Galactic binaries is vast; a good X-ray review remains Ref. [322]. Here we highlight only a few of those aspects that make for an appreciation of GeV–TeV γ-ray binaries in the Galaxy, of which 6 are well-established, namely

1. LS 5039 ($P = 3.9$ days);
2. LSI $+61°$ 303 ($P = 26$ days);

3. PSR B1259-63 ($P = 3.5$ years);
4. Cygnus X-3 ($P = 4.8$ h);
5. 1FGL J1018.6-5856 ($P = 16.6$ days) [300]; and
6. Cygnus X-1 ($P = 5.6$ days).

These sources are exclusively identified with high-mass X-ray binary (HMXB) systems and are identified, not merely associated, with the named system by period folding or time differencing with the orbital period. With its long orbital period, this is harder for PSR B1259-53, but γ-ray flaring near periastron strengthens the identification. PSR B1259-63 is an example of a Be X-ray binary system, where the compact object, a young pulsar with a spin period of 48 s, is in a highly eccentric orbit around a luminous B star that displays emission lines in its optical spectrum (thus the designation "Be"). Flares take place near periastron at all frequencies, when the pulsar passes through the equatorial stellar wind [318], and now also at GeV energies [294].

High mass and low mass refer to the companion star, not the compact object. HMXBs are X-ray binaries composed of a compact object—white dwarf, neutron star, or black hole—and a young massive O or B main-sequence or Wolf-Rayet star. The low-mass X-ray binaries (LMXBs) typically have older, less massive companions $\lesssim 2M_{\odot}$. If the companion compact object is a pulsar, it can vary in age from young and energetic to old and slowed down (and/or spun up). The fuel illuminating galactic X-ray binaries is accretion onto the compact object, which is transferred, primarily, through stellar wind in HMXBs and Roche-lobe overflow in LMXBs. X-ray variability is modulated by orbital effects, the most important being the change in the binary orbital separation distance. In only the case of PSR B1259-63 are X-ray pulsations at the pulsar period seen, which are also found for intermediate mass X-ray binaries like Her X-1 and other HMXBs at X-ray energies.

Microquasars are X-ray binaries with radio jets. HMXBs number around 7, including SS 433 and V4641 Sgr to the list above. There are 9 LMXB microquasars in the Paredes list from 2005 [331]. Some of the most famous are GX 339-4, Circinus X-1 and Scorpius X-1.

Like the radio structures seen in extragalactic AGNs (e.g., 3C 84), jetted structures have been found in unusual X-ray binary sources, e.g., the Hertz-Grindlay source 1E 1740.7-2942 lying $\sim 0.7°$ away from the Galactic center [326]. The analogy with extragalactic radio sources was strengthened when this source [327] and other LMXB sources such as GRS 1915+107 [328] were shown to exhibit superluminal effects. No LMXBs are yet detected at GeV or TeV energies.

Prior to Fermi, three Galactic binaries were detected at TeV energies, namely LSI +61° 303 with MAGIC and VERITAS, and LS 5039 and PSR B1259-63 with HESS [295]. An isolated flare from Cyg X-1 was observed with MAGIC. No solid EGRET identification with a Galactic binary source had been made, though both LS 5039 [332] and LSI +61° 303, even in the COS-B days [18], were suspected to be γ-ray sources.

A central debate regarding γ-ray binaries is whether they are scaled-down quasars, as suggested by the microquasar label, or colliding stellar wind/pulsar wind systems. Leptonic microquasar models [297, 302, 330] invoke electrons or hadrons in jets

accelerated up to TeV energies with mildly relativistic outflows, accounting for the extended radio emission. Confirming evidence would be VHE emission from X-ray binaries with definite black holes, e.g., Cyg X-1, V 4641, or GRS 1915+105.

The other class of pulsar/star model invokes the rotational energy of the neutron star and the interactions of the pulsar and stellar winds [305, 323]. This model surely applies to PSR B1259-63. Insofar as LS 5039 and LSI +61° 303 have compact objects with $M < 4M_\odot$, the question of whether the compact object is a black hole or neutron star is unresolved. The confirming evidence for a neutron star is, of course, detection of pulsations.

When considering the different sources, it is useful to have a theoretical picture in mind. Most developed and perhaps most likely is a leptonic model, where electrons, which are accelerated in the extreme environment near the compact object, Compton scatter the photons of the hot star. Those γ rays are in turn attenuated by the stellar radiation. To first order, a source at superior conjunction should be GeV luminous (more head-on collisions for Compton scattering) and less TeV luminous (due to $\gamma\gamma$ attenuation). It should anti-correlate with the TeV emission, which has least $\gamma\gamma$ opacity and therefore should be brightest at inferior conjunction.

LS 5039

Though sitting on the edge of the Galactic plane, LS 5039 is well-detected and identified by phase-folding [290]. Its Fermi-LAT light curve is remarkably anti-correlated with the HESS TeV emission. At superior conjunction, when the TeV flux is in a low power-law state, the Fermi GeV spectrum is in a high state. At inferior conjunction, the roles are reversed. Only quasi-simultaneity is possible, because Fermi will necessarily integrate over longer times to achieve the same sensitivity as a TeV telescope. The joint spectra are also remarkable, with the GeV high state fit by an exponentially cut-off power law, and taking place at the same phase as the low-state power-law HESS spectrum. In reverse, the low-state GeV power law takes place when the TeV flux is elevated.

LSI +61° 303

The Fermi LAT light curve of this source is a rather sharp-peaked sinusoid, with peaks of emission around periastron [291]. The GeV radiation is modulated with a maximum close to periastron, and minimum close to apastron. TeV emission peaks at different orbital phase, revealing a flux anti-correlation with the MAGIC TeV emission that might be expected in a leptonic model. The phase-averaged Fermi spectrum is well described by an exponentially cutoff power-law with index ≈ 2.2 and cutoff energy $\approx 6.3\,\mathrm{GeV}$. This spectrum is not found to vary strongly with phase. The non-simultaneous MAGIC and VERITAS data corresponding to different phases look, if they correspond to the time of Fermi observations, like a distinct spectral component, with the hard component extending to TeV energies with a spectral index ≈ 1.9 overtaking the exponentially decaying GeV spectrum.

Cygnus X-3

Cyg X-3 is a HMXB microquasar with an orbital period of 4.8 h—short for an HMXB—at a distance of ≈7 kpc from the Earth. It is found in a complicated region, the Cygnus arm, but its period readily identifies it in γ rays. Its companion is the Wolf-Rayet star V1521 Cyg. It exhibits bright radio flares and goes through active states. It is not known whether it contains a black hole or neutron star binary. Four episodes of flares were observed with AGILE [341] a few days before radio/X-ray flares. The Fermi LAT light curve exhibits periods of scattered activity, and periods of quiescence that appear to be anti-correlated with the X-ray emission [289]. The radio and γ-ray activity are also correlated.

1FGL J1018.6-5856

Period-folding analysis has resulted in the recent discovery of a new high-mass X-ray/γ-ray binary, 1FGL J1018.6-5856 [300, 317]. The 100 MeV–200 GeV Fermi-LAT data is modulated with the 16.58 ± 0.04 day orbital period. The Swift XRT finds an X-ray source coincident with the position of the Fermi source. The digitized sky survey (DSS2) image of this region shows a star with spectrum very similar to that of the gamma-ray binary LS 5039, and the Australia Telescope Compact Array detects a radio source at 9 GHz. Chandra and XMM-Newton observations [333] find a coincident X-ray source from which the column density and reddening can be determined. With 2MASS optical observations, this implies a star of type of O6V at a distance of 6–12 kpc, large by comparison with the distances of LS 5039 and LSI.

PSR B1259-63

PSR B1259-63 may help us understand these systems best by virtue of the compact object being unambiguously a neutron star. The pulsar has a 47.76 ms rotation period and a massive, $\approx 10 M_\odot$ Be star companion, SS 2883. The stars orbit with a 3.5 year orbital period and eccentricity $= 0.87$, guaranteeing episodes of strong interactions near periastron passage. Because of its inclination and the geometry of the equatorial stellar wind, the pre- and post-periastron passages are not symmetrical. EGRET on CGRO failed to detect it during 7 weeks of observation around the 1994 periastron passage [340], but flaring at radio, optical, X-ray, and HESS TeV [295] has been monitored in previous passages.

So it was with great anticipation that observers at all wavelengths waited for the 2010 December passage. PSR B1259-63 did not fail to perform: the light curves measured by an array of instruments at different wavelengths and epochs are shown in Fig. 22, and reveal the complexity of the encounter [294]. It was only weakly detected with the Fermi LAT on the approach to periastron and up to 20 days after (MJD 55517-55563), during which the pre- and near-periastron spectrum in Fig. 22 was measured. The system flared at GeV energies starting at ≈30 day post-periastron,
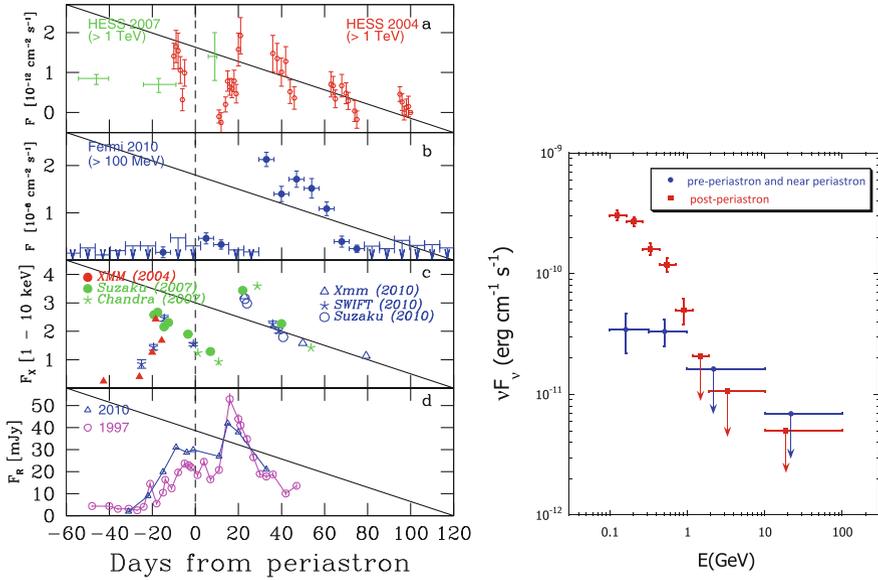
Fig. 22 *Left* Radio, X-ray, GeV, and TeV light curves of PSR B1259-63 at different epochs around the 2011 perastron passage on 2010 December 15 (MJD 55545), as labeled [294]. The 120 days shown represents ≈13% of the orbital period. *Right* Fermi spectra of the system in the pre- and near-periastron period (*blue circles*) and in the post-periastron flare (*red squares*)

with the Fermi LAT measuring during period MJD 55575-55624 shown as the post-periastron spectrum in Fig. 22. The reason that the GeV flare seems to lag behind the peak of the radio and X-ray emissions at ≈20 day post-periastron is not understood.

Though the time of 2010 periastron passage was not ideal for TeV observations due to moonlight conditions, additional post-periastron TeV data for the most recent passage will help illuminate this system. We can reliably predict interesting events in the future.

7.2 Misaligned Blazars and Radio Galaxies

The misaligned AGN (MAGN) population detected with Fermi using the first 15 months of science survey data consists of 11 sources (Fig. 7; [89]), including 7 FRI and 2 FRII radio galaxies, which are thought to be the parent population of BL Lac objects and FSRQs, respectively, in addition to 2 steep spectrum radio quasars that are believed to be slightly misaligned FSRQs. The MAGNs are associated with objects in the Third Cambridge (Revised) and Molonglo radio catalogs. The MAGN sources in the 3CRR catalog have large core dominance parameters compared to the general 3CRR source population, implying that the beamed component makes an

appreciable contribution to the γ-ray flux. This is furthermore supported by the fact that three of these sources—3C 78, 3C 111, and 3C 120—do not appear in 2LAC, evidently due to variability.

Why the ratio of measured γ-ray luminosities of FRI galaxies and BL Lac objects span a much larger range than the comparable ratio for FRII radio galaxies and FSRQs is an open question, if in fact not due to statistics. The simplest possibility is that there are a lot more nearby FRIs, and limitations on Fermi sensitivity will therefore favor detection of these nearby sources. At $z \ll 1$, the luminosity distance $d_L \approx 4200z(1+z)$ Mpc. As noted earlier, the Fermi-LAT reaches a limiting energy-flux sensitivity of $\approx 5 \times 10^{-12}$ erg cm^{-2}s^{-1} for two years of observations that is, unlike integral photon flux, only weakly dependent on source spectral index [92]. Fermi can thus only detect sources with luminosity $L_\gamma \gtrsim 10^{46} z^2 (1+2z)$ erg s^{-1}.

The number of nearby FRI and FRII galaxies depends on the space density of these objects which, in the study by Gendre, Best, and Wall [278], is inferred from 1.4 GHz NVSS-FIRST radio observations (recall that the Fanaroff-Riley dichotomy is based on 178 MHz luminosities $P_{178\,\mathrm{MHz}}$). From their Fig. 12, FRI and FRII radio galaxies have a local ($z < 0.3$) space density of ≈ 35000 Gpc^{-3} and ≈ 2000 Gpc^{-3}, respectively, at $P_{1.4\,\mathrm{GHz}} > 10^{22}$ W Hz^{-1} sr^{-1} or $L_r \gtrsim 10^{40}$ erg/s, implying that the local space density of FRIs exceeds that of FRIIs by ≈ 20.

The volume of the universe within $z < 0.15$ ($d_L < 700$ Mpc) is ≈ 1 Gpc^3, so there are thousands of FRIs and FRIIs within $z \cong 0.1$, yet other than 3C 111, all LAT-detected radio galaxies within this volume are FRI sources (see Fig. 29; [303]). Sensitivity limitations and the abundance of nearby FRIs could explain the MAGN population statistics, but low apparent-luminosity off-axis FRIIs, which are far more radio luminous than BL Lac objects and FRI galaxies, are lacking. With only a few thousand randomly aligned sources within $z = 0.1$, a narrower γ-ray beaming cone in FSRQs, with a more rapid fall-off in off-axis flux, makes detection of these nearby sources far less likely than a broader γ-ray emission cone in BL Lac objects, as expected for the different beaming factors from SSC emission and external Compton processes [301, 311]. This could also reflect differences in jet structure between FSRQs and BL Lac objects [299, 315], or extended jet or lobe emission in FRIs [296] that is missing in FRII galaxies.

Both core and lobes can significantly contribute to the measured γ-ray luminosities. In the case of Centaurus A, the values of L_γ of the core and lobes are comparable, with the lobe emission primarily attributed to Compton-scattered CMBR [46]. The significant or dominant lobe component means that the core luminosity of misaligned AGNs can be less than the measured L_γ unless the lobe emission is resolved.

One-zone synchrotron/SSC models with $\delta_D \gtrsim 10$ give good fits to the long-term average spectra of other HSP BL Lac objects such as Mrk 421 and Mrk 501. The SEDs of radio galaxies, in contrast, are fit with much lower Doppler factors. The SED of the core of Cen A, for instance, can be fit with $\delta_D \approx 1$ and bulk Lorentz factors $\Gamma \approx$ few [292]. Likewise, the SEDs of the FR1 radio galaxies M87 and NGC 1275 [264] and M87 [288] are well fit with $\delta_D \approx 2$ and $\Gamma \sim 4$.

The much larger values of Γ for BL Lac objects than for their putative parent population, the misaligned FR1 radio galaxies, is contrary to simple unification

expectations. Moreover, even though the γ-ray luminosities from FR1 radio galaxies are much smaller than that of BL Lac objects (Fig. 7), they are still larger than expected by debeaming the radiation of BL Lac objects with $\Gamma \gtrsim 20$. Additional soft target photons that can be Compton scattered to high energies result in a reduction of the value of δ_D compared to those implied by the one-zone synchrotron/SSC model. These target photons can be produced in a structured jet, as in the spine and sheath model [299]. Another soft photon source arises if blazar flows decelerate from the inner jet to the pc scale [312], seemingly in accord with the mildly relativistic flows at the sub-pc scale found in radio observations of Mrk 421 and Mrk 501.

7.3 The EBL

In the most general sense, the EBL refers to the cosmic microwave background radiation (CMBR), the radiations from all the past stars and black holes, the glow from annihilating and decaying dark matter, and any residual emissions from exotic particle decays. The EBL at infrared, optical and UV frequencies originates predominantly from stellar emissions, either directly or after being absorbed and reradiated by dust. The present EBL is a consequence of the star formation history of the universe, light absorption and re-emission by dust, and different types of dust extinction in various classes of galaxies throughout cosmic time. The IR and optical/UV EBL attenuates TeV and 10–100 GeV γ-rays, respectively. Knowledge of the absorption due to the EBL is therefore necessary to infer the intrinsic, escaping spectra of extragalactic γ-ray sources.

Directly measuring the EBL at IR and optical wavelengths is difficult because of the interfering foreground zodiacal light and Galactic synchrotron radiation. Many attempts have been taken to predict the EBL intensity. Empirical methods [309, 336, 338] sum optical/IR emissions from galaxy-counts using, e.g., luminosity-dependent galaxy SEDs, and extrapolating to high redshift where data is lacking. Semi-analytic models of the EBL [304, 316, 334] model galaxy formation following mergers of dark matter halos, with effects of supernova feedback, dust attenuation, and metal production considered. Other models [145, 319, 320] are based on integrating stellar light with dust absorption, employing arguments for the star formation rate and the initial mass function of the stars. A lower bound to the EBL intensity can be determined by galaxy counts at different frequencies and redshifts. The upper panel in Fig. 23 shows a comparison of observations with models, and inferred constraints. Curves are model C from [145] (solid black curve); best fit model from [320] (red dotted curve); [309] (dashed green curve); the fiducial model of [316] (dashed blue curve); and the fast evolution and baseline models from [338] (upper and lower dot-dashed violet curves, respectively). Orange curves and inverted triangles are upper limits derived from blazar observations by [324] and [310], respectively.

Gamma-ray astronomy offers a special technique to probe the EBL at IR through UV frequencies, because photons of the EBL attenuate γ rays via pair production through the $\gamma\gamma \rightarrow e^+e^-$ process. The threshold condition implies that 400 GeV
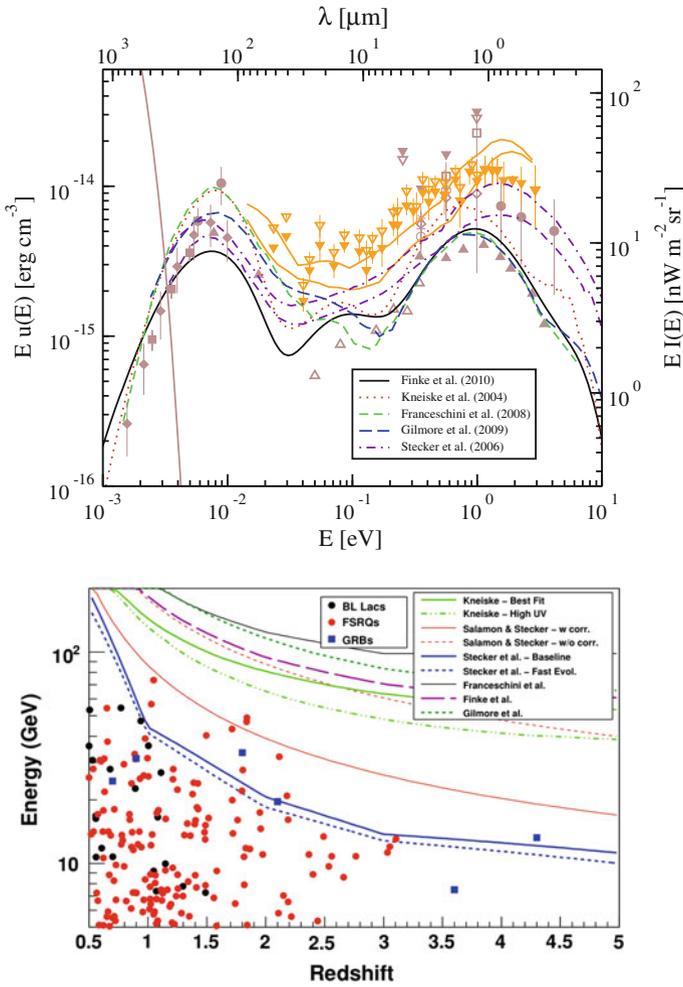
Fig. 23 *Upper* EBL observations, models, and constraints [145]. See text for references and details. *Lower* Highest energy photons observed from blazars and GRBs with the Fermi-LAT as a function of the source redshift. *Curves* giving the $\tau_{\gamma\gamma} = 3$ optical depth for EBL attenuation are shown for different models as labeled [293]

photons can pair produce on soft EBL photons with $1\,\mu$ wavelength, and $4\,\mathrm{TeV}$ photons with $10\,\mu$ EBL photons, etc., so that

$$\lambda(\mu) \Leftrightarrow \frac{E_1}{0.4\,\mathrm{TeV}}.$$

Analysis of the effects of this process have already ruled out high-intensity models of the EBL at IR/optical frequencies [293]. Ruling out other EBL models will require longer integration times to improve implementation of tests like the following:

1. *Flux ratio method.* By assuming that intrinsic blazar spectra are independent of redshift, then the ratio of high to low energy fluxes will decrease with redshift due to enhanced EBL attenuation [298]. It is necessary to use separate blazar populations in this test since Fermi has established significant spectral differences between FSRQs and BL Lacs, but this test is not yet sensitive enough to identify any trend of spectral softenings due to EBL absorption.

2. *Deabsorbed γ-ray spectrum.* By extrapolating the Fermi-LAT spectrum of an individual source to high energies, and assuming that the spectrum does not harden with energy, then deabsorbing the spectrum limits the range of EBL models that are compatible with this underlying assumption [314]. Upper limits are also placed on the allowed EBL intensity by assuming that the intrinsic source spectrum is limited in hardness by some theoretically determined value [324, 337], which depends on details of the particle acceleration mechanism [339]. Specific approaches determine the likelihood of different EBL models to be compatible with the measured γ-ray spectrum by assuming an intrinsic (usually power-law) source spectrum.

3. *Distribution of highest energy photons.* The probability of detecting high-energy photons depends on the γγ optical depth associated with a given EBL model. The detection of such photons, though rare in number, can place severe demands on the apparent power budget of a source. If the absolute power or energy if separately determined, e.g., by using optical breaks to infer the jet opening angle in GRBs, then the required energy release may be extreme and call the EBL model into question (see the lower panel in Fig. 23).

4. *Detection of lobes of radio galaxies.* Photons of the EBL are Compton scattered to GeV energies by relativistic electrons in the radio lobes of nearby extended radio galaxies. Thus measurements of the GeV lobe flux can be used to determine the level of the EBL [313]. The first radio galaxy with radio lobes resolved by Fermi is Cen A [46]. Due to the low significance of the Fermi detection and the lack of data at energies ∼2 GeV, the method cannot yet be applied. Longer exposures should reduce the error bars below 2 GeV and detect photons in the critical regime above ∼2 GeV where the emission from the up-scattered EBL is expected to dominate. Besides Cen A, Cen B and Fornax A—two other radio galaxies with lobes—have recently been detected [92]. These objects, as well as NGC 6251 [325], are potential targets for resolving extended lobe emission to measure or constrain the EBL.

Implicit in the above methods is the assumption that the γ rays are made at the source. This need not be the case if blazars or GRBs are sources of UHECRs. UHECR protons that are able to escape from the blazar and structured regions without deflection can deposit energy in transit through the IGM via photopair production. The cascade radiation can produce persistent TeV emission that would confuse the interpretation of the attenuation due to the EBL. This is less likely for UHECRs accelerated by GRBs, where deflection of the UHECRs would suppress the emission generated by the UHECRs. This is currently an active area of research; see, e.g., [157, 306–308, 329, 335].

8 Fermi Observations of Gamma Ray Bursts

The effective lifetime for GRB studies using EGRET's spark chamber on the *Compton Gamma-ray Observatory* ended ≈4.5 years into mission, after 1996 [20].[20] The depletion of spark-chamber gas was mitigated through the introduction of a narrow-field mode suitable for pointed observations. This made the chance of catching a GRB, proportional to EGRET's FoV, too improbable without rapid automated slewing, which was not possible for CGRO. Consequently EGRET only detected a total of five spark-chamber GRBs, all early in the mission [352]. These are GRB 910503, GRB 910601, the superbowl GRB 930131, the famous long-lived GRB 940217 [358], and GRB 940301. In the wide-field mode, EGRET was sensitive to ≈1/25th of the full sky, which is ≈1/5th as large as the FoV of Fermi [10].

8.1 Fermi LAT Observations of GRBs

Since Fermi operations began, 13 GRBs were reported as significantly detected in the LAT by the Fermi Collaboration from early August 2008 through calendar year 2009.[21] The year 2010 saw a dearth of bright LAT GRBs, but the rate has picked up in 2011 with the detection of a few remarkable events like GRB 110721A and GRB 110731A, the former of which displays classic "fast-rise, exponential decay"-type GBM light curves with prompt LAT emission [369, 371], and the latter of which is the first long-duration GRB jointly detected with Fermi-LAT and Swift from the prompt phase into the afterglow [348, 364]. In the meantime, the development of the LLE technique ([24]; Sect. 1.2) has found LAT emission from several GRBs in the 30–100 MeV range. Given the addition of these GRBs, the rate of GRB detection with the LAT is about 1 per month, and the number of Fermi-LAT GRBs has reached nearly 30. Because analysis is in progress on 2011 GRBs, and there were so few GRBs in 2010, here we review GRB observations only through 2009.

All LAT GRBs during this period are also GBM GRBs and comprise the brighter GBM GRBs, as already expected from a comparison between EGRET and BATSE GRBs in terms of fluence [360]. The 13 Fermi LAT GRBs observed through 2009 include 11 long GRBs and 2 short bursts, namely GRB 081024B and GRB 090510 ($z = 0.903$). The most studied—because they are brightest—GRBs are GRB 090902B ($z = 1.822$) [342], which provides the first strong evidence for a hard spectral component in long GRBs; GRB 080916C ($z = 4.35$) [120], the first bright long GRB; and GRB 090926A ($z = 2.106$), a burst with a narrow spike from the lowest to highest energies in an SED that requires both a Band function and a hard

---

[20] Very bright GRBs like GRB 990123 [349] could still be detected far off the COMPTEL and OSSE axes while making a signal in the Energetic Gamma Ray Experiment Telescope's Total Absorption Shower Counter. EGRET TASC and BATSE data were used to made the discovery of the additional hard component in GRB 941017 [355].

[21] See http://fermi.gsfc.nasa.gov/ssc/observations/types/grbs.
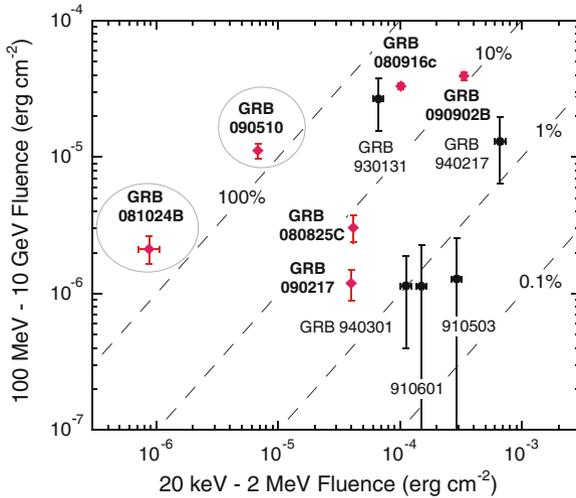
Fig. 24 Fluence-fluence diagram showing 6 Fermi [345] GRBs (*red* data points) and 5 EGRET spark-chamber [360] GRBs (*black* data points). The EGRET fluence is measured from 100 MeV to 5 GeV, whereas the Fermi LAT fluence is measured from 100 MeV to 10 GeV. Short hard GRBs are *circled*

power-law component to fit. GRB 090926A also reveals an extraordinary spectral softening at $\gtrsim 1$ GeV in its time-integrated spectrum when the hard LAT spectral component is bright. These long GRBs had GBM fluences $\mathcal{F}$ in the 20–2000 keV range $\gtrsim 10^{-4}$ erg cm^{-2} (Fig. 24). The bright, short GRB 090510, with $\mathcal{F} \cong 10^{-5}$ erg cm^{-2}, also shows (like GRB 090902B and GRB 090926A) a distinct high-energy power-law spectral component in addition to a Band component [124]. Its short duration, large distance, and the detection of a 31 GeV photon permit strict tests on quantum gravity theories that predict a dependence of the speed of light in vacuo that is linear with energy [343].

Besides these notable GRBs are the less well-known and also less fluent long duration GRBs 090323 ($z = 3.57$), 090626 and 090328 ($z = 0.736$), the widely off-LAT-axis GRB 081215A, the first LAT GRB 080825C [344], the unusual GRB 090217 [346] showing none of the typical properties of LAT GRBs, and the unremarkable LAT GRBs 091003A and GRB 091031. The weakest fluence GRB of the sample is the first short GRB detected at LAT energies, GRB 081024B [345], with $\mathcal{F} \approx 4 \times 10^{-7}$ erg cm^{-2}. The weakness of this GRB could be related to the high $E_{pk} \approx 2$–3 MeV of its Band-function component, but the time-averaged $E_{pk} \cong 4$ MeV for GRB 090510 between 0.5 and 1 s after trigger is even higher [124].

For those GBM GRBs occurring within the LAT FoV, detection of GRBs with the LAT is almost guaranteed when $\mathcal{F} \gtrsim 10^{-4}$ erg cm^{-2}. The detection rate slips to less than 50 % when $\mathcal{F} \approx 3 \times 10^{-5}$ erg cm^{-2}, and becomes highly improbable for $\mathcal{F} \lesssim 10^{-5}$ erg cm^{-2}. This behavior undoubtedly reflects a distribution in the ratios of $\gtrsim 100$ MeV LAT to GBM energy fluence [356].

In the first 16 months of Fermi science operations, $\lesssim 1$ GRB per month was detected with the Fermi LAT, or ≈ 9 GRB/year, with LAT detecting short GRBs at ≈ 10–20% of the rate of long GRBs. GRBs are detected with the GBM at a rate of 250 GRB/year, or ≈ 500 GRB/year (full-sky). When corrected for FoV, EGRET detected ≈ 25 GRB/year (full sky), while the Fermi LAT detects ≈ 50 GRB/year (full sky). Given the much larger effective area of Fermi than EGRET, by a factor ≈ 6 [$\approx (8000$–$9000\,\mathrm{cm}^2)/(1200$–$1500\,\mathrm{cm}^2)$], this small rate increase is something of a surprise, compounded by the ongoing sparse period of Fermi LAT detections of GRBs in 2010. Part of this difference is the stronger detection criteria of Fermi LAT than EGRET. But an improvement in flux sensitivity by a factor ≈ 6, with an accompanying rate increase by only a factor ≈ 2–3 suggests that LAT GRBs are being sampled in a portion of their $\log N - \log \mathcal{F}$ distribution that is flattened by cosmological effects. This is consistent with the known redshifts of LAT GRBs, which range from ≈ 0.7 to $z = 4.35$, with a very rough average redshift of $\langle z \rangle = 2$ for long GRBs and $\langle z \rangle \approx 1$ for short GRBs (based only on GRB 090510). If typical, both classes of GeV-emitting GRBs would be subject to strong cosmological effects on the fluence and flux distributions.

Fluence-Fluence Diagram

Figure 24 shows the fluence-fluence diagram for the 5 EGRET spark-chamber [360] and for 6 Fermi [345] GRBs. Most GBM GRBs have $\mathcal{F} \lesssim 10^{-5}\,\mathrm{erg\,cm}^{-2}$, and are only rarely detected with the LAT. Because of the small number of LAT GRBs, it is not yet clear whether there is a systematic difference between fluence ratios of EGRET and Fermi LAT-detected GRBs. The weakest Fermi LAT GRBs in terms of GBM fluence are both short duration GRBs. This could indicate a preference for short GRBs to have a larger ratio of LAT to GBM fluences than long GRBs, depending on possible triggering biases, e.g., increased LAT background for long GRBs.

The apparent isotropic energies $\mathcal{E}_{iso}$ of GBM and LAT emission for LAT GRBs with known redshifts are in several cases $\gtrsim 10^{54}$ erg. For GRB 080916C, $\mathcal{E}_{iso} \approx 10^{55}$ erg. The LAT GRBs tend to have the largest energies of all measured GRBs, and as a result are good for radio studies [137].

Fermi LAT GRB Phenomenology

Some distinct behaviors have been identified in Fermi GRBs, namely:

- Extended (long-lived) LAT (100 MeV–GeV) emission compared to the GBM (20 keV–2 MeV) emission, known already from EGRET observations, especially GRB 940217 [358].
- Delayed onset of the LAT emission compared to the GBM emission, in both long and short classes of GRBs.

- Power-law temporal decay profiles of the LAT extended emission, decaying typically $\propto t^{-1.5}$ [354].
- Appearance of separate power-law spectral components with photon number index harder than -2.
- Delayed onset of the lowest energy GBM emission at $\approx 10\,\mathrm{keV}$, seen for example in GRB 090902B and GRB 090926A.
- Quasi-thermal Band function components with steep Band β found, e.g., in GRB 090902B at $E \gtrsim 1\,\mathrm{MeV}$ [342].

The onsets of the $>100\,\mathrm{MeV}$ emission appear to be delayed by $\sim 0.1 t_{90}$ compared to the $100\,\mathrm{keV}$–MeV emission (with t_{90} measured, e.g., in the 50–300 keV GBM/BATSE range). This is one of the key and unanticipated results on GRBs from Fermi, and it appears to operate equally for both the long- and short-duration LAT GRBs. There have as yet been no LAT detections of members of the low-luminosity/sub-energetic class of GRBs that includes GRB 980425 and GRB 030329, nor have any X-ray flashes or X-ray luminous GRBs been detected with the LAT. Because GBM's primary triggering modes are similar to BATSE, high E_{pk}, relatively low-z GRBs are more likely to be detected compared to *Swift*.

The luminosity function can be constructed from observations of GRBs. This function gives information about the rate density, the opening angle, and the total nonthermal energy radiated by GRBs. A description of the luminosity function for GRBs, and how the luminosity density can be evaluated for different GRB models is now given.

8.2 GRB Luminosity Function

The luminosity function $Y(L)$ describes the distribution of apparent isotropic luminosities of members of a common source class. For GRBs, one can construct the luminosity function for, at least, the long duration (classical) GRBs, the short hard GRBs, and the X-ray flashes and sub-energetic (low-luminosity) GRBs. As defined by [361], the luminosity function

$$Y_{\mathrm{LZVD}}(L) = L\,\frac{dN}{dL} = \frac{dN}{d\ln L} = Y_0\left[\left(\frac{L}{L_b}\right)^{\alpha} + \left(\frac{L}{L_b}\right)^{\beta}\right]^{-1} H(L; L_l, L_u), \quad (146)$$

with normalization

$$Y_0^{-1} = L_b \int_{L_l/L_b}^{L_u/L_b} dx\,(x^{\alpha} + x^{\beta})^{-1}.$$

As defined by [151], the luminosity function

$$Y_{\mathrm{GPW}}(L) = c_0 \begin{cases} \left(\frac{L}{L_*}\right)^{\alpha}, & L_*/\Delta_1 < L < L_* \\ \left(\frac{L}{L_*}\right)^{\beta}, & L_* < L < \Delta_2 L_* \end{cases}, \qquad (147)$$

with normalization

$$c_0^{-1} = \alpha^{-1}(1 - \Delta_1^{-\alpha}) + \beta^{-1}(\Delta_2^{\beta} - 1).$$

Generally, the luminosity function depends on z, so that $Y(L) = Y(L; z)$.

Luminosity Density from the Luminosity Function

The luminosity density $\ell = d\mathcal{E}/dVdt$ at redshift z, referred to comoving volume, is given by

$$\ell(z) = \dot{n}(z)\langle \Delta t \rangle \int_0^\infty d(\ln L) \, L \, Y(L; z), \qquad (148)$$

where $\dot{n}(z)$ is the comoving rate density and the integration over dL or $d \ln L$ depends on the definition of the luminosity function $Y(L)$. The local luminosity density, which depends on the low-redshift luminosity function and the local rate density, is

$$\ell_0 = \dot{n}_0\langle \Delta t \rangle \int_0^\infty d(\ln L) \, L \, Y(L). \qquad (149)$$

The relation of mean duration $\langle \Delta t \rangle$ to luminosity is normalized so that the fluence $\mathcal{F} = L\langle \Delta t \rangle$.

For the model of Ref. [151], $Y(L)$ is dimensionless (per unit $\ln L$), and the expression for the local luminosity density is analytic, given by

$$\ell_{\mathrm{GPW}} = \dot{n}_0 c_0 \langle \Delta t \rangle L_* \big[(1 - \Delta_1^{-(1+\alpha)})/(1 + \alpha) + (\Delta_2^{1+\beta} - 1)/(1 + \beta) \big], \quad (150)$$

with $c_0^{-1} = \alpha^{-1}(1 - \Delta_1^{-\alpha}) + \beta^{-1}(1 - \Delta_1^{-\beta})$.

For the model of Ref. [361], $Y(L)$ has dimensions of L^{-1}, and the local luminosity density requires a simple numerical integration of the expression

$$\ell_{\mathrm{LZVD}} = \frac{\dot{n}_0 Y_0 \langle \Delta t \rangle}{L_*} \int_0^\infty dL \, L \left[\left(\frac{L}{L_*}\right)^{\alpha} + \left(\frac{L}{L_*}\right)^{\beta} \right]^{-1}, \qquad (151)$$

with

$$Y_0^{-1} = \left[L_B \int_{L_{low}/L_b}^{L_u/L_b} dx \, (x^{\alpha_1} + x^{\alpha_2})^{-1} \right]^{-1}$$

Elementary Cosmology: Following, e.g., [361], we write the differential rate of GRBs with redshift z between z and $z + dz$ and luminosity L in the range L to $L + dL$ by

$$\frac{dN}{dtdzdL} = \frac{\dot{n}(z)}{1+z} \frac{dV(z)}{dz} Y(L). \tag{152}$$

The comoving volume element is

$$\frac{dV(z)}{dz} = \frac{c}{H_0} \frac{4\pi d_L^2}{(1+z)^2 \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}}. \tag{153}$$

To derive this, note that $dV_* = cdt_* dA_*$, so $dV_*/dz = c|dt_*/dz|dA_*$. With $dA_* = (R_*\chi)^2 d\Omega_* = d_L^2(z)d\Omega/(1+z)^4$ (Ref. [4], Eq. (4.43); see below). The cosmology of the universe we inhabit is a flat ΛCDM universe, that is, a dark-matter dominated flat universe with nonzero cosmological constant, well-described by

$$\frac{dt_*}{dz} = \frac{-1}{H_0(1+z)\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}}. \tag{154}$$

This defines, noting that $V = (1+z)^3 V_*$, Eq. (153). Our standard values are $H_0 = 72 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$, $\Omega_m = 0.73 = 1 - \Omega_\Lambda$.

The proper distance, though not directly measurable, is the distance between two objects that would be measured at the same time t. The proper distance at the present epoch is just the comoving coordinate, so

$$d_{prop} = \chi = ct = c \int_0^{\chi/c} dt = c \int_0^z dz' \left| \frac{dt_*}{dz'} \right| (1+z'). \tag{155}$$

The energy flux for a source isotropically radiating luminosity $L = d\mathcal{E}/dt$ at proper distance d_{prop} is related to the energy flux from a source with isotropic luminosity $L_* = d\mathcal{E}_*/dt_*$ at luminosity distance d_L through the relation

$$\Phi = \frac{d\mathcal{E}}{dAdt} = \frac{d\mathcal{E}/dt}{4\pi d_{prop}^2} = \frac{d\mathcal{E}_*/dt_*}{4\pi d_L^2}. \tag{156}$$

The fluence $\mathcal{F} = \int_{t_1}^{t_2} dt \, \Phi(t)$ measured over some time interval t_1–t_2 is therefore related to the apparent isotropic energy release through the expression

$$\mathcal{E}_* = \frac{4\pi d_L^2 \mathcal{F}}{1+z}. \tag{157}$$

For an expanding universe, $d\mathcal{E}_* = \epsilon_* dN = \epsilon(1 + z)dN = (1 + z)d\mathcal{E}$, and $dt_* = dt/(1 + z)$. From the definition of d_L, the energy flux

$$\Phi = \frac{d\mathcal{E}}{dA dt} = \frac{L_*}{4\pi d_L^2} = \frac{dA}{4\pi d_L^2} \left| \frac{d\mathcal{E}}{dE_*} \right| \left| \frac{dt_*}{dt} \right| \frac{d\mathcal{E}}{dA dt} = \frac{(1 + z)^2 dA}{4\pi d_L^2} \Phi . \tag{158}$$

Therefore $dA = R^2\chi^2 d\Omega = 4\pi d_L^2/(1+z)^2$. At emission time t_*, $dA_* = R_*^2\chi^2 d\Omega_*$, and $dA = R^2\chi^2 d\Omega$, so that $dA_*/dA = 1/(1 + z)$. This implies from the definition of the luminosity distance and the angular diameter $R_*\chi$, with $d\mathcal{E}/dt = (1 + z)^{-2}(d\mathcal{E}_*/dt_*)$, and letting $d\Omega = d\Omega_*$,

$$d_L = d_L(z) = (1 + z)d_{prop} = \left(\frac{R}{R_*} \right)^2 (R_*\chi) = (1 + z)^2 d_A$$

$$= \frac{c}{H_0} (1 + z) \int_0^z dz' \frac{1}{\sqrt{\Omega_m(1 + z')^3 + \Omega_\Lambda}} \rightarrow \frac{cz}{H_0} \left[1 + \left(1 - \frac{3\Omega_m}{4} \right)z + \mathcal{O}(z^2) \ldots \right] \tag{159}$$

$\cong 4170z(1 + 0.8z)$ Mpc for a flat universe.

Compare this derivation from Ref. [4]. The directional event rate, or event rate per sr, is

$$\frac{d\dot{N}}{d\Omega} = \frac{1}{4\pi} \int dV_* \, \dot{n}_*(t_*) \left| \frac{dt_*}{dt} \right| = c \int_0^\infty dz \left| \frac{dt_*}{dz} \right| \frac{(R_*\chi)^2 \dot{n}_*(z)}{(1 + z)}, \tag{160}$$

where the burst emissivity $\dot{n}_*(z)$ gives the rate density of events at redshift z. Volume density in comoving and proper coordinates is related by the expression $\dot{n}_*(\epsilon_*; z) = \dot{n}_{co}(\epsilon_*; z)(1 + z)^3$. Assuming separability of source emission properties and the rate density, then $\dot{n}_{co,i}(z) = \dot{n}_{0i} \Sigma_i(z)$, where $\Sigma_i(z)$ is the structure formation history (SFH) of sources of type i, defined so that $\Sigma_i(z \ll 1) = 1$, and $\dot{n}_{0i}$ is the local ($z \ll 1$) rate density of bursting sources of type i. Thus

$$\frac{d\dot{N}}{d\Omega} = \frac{c}{H_0} \int_0^\infty dz \frac{d_L^2(z)\dot{n}_i \Sigma_i(z)}{(1 + z)^3\sqrt{\Omega_m(1 + z)^3 + \Omega_\Lambda}}. \tag{161}$$

This expression can be easily generalized to accommodate spectral behavior.

Star formation rate functions: The redshift-dependent rate density by Porciani & Madau (2001; SFR2) that has roughly constant star formation at $z > 2$ [365] is described by

$$\dot{n}_{PM}(z) = 23\dot{n}_0 \frac{\exp(3.4z)}{\exp(3.4z) + 22}. \tag{162}$$

The analytic function

$$\Sigma_{SFR}(z) = \frac{1 + a_1}{(1 + z)^{-a_2} + a_1(1 + z)^{a_3}}, \tag{163}$$

Fig. 25 Star formation rate factors [158, 357, 374]
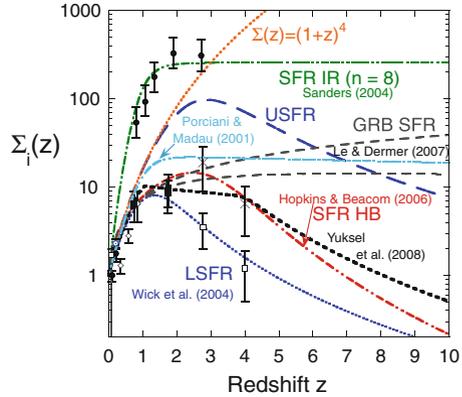
[374], with $a_1 = 0.005$, $a_2 = 3.3$, and $a_3 = 3.0$ for the lower star formation rate (LSFR) and $a_1 = 0.0001$, $a_2 = 4.0$, and $a_3 = 3.0$ for the upper star formation rate (USFR) describes extreme ranges of optical/UV measurements without and with dust extinction corrections, respectively.

The SFR history (SFR3) of Hopkins and Beacom (2006) [357], which is intermediate between the LSFR and USFR, is given by their analytic fitting profile

$$\Sigma_{\mathrm{HB}}(z) = \frac{1 + (a_2 z / a_1)}{1 + (z/a_3)^{a_4}}, \tag{164}$$

where $a_1 = 0.015$, $a_2 = 0.10$, $a_3 = 3.4$, and $a_4 = 5.5$ are best fit parameters (Fig. 25). An update of the SFH fit of [357] is contained in Ref. [375] in the form of a continuous broken power law,

$$\dot{\rho}_*(z) = \dot{\rho}_0 \left[(1+z)^{a\eta} + \left(\frac{1+z}{B} \right)^{b\eta} + \left(\frac{1+z}{C} \right)^{c\eta} \right]^{1/\eta}, \tag{165}$$

where $a = 3.4$ ($b = -0.3$, $c = -3.5$) is the logarithmic slope of the first (middle, last) piece breaking at $z_1 = 1$ and $z_2 = 4$, and the normalization is $\dot{\rho}_0 = 0.02 M_\odot \, \mathrm{Mpc}^{-3} \, \mathrm{year}^{-1}$. Using $\eta = -10$ smooths the transitions, while $\eta \to \infty$ recovers the kinkiness of the original form. Here $B = (1 + z_1)^{1-a/b} \cong 5000$, and $C = (1 + z_1)^{(b-a)/c} (1 + z_2)^{1-b/c} \cong 9$.

Figure 25 shows star formation rate factors used in different models for GRBs. For data, see [367, 374]. An enhanced rate of GRBs compared to the global star formation rate is required to fit Swift statistical data [158], as confirmed in other studies [372, 375].

Luminosity Function in Le & Dermer Model

We approximate the temporally-evolving SED of, for example, a GRB or blazar by the expression

$$\nu F_\nu \equiv f_\epsilon(t) \cong f_{\epsilon_{pk}} S(x) H(\mu; \mu_j, 1) H(t; 0, \Delta t) , \tag{166}$$

and

$$S(x) = x^a H(1-x) + x^b H(x-1), \ x \equiv \frac{\epsilon_*}{\epsilon_{*,pk}} = \frac{\epsilon}{\epsilon_{pk}} = \frac{\epsilon'}{\epsilon'_{pk}}. \tag{167}$$

The bolometric fluence of the model GRB for observers with $\theta \le \theta_j$ is

$$\mathcal{F} = \int_{-\infty}^{\infty} dt \int_0^{\infty} d\epsilon \, \frac{f_\epsilon(t)}{\epsilon} = \lambda_b \, f_{\epsilon_{pk}} \, \Delta t, \tag{168}$$

where λ_b is a bolometric correction to the peak measured νF_ν flux. If the SED is described by Eq. (167), then $\lambda_b = (a^{-1} - b^{-1})$, and is independent of ϵ_{pk}. The beaming-corrected γ-ray energy release $\mathcal{E}_{*\gamma}$ for a two-sided jet is

$$\mathcal{E}_{*\gamma} = 4\pi d_L^2 (1 - \mu_j) \, \frac{\mathcal{F}}{1 + z}, \tag{169}$$

where the luminosity distance is given in Eq. (159). Substituting Eq. (168) for $\mathcal{F}$ into Eq. (169) gives the peak flux

$$f_{\epsilon_{pk}} = \frac{\mathcal{E}_{*\gamma}}{4\pi d_L^2(z)(1 - \mu_j)\Delta t_* \lambda_b}. \tag{170}$$

Finally, by substituting Eq. (170) into Eq. (166), the energy flux becomes

$$f_\epsilon(t) = f_{\epsilon_{pk}} H(\mu; \mu_j, 1) H(t; 0, \Delta t) S(x). \tag{171}$$

The observed directional event rate for bursting sources with νF_ν spectral flux greater than $\hat{f}_{\bar{\epsilon}}$ at observed photon energy $\bar{\epsilon}$ is given by

$$\frac{d\dot{N}(> \hat{f}_{\bar{\epsilon}})}{d\Omega} = \frac{c}{H_0} \int_{\hat{f}_{\bar{\epsilon}}}^{\infty} df'_\epsilon \int_0^{\infty} d\mathcal{E}_{*\gamma} \int_0^{\infty} d\epsilon_{pk*} \int_0^{\infty} d(\Delta t_*) \int_{\mu_{jmin}}^{\mu_{jmax}} d\mu_j \int_{\mu_j}^1 d\mu$$

$$\times \int_0^{\infty} dz \, \frac{d_L^2(z) \dot{n}_{co}(z) Y(\mathcal{E}_{*\gamma}, \mu_j, \epsilon_{pk*}, \Delta t_*)}{(1+z)^3 \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}} \, \delta\left(f'_\epsilon - f_{\epsilon_{pk}} S(x)\right). \tag{172}$$

For discrete values of $\mathcal{E}_{*\gamma}$, ϵ_{pk*}, and Δt_*, the property distribution function

$$Y(\mathcal{E}_{*\gamma}, \mu_j, \epsilon_{pk*}, \Delta t_*) = g(\mu_j)\,\delta(\mathcal{E}_{*\gamma} - \mathcal{E}_{*\gamma 0})\,\delta(\epsilon_{pk*} - \epsilon_{pk*0})\,\delta(\Delta t_* - \Delta t_{*0})\,, \quad (173)$$

where $g(\mu_j)$ is the jet opening angle distribution, and $\hat{f}_\epsilon$ is the instrument's detector sensitivity. The detector threshold for *Swift* and pre-*Swift* GRB detectors is taken to be $\sim 10^{-8}$ and $\sim 10^{-7}$ ergs cm^{-2} s^{-1}, respectively.

In the model of Le & Dermer (2007) [158], the form for the jet opening angle $g(\mu_j)$ is chosen to be

$$g(\mu_j) = g_0\,(1 - \mu_j)^s\,H(\mu_j; \mu_{\text{jmin}}, \mu_{\text{jmax}}), \quad (174)$$

where s is the jet opening angle power-law index; for a two-sided jet, $\mu_{\text{jmin}} \geq 0$. Normalized to unity, Eq. (174) gives

$$g_0 = \frac{1 + s}{(1 - \mu_{\text{jmin}})^{1+s} - (1 - \mu_{\text{jmax}})^{1+s}}. \quad (175)$$

Integrating over μ in Eq. (172) gives the factor $(1 - \mu_j)$ describing the rate reduction due to the finite jet opening angle. Hence, Eq. (172) becomes

$$\frac{d\dot{N}(>\hat{f}_\epsilon)}{d\Omega} = \frac{c}{H_0} \int_{\hat{f}_\epsilon}^\infty df'_\epsilon \int_{\mu_{\text{jmin}}}^{\mu_{\text{jmax}}} d\mu_j\, g(\mu_j)(1 - \mu_j)$$

$$\times \int_0^\infty dz\, \frac{d_L^2(z)\,\dot{n}_{co}(z)\,\delta\left(f'_\epsilon - f_{\epsilon_{pk}} S(x)\right)}{(1+z)^3\,\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}}, \quad (176)$$

where $f_{\epsilon_{pk}}$ is given by Eq. (170). The redshift, size, and opening angle distributions can then be calculated for comparison with data.

Fitting to redshift and opening-angle distributions of pre-Swift and Swift GRBs using this model gives a good fit, assuming a flat νF_ν spectrum with bolometric factor $\lambda_b = 5$, for the following parameters:

$$\Delta t_* = 10\,\text{s},\, \mathcal{E}_{*\gamma} = 4 \times 10^{50}\,\text{erg s}^{-1},\, s = -1.3,\, \mu_{min} = 0.765,\, \mu_{max} = 0.99875, \quad (177)$$

i.e., $\theta_{min} = 0.05\ (2.9°)$, $\theta_{max} = 0.7\ (40°)$. The "*true*" event rate $\dot{n}_{GRB} = (7.5–9.6)$ Gpc^{-3} year^{-1}. This can be compared with beaming corrected rates of $\cong 75 \times 0.5$–40 Gpc^{-3} year^{-1} [151]. The local emissivity in all forms of electromagnetic radiation from long GRBs is automatically implied in this model, namely

$$\ell_{LD} = \dot{n}_{GRB}\mathcal{E}_{*\gamma} \approx (3–4) \times 10^{43}\,\text{erg s}^{-1}\,\text{Mpc}^{-1} \quad (178)$$

Luminosity function: We derive the luminosity function for the Le & Dermer (2007) model, which can be checked in the limit $z \ll 1$ to agree with Eq. (178). The isotropic luminosity $L_{*,iso} = 4\pi d_L^2 \Phi = 4\pi d_L^2 \int_0^\infty f_\epsilon / \epsilon$. From Eqs. (166) and (170),

$$f_{\epsilon_{pk}} = L_{*,iso}(t) S(x)/\lambda_b,$$

where

$$L_{*,iso}(t) = L_{*,iso} \, H(\mu; \mu_j, 1) H(t; 0, \Delta t)$$
$$= \frac{\mathcal{E}_{*\gamma}}{\Delta t_*(1 - \mu_j)} \, H(\mu; \mu_j, 1) H(t; 0, \Delta t)$$

is the apparent isotropic luminosity that would be measured within the beaming cone during time t from 0 to Δt. Thus the relation between the apparent luminosity $L = L_{*,iso}$ and $l_* = \mathcal{E}_{*\gamma}/\Delta t_* = 4 \times 10^{40}$ erg s^{-1} is, obviously, $L = l_*/(1 - \mu_j)$ for a two-sided jet ($0 < \mu_j \leq 1$). Thus $dL/d\mu_j = L/(1 - \mu_j)$, and

$$\frac{dN}{dVdtdL} dL = \frac{dN}{dVdtd\mu_j} d\mu_j (1 - \mu_j), \tag{179}$$

noting the extra factor $(1 - \mu_j)$ which accounts for the smaller opening angle of the more apparently luminous jets.

From the formulation of the model,

$$\frac{dN}{dVdtd\mu_j} = \dot{n}_i g(\mu_j) = \dot{n}_i g_0 (1 - \mu_j)^s H(\mu_j; \mu_j^{min}, \mu_j^{max}). \tag{180}$$

Therefore

$$\frac{dN}{dVdtdL} = \left| \frac{d\mu_j}{dL} \right| \frac{dN}{dVdtd\mu_j}(1 - \mu_j) = \dot{n}_i g_0 \frac{(1 - \mu_j)^{s+2}}{L} H(\mu_j; \mu_j^{min}, \mu_j^{max})$$
$$= \frac{g_0 \dot{n}_i}{l_*} \left(\frac{l_*}{L} \right)^{s+3} H\left(L; \frac{l_*}{1 - \mu_j^{min}}, \frac{1}{1 - \mu_j^{max}} \right). \tag{181}$$

The luminosity function in this model is

$$\Phi_{LD}(L) = \frac{dN}{dL} = \frac{g_0}{l_*} \left(\frac{l_*}{L} \right)^{s+3} H\left(L; 1.7 \times 10^{51} \text{erg s}^{-1}, 3.2 \times 10^{53} \text{erg s}^{-1} \right). \tag{182}$$

Therefore $L^2 \Phi(L) \propto L^{-1-s} \propto L^{0.3}$. From Eq. (175), $g_0 = 0.051$.

Integrating $\Phi(L)$ over L gives the emissivity according to Eq. (149). We find that integration of Eq. (182) recovers the luminosity density given by Eq. (178). The luminosity function in the form $\dot{n}_i L^2 \Phi_{LD}(L)$ is plotted in Fig. 26. The model is

overconstrained, which problem can easily be relaxed by changing the functional dependence of the jet opening angle.

Local Luminosity Density of GRBs

The "νF_ν" intensity $\epsilon I_\epsilon (\text{erg/cm}^2\text{-s-sr}) = m_e c^2 \epsilon^2 \phi(\epsilon)/4\pi$ for unbeamed sources, where $\phi(\epsilon)$ is the differential photon flux, is given by

$$\epsilon I_\epsilon = \frac{m_e c^3}{4\pi} \int_0^\infty dz \left| \frac{dt_*}{dz} \right| \frac{\epsilon_*^2 \dot{n}_{co}(\epsilon_*; z)}{1+z}. \tag{183}$$

(Eq. (4.57) in Ref. [4]). This equation also applies to beamed sources provided that the apparent comoving event rate density $\dot{n}_{co}(z) = \int_0^\infty d\epsilon_* \dot{n}_{co}(\epsilon_*; z)$ is increased by the inverse of the beaming factor, f_b, to get the true source rate density (see Eq. (11)).

The photon luminosity density at redshift z is given by

$$\ell(z) = m_e c^2 \int_0^\infty d\epsilon_* \, \epsilon_* \dot{n}_{co}(\epsilon_*; z) \tag{184}$$

and the local luminosity density $\ell_0 \equiv \ell(z \ll 1)$. Assuming that the only property of long-duration GRBs that is redshift dependent is the rate density, then $\ell(z) = \ell_0 \Sigma(z)$. The integrated intensity $\mathcal{I} = \int_0^\infty d\epsilon I_\epsilon$ is related to the mean photon flux $\Phi = d\mathcal{E}/dAdt$ through the relation $\mathcal{I} = \Phi/4\pi$, so

$$\ell_0 = k \frac{H_0 \Phi}{c}, \tag{185}$$

where

$$k^{-1} \equiv \int_0^\infty dz \frac{\Sigma(z)}{(1+z)^2 \sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}}. \tag{186}$$

Table 10 gives values for k obtained by numerically integrating Eq. (186), using the SFR factors shown in Fig. 25 with the standard ΛCDM cosmology with Hubble constant $H_0 = 72$ km/s-Mpc, $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$. The integration was truncated at a redshift $z_{max} = 10$.

Table 10 Constant k, from Eq. (186), for different star formation rate factors

SFR	k
Constant comoving	1.85
LSFR [374]	0.52
USFR [374]	0.13
HB [357]	0.37
LD [158]	0.33–0.37

The mean flux Φ over the full sky can be estimated from the total fluence per year from GRBs, given the number of GRBs per year. For 1293 GRBs in 4B catalog the total fluence (20–300 keV) is 63.80×10^{-4} erg/cm^2 and the total fluence >20 keV is 153.63×10^{-4} erg/cm^2 (M. Gonzalez, private communication, 2003). For the 4th BATSE Catalog, considering 666 Burst/year full sky with 1293 bursts implies $1293/(666 \text{ GRB/year}) = 1.94$ years $= 61225300$ s, so that the mean 20–300 keV flux 1.04×10^{-10} erg/cm^2-s or 3.3×10^{-3} erg/cm^2-year. Likewise, for the total >20 keV fluence was, we obtain 2.54×10^{-10} erg/cm^2-s or 7.9×10^{-3} erg/cm^2-year Band (2002) [347] obtains 550 GRBs/year for BATSE exposure, which brings the flux down to 6.5×10^{-3} erg/cm^2-year. Subtracting 10–20 % for short GRBs gives $(5–6) \times 10^{-3}$ erg/cm^2-year.

From Eq. (185), writing $\Phi_{-2} = \Phi/(10^{-2} \text{ erg/cm}^2\text{-year})$ we find $\ell_0 = 2.3 \times 10^{43} k\Phi_{-2}$ erg/Mpc^3-year. With $\Phi_{-2} \approx 0.5$ and $k \cong 0.35$, then

$$\ell_0 = 0.5 \times 10^{43} \left(\frac{k}{0.35}\right)\left(\frac{\Phi_{-2}}{0.6}\right) \text{ erg/Mpc}^3\text{-year}, \tag{187}$$

in agreement with the results of Eichler et al. (2010) [353].

This value is at least a factor of 5 smaller than values obtained through statistical treatments of GRB data [151, 158, 361], as can be seen from Table 11. Here we give the local luminosity densities for long GRB models derived either numerically or analytically from the luminosity functions derived in these models.

The reason for the discrepancy between the local emissivity of long duration GRBs derived from BATSE data and inferred from Swift data is suggested by Fig. 26. This figure compares different models for the local GRB luminosity functions, plotted in the form of a differential emissivity function $\dot{n}_0 \Delta t L Y(L)$, where $\dot{n}_0$ is the local rate density, and the characteristic duration of a GRB in the explosion frame is $\Delta t = 10$ s. Because the treatments of Schmidt (2001) [368] and Guetta, Piran, & Waxman (2005) [151] are based on BATSE observations in the 50–300 keV range, a bolometric correction factor $=6.3$ is applied to normalize them to the 1 keV–10 MeV range (D. Wanderman, private communication, 2010). In this case, the implied luminosity density from the treatments of the luminosity function agree within factors of ≈ 3.

Table 11 Local luminosity density of long GRBs for different models

Model	ℓ_0 (10^{44} erg/Mpc^3-year)	Photon energy range
Schmidt (2001) [368]	0.03	50–300 keV
Guetta et al. (2005) [151]	0.05–0.08	50–300 keV
Le & Dermer (2007) [158]	0.30–0.38	All photon energies
Liang et al. (2007) [361]	2.0	1 keV–10 MeV
Wanderman & Piran (2010) [372]	1.3[a]	1 keV–10 MeV

[a]Normalizing to the Swift durations brings this value down by a factor of ≈ 2
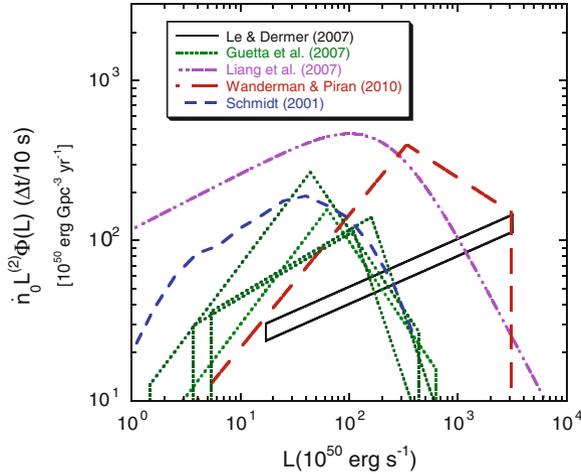
Fig. 26 Long duration GRB luminosity function, plotted in the form of a differential luminosity density $\ell = \dot{n}_0 L\, Y(L)\, \Delta t$, for various models

This figure show that the integrated luminosity density depends sensitively on the maximum luminosity taken in the construction of the luminosity function. When redshift information is available, very luminous GRBs are found with apparent luminosities reaching 10^{53} erg/s. To fit these luminous GRBs, the luminosity function must extend to large luminosities, and these very rare, very luminous GRBs make a significant contribution to the integrated luminosity density. For such luminosity functions, the large fluence GRBs make up a large, even dominant part of the total GRB fluence, but such high fluence events are so rare that the chances to now of such a GRB with large, $\gg 10^{-2}$ erg/cm^2, fluence having been detected is low.

The local electromagnetic emissivity of long GRBs, given in Table 11, is therefore found to be $\lesssim 10^{44}$ erg/year-Mpc^3 [350]. This can be compared with the minimum local emissivity of $\ell_{UHECR} \approx 4 \times 10^{44}$ erg/s-Mpc required to power UHECRs, as seen in Fig. 27, and discussed in more detail in the next section. Related estimates can be made for other source classes [361, 363], including the short hard GRB class, the X-ray flashes, the sub-energetic GRBs, and engine-driven supernovae with relativistic outflows that lack GRB-type emissions. These classes, though they have much larger local rate densities than long-duration GRBs, are not much preferred over long-duration GRBs on the basis of their local luminosity density. The so-called low-luminosity GRBs, for example, have a large luminosity density in ejecta kinetic energy, $\gtrsim 10^{46}$ erg Mpc^{-3} year^{-1} [362, 373], but smaller or comparable nonthermal γ-ray luminosity densities than long-duration GRBs.
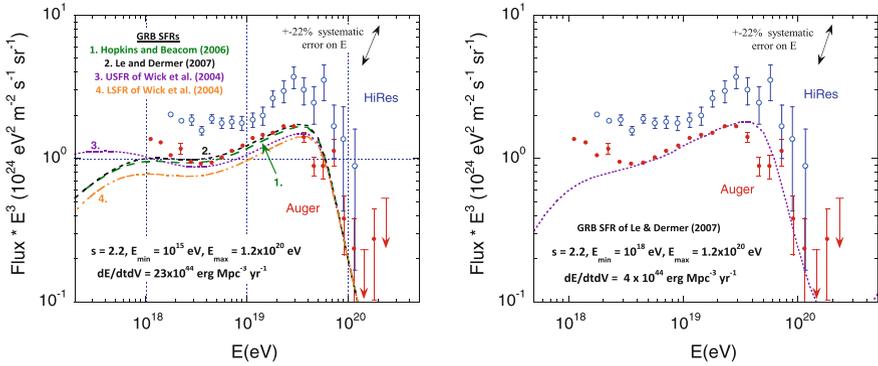
Fig. 27 Local luminosity density requirements on different models for UHECR production from long-duration GRBs with rate density following various star-formation-rate functions shown in Fig. 25

8.3 Closure Relations

The relativistic blast wave theory for the afterglow has become the industry standard model. Familiarity with the derivation of the closure relations is required for basic knowledge of GRB physics. The reader may see[22] for the derivation, [4] for review, and [366] for the essential treatment.

Interpreting the delayed onset of the Fermi-LAT radiation as due to external shock emission, one approach is to suppose that the blast wave decelerates adiabatically in a uniform surrounding medium [359], with closure relation $\nu F_\nu \propto t^{(2-3p)/4} \nu^{(2-p)/2}$, where p is the electron injection index. A value of $p \approx 2.5$ gives a reasonable fit to the data. Another regime to consider is a radiative GRB blast wave [354], where the comparable closure relation is $\nu F_\nu \propto t^{(2-6p)/7} \nu^{(2-p)/2}$, with $p \approx 2$ giving a plausible fit to the data. The adiabatic model requires unusually low densities and magnetic fields for GRB 080916C, and the radiative model supposes pair loading can help achieve strong cooling.

Alternate leptonic models for Fermi LAT GRBs include photospheric models with the photospheric emission passing through shocked plasma in the colliding shells or external shocks of the GRB outflow [370]. A joint Fermi-*Swift* paper examines leptonic afterglow models for GRB 090510 [351].

---

[22] http://heseweb.nrl.navy.mil/gamma/~dermer/notes/index.htm

9 Fermi Acceleration, Ultra-High Energy Cosmic Rays, and Black Holes

Fermi acceleration is intrinsically coupled to the presence of a magnetic field that governs particle motions. The energy is gained stochastically in second-order Fermi scenarios, for example, via resonant pitch-angle scattering, or the particle energy is gained systematically by those few particles that diffuse back and forth across a shock front multiple times before convecting downstream (Sect. 5.5). In either case, the energy gain rate is related to the Larmor rate v/r_L. The magnetic field and the size scale set a fundamental limit on the maximum particle energy through the Hillas criterion, Eq. (78).

9.1 Maximum Particle and Synchrotron Photon Energy

In Fermi acceleration scenarios, the maximum particle energy is obtained by balancing acceleration rate with the energy-loss rate. For relativistic electrons, the maximum comoving electron Lorentz factor is $\gamma'_{max} = \sqrt{6\pi e/\phi \sigma_T} B'$. The maximum synchrotron photon energy for a relativistic jet source at redshift z is, from Eq. (145),

$$\epsilon_{syn,max} \cong \frac{27}{8\alpha_f} \frac{\Gamma}{\phi(1+z)},$$
(188)

or

$$E_{syn,max} \cong 240 \, \frac{(\Gamma/1000)}{\phi(1+z)} \, \text{GeV}.$$
(189)

The delayed appearances of the highest energy photons from Fermi LAT GRBs, which typically happens at times after the GBM t_{90}, calls into question a synchrotron origin for these photons, unless Γ remains large at the end of the prompt phase of GRBs.

The maximum energy for protons and ions is obtained from the Hillas criterion, written as $r'_L = E'/\beta Q B' < R'$. The maximum escaping particle energy $E_{max} = \Gamma E'_{max} < \Gamma E' = Z\beta e B'R$, noting that $R' = R/\Gamma$ from length contraction of the stationary frame size scale as measured in the comoving frame. Relating the magnetic field energy density by a factor $\epsilon_B < 1$ times the proper frame energy density associated with the wind luminosity L, then $B'^2 = 2\epsilon_B L/(\beta c R^2 \Gamma^2)$, and

$$E_{max} < \left(\frac{Ze}{\Gamma}\right) \sqrt{\frac{2\beta \epsilon_B L}{c}},$$
(190)

implying

$$L_\gamma \gtrsim \left(\frac{3 \times 10^{45}}{Z^2 \beta} \right) \Gamma^2 \, E_{20}^2 \, \text{erg/s} \qquad (191)$$

[303, 384, 399], noting that the apparent γ-ray luminosity $L_\gamma < L$.

Equations (72) and (74) imply

$$E_{ionsyn, \, max} \cong m_e c^2 \frac{A}{\phi Z^2} \frac{\Gamma}{1+z} \frac{m_p}{m_e} \left(\frac{27}{8\alpha_f} \right) \cong 0.44 \frac{A(\Gamma/1000)}{\phi Z^2 (1+z)} \, \text{PeV}. \quad (192)$$

Thus proton or ion synchrotron can make a γ-ray component in jet sources provided the ion power and energy is sufficiently great.

9.2 L-Γ Diagram

Using the minimum Lorentz or Doppler factors implied from γ-ray opacity arguments, Eqs. (45) and (46), along with the apparent γ-ray luminosity corresponding to the time that Γ_{min} is measured, we can construct an L-Γ diagram.

Figure 28 shows such a plot for blazars and GRBs [303]. Note that GRB 090510A is a short hard GRB, whereas GRB 080916C and GRB 090926A are long-duration GRBs. Except for PKS 2155-304, which uses HESS data for the giant outbursts in 2006, all the data were measured with the Fermi LAT. As can be seen, the powerful
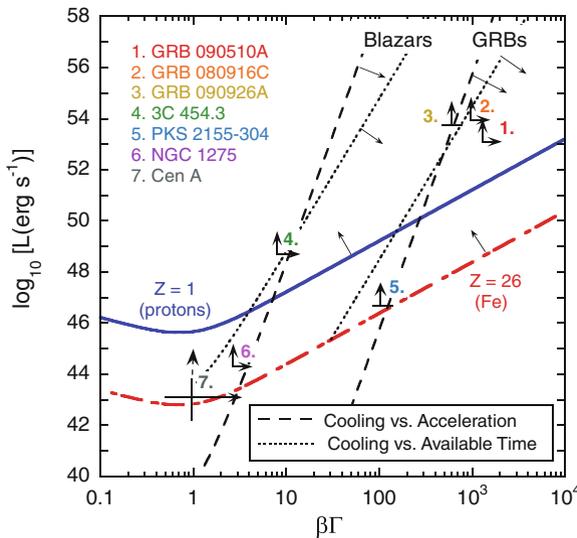
Fig. 28 Data show apparent isotropic L versus Γ_{min} for blazars, radio galaxies, and GRBs. *Solid* and *dot-dashed curves* plot the constraint given by Eq. (191)

Fermi GRBs have more than adequate luminosity to accelerate either protons or ions to ultra-high energies, even with Lorentz factors $\Gamma_{min} \sim 10^3$ as inferred from $\gamma\gamma$ opacity arguments. By comparison, the blazars and radio galaxies have smaller apparent luminosities and also smaller Γ_{min}. Acceleration of high-Z ions like Fe to ultra-high energies is possible for these sources on the basis of Eq. (190), but acceleration of protons appears unlikely, except possibly during flaring episodes.

9.3 *Luminosity Density of Extragalactic γ-Ray Jet Sources*

Maintaining the intensity of UHECRs against photopion losses with the CMB requires a luminosity density in UHECRs of $\sim 10^{44}$ erg/Mpc^3-year [398]. This value can easily be obtained by dividing the energy density of UHECRs at $\approx 10^{20}$ eV, which is $\sim 10^{-21}$ erg/cm^3, by the photopion loss timescale $t_{p\pi} \sim 100$ Mpc/c. Because $t_{p\pi}$ increases rapidly with decreasing energy, the required luminosity density remains at this level even when fitting to the ankle of the cosmic-ray spectrum at $E \approx 4 \times 10^{18}$ eV [381], though precise values of the local luminosity density depends in detail on the assumed cosmic star-formation rate factor applicable to GRBs.

The standard argument for inferring the luminosity density of a source class is to assume that the nonthermal γ-ray luminosity represents, to an uncertain factor of order unity, the luminosity in UHECRs. This seems reasonable because some radiative losses into γ rays are likely to occur during acceleration. If those losses are small, then the UHECR power could exceed the γ-ray luminosity, but if the particles experience severe radiative losses during acceleration, then they could not be the sources of the UHECRs [398]. This estimate is furthermore independent of beaming factor, because for every beamed source we detect, a proportional number of misdirected sources will point away from us. The argument suffers, however, from the likelihood that a large percentage of the γ radiation is produced by ultra-relativistic leptons (especially in the case of blazars where correlated variability with lower energy bands can be monitored). Moreover, it is not certain that MeV γ rays from long duration GRBs, which comprise the bulk of the energy output, is entirely nonthermal, inasmuch as a thermal/photospheric interpretation can potentially resolve the line-of-death problem that plagues nonthermal synchrotron interpretations of GRBs [396]. Thus the nonthermal LAT emission might be a more appropriate luminosity with which to define the nonthermal luminosity density of GRBs [353].

The luminosity density of GRBs has been evaluated in detail in Sect. 8.2. The luminosity density of radio galaxies and blazars, based on the average 100 MeV–100 GeV flux measured in the 1LAC [45], is shown in Fig. 29.
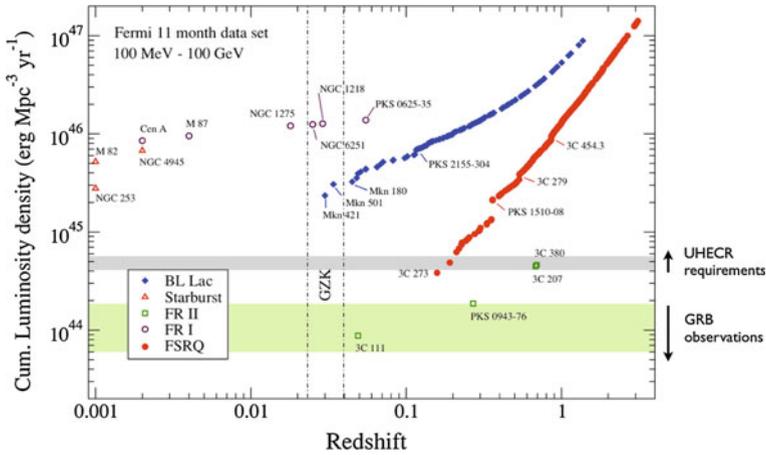
Fig. 29 γ-ray luminosity density inferred from Fermi observations of various classes of γ-ray emitting AGNs [303], compared with the local luminosity density of long-duration GRBs (Table 11), and UHECR source requirements

9.4 Origin of UHECRs

The sources of the UHECRs are unknown. With the anisotropy of the UHECRs observed by Auger becoming less significant than originally reported, leaving the cluster of UHECRs in the vicinity of Cen A as the only outstanding hotspot [6], even Galactic sources have been considered as plausible UHECR candidates [380]. Indeed, a wide variety of possible source classes could contribute to the UHECRs, including magnetars and young pulsars [376, 386], particle acceleration in structure-formation shocks [395], and new physics candidates. Criteria that the sources of UHECRs should meet are [382]

- extragalactic origin;
- mechanism to accelerate to ultra-high energies;
- adequate luminosity density;
- sources within the GZK radius; and
- UHECR survival during acceleration, escape, and transport.

On this basis, we [4] suggest that the very luminous, very energetic extragalactic blazar and GRB jet sources are the most plausible candidates, with the BL Lac objects and FR1 radio galaxies perhaps favored [303]. Electromagnetic signatures of ultra-relativistic hadrons in GRBs and blazars are not clearcut. High-energy neutrino detections from transient or bursting γ-ray sources will be crucial to finally solve this puzzle.

9.5 Black Holes, Jets, and the Extreme Universe

Black Hole Physics

The Kerr metric for an uncharged rotating black hole, written in the Boyer-Lindquist coordinates $\{t, r, \theta, \varphi\}$ useful for asymptotic analysis in the exterior region of the horizon of the black hole, takes the form [397]

$$ds^2 = (\beta^2 - \alpha^2)\, dt^2 + 2\, \beta_\varphi\, d\varphi\, dt + \gamma_{rr}\, dr^2 + \gamma_{\theta\theta}\, d\theta^2 + \gamma_{\varphi\varphi}\, d\varphi^2, \quad (193)$$

where the metric coefficients are given by

$$\beta^2 - \alpha^2 = g_{tt} = -1 + \frac{2Mr}{\rho^2},$$

$$\beta_\varphi \equiv g_{t\varphi} = \frac{-2Mra\sin^2\theta}{\rho^2}, \quad \gamma_{rr} = \frac{\rho^2}{\Delta},$$

$$\gamma_{\theta\theta} = \rho^2, \quad \text{and} \quad \gamma_{\varphi\varphi} = \frac{\Sigma^2 \sin^2\theta}{\rho^2}.$$

Here, α is the lapse function and β is the shift vector,

$$\rho^2 = r^2 + a^2 \cos^2\theta, \quad \Delta = r^2 - 2Mr + a^2, \quad \text{and} \quad \Sigma^2 = (r^2 + a^2)^2 - \Delta\, a^2 \sin^2\theta.$$

Additionally

$$\alpha^2 = \frac{\rho^2 \Delta}{\Sigma^2}, \quad \beta^2 = \frac{\beta_\varphi^2}{\gamma_{\varphi\varphi}}, \quad \text{and} \quad \sqrt{-g} = \alpha\sqrt{\gamma} = \rho^2 \sin\theta.$$

The parameter a is the angular momentum per unit mass of the Kerr black hole, so aM is the angular momentum of the black hole. The horizons are located at $r_\pm = M \pm \sqrt{M^2 - a^2}$, as shown in Fig. 30.

The set of points described by $g_{tt} = 0$ defines the ergosphere

$$r_{erg}(\theta) = M + \sqrt{M^2 - a^2 \cos^2\theta}. \quad (194)$$

Note that $g_{tt} > 0$ in the region

$$r_+ < r < r_{erg}(\theta), \quad (195)$$

and consequently $\tilde{t} \equiv \partial_t$ becomes spacelike in the above region. When $r > r_{erg}$, $\tilde{t}$ is timelike and future directed. See Fig. 30.

In elementary electrodynamics [387], the Poynting vector in flat space far away from a gravitating object is given by
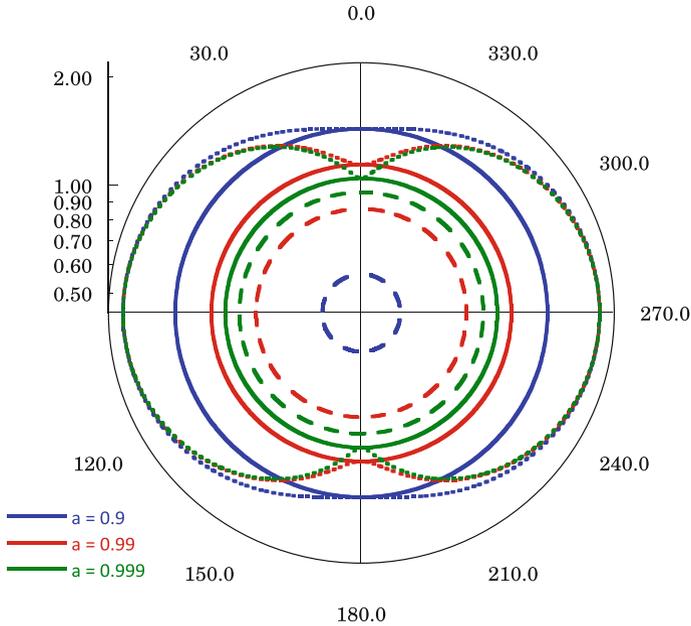
Fig. 30 Sketch of locations of the horizons r_+ and r_- and the ergosphere $r_{erg}(\theta)$ in the Kerr geometry

$$S = \frac{c}{4\pi} (E \times H), \tag{196}$$

and

$$\frac{d}{dt}\mathcal{E}_{tot} = \frac{d}{dt}(\mathcal{E}_{field} + \mathcal{E}_{mech}), \tag{197}$$

where

$$\mathcal{E}_{field} = \frac{1}{8\pi} \int_V d^3x \, (E^2 + B^2), \text{ and} \tag{198}$$

$$\frac{d}{dt}\mathcal{E}_{mech} = \int_V d^3x \, J \cdot E \tag{199}$$

give the field and mechanical energies, respectively. Thus we see that $J \cdot E$ is an emissivity, so conservation of energy requires

$$\frac{\partial u}{\partial t} + \nabla \cdot S = -J \cdot E, \tag{200}$$

where the field energy density is $u = (E \cdot D + B \cdot H)/8\pi$.

Electromagnetic energy extraction through the Blandford-Znajek process involving Penrose processes in curved space can be expressed in the case of a stationary axisymmetric force-free magnetosphere as

$$\frac{d^2\mathcal{E}}{dAdt}|_{\text{BZ}} = S^r \sqrt{\gamma_{rr}} = -H_\varphi \Omega B^r \sqrt{\gamma_{rr}} \tag{201}$$

[388, 391]. The fields must satisfy the Znajek regularity condition [400]

$$H_\varphi\Big|_{r_+} = \frac{\sin^2\theta}{\alpha} B^r (2Mr\,\Omega - a)\Big|_{r_+} \tag{202}$$

at the event horizon. We can suppose that the poloidal function Ω defining the fields is radially independent, so

$$\frac{d\mathcal{E}}{dt}|_{\text{BZ}} = 4\pi Mr_+ \int_0^\pi d\theta \sqrt{\gamma} \, \frac{\sin^2\theta}{\alpha} \Omega(\Omega_H - \Omega)B_+^{r2}, \tag{203}$$

where B_+^r is the radial magnetic field at the event horizon and

$$\Omega_H = \frac{a}{2Mr_+} \tag{204}$$

is the angular velocity of the event horizon [389].

The form of Ω in Eq. (203) is found by solving the constraint equation, originally derived by Blandford and Znajek [378], governing fields and currents in a black-hole magnetosphere under the force-free condition. The constraint equation in the $3 + 1$ formalism takes the form [391]

$$\frac{1}{2\Lambda} \frac{dH_\varphi^2}{d\Omega} = \alpha(\rho_c \Omega \gamma_{\varphi\varphi} - J_\varphi), \tag{205}$$

where the charge density ρ_c and current density J is expressed in terms of Ω and metric coefficients. By satisfying the Znajek regularity condition, Eq. (202), we [391] obtained solutions

$$\Omega_\pm \equiv \frac{a}{2Mr_+ \pm \rho_+^2}, \tag{206}$$

$(\rho_+^2 = r_+^2 + a^2 \cos^2\theta)$ by considering the behavior of the constraint equation in the far-field limit for radially-independent Ω. Thus

$$\Omega_+ \equiv \frac{a}{2Mr_+ + \rho_+^2}, \tag{207}$$

and

$$\Omega_- \equiv \frac{1}{a \sin^2 \theta}. \tag{208}$$

The Ω_+ solution implies a rate of electromagnetic energy extraction of

$$\frac{d\mathcal{E}}{dt}|_{\mathrm{BZ}} = \frac{\pi Q_0^2}{a r_+} \left(\arctan \frac{a}{r_+} - \frac{a}{2M} \right) \ \rightarrow \ \frac{8\pi}{3} B_0^2 M^2 \times \begin{cases} \left(\frac{a}{2M} \right)^2, & a/2M \ll 1 \\ \left(\frac{2}{9} \right) \left(\frac{\pi}{4} - \frac{1}{2} \right), & a \rightarrow M, \end{cases} \tag{209}$$

after relating Q_0 and B_0 [4]. The Ω_+ solution generalizes the Blandford-Znajek [378] split monopole solution for all values of $0 < a < M$. Inspection of this solution shows that Poynting flux outflow is greatest along the equator, so the energy flux has a pancake geometry. The jet geometry, as found ubiquitously in nature, is not described by this solution.

The exact Ω_- solution implies an electromagnetic energy flux peaking towards the poles of the rotating black hole, but describes an inward energy flux, so is unphysical. The Ω_- solution has an additional freedom in the assignment of an arbitrary poloidal field Λ in the prescription for the poloidal magnetic field

$$B_P = \frac{\Lambda}{\sqrt{\gamma}} (-\Omega_{,\theta} \partial_r + \Omega_{,r} \partial_\theta) \tag{210}$$

[392]. Energy extraction is accomplished not only by Poynting flux outflow, but also by the matter currents that sustain the black-hole magnetosphere. Analysis of the timelike geodesics consistent with the Znajek regularity condition [393], and symmetry transformation of the Ω_- solution yields a dual class of exact solutions with positive energy extraction of Poynting and matter outflows. The radial magnetic field for this dual class of solutions takes the form

$$B^r = \frac{2}{a} \Lambda \frac{\cos \theta}{\sqrt{\gamma} \sin^3 \theta}. \tag{211}$$

The extra freedom in the solutions with general θ-dependent Λ provides formal solutions devoid of strict physical content, and with arbitrary a dependence.

Performing dimensional analysis of energy extraction using Eq. (203) (see [4, 391, 394] for detailed treatments) yields an estimate of the BZ power given by

$$\frac{d\mathcal{E}}{dt}|_{\mathrm{BZ}} \approx 4\pi M r_+ \frac{\sqrt{\gamma}}{\alpha} \Omega (\Omega_H - \Omega) B_+^{r2} \approx \pi M r_+ \frac{\sqrt{\gamma}}{\alpha} \Omega_H^2 B_+^{r2}, \tag{212}$$

where B_+^r is the radial component of the magnetic field threading the event horizon. Note that the expression peaks at $\Omega = \Omega_H/2$ [390]. Along the equatorial direction, $\theta \cong \pi/2$, $\sqrt{\gamma}/\alpha \cong \Sigma^2/\Delta$. Supposing the expression is evaluated on a size scale somewhat larger than the event horizon, to cancel the divergence, we have

$$\frac{d\mathcal{E}}{dt}|_{BZ} \approx \pi c \left(\frac{a}{M}\right)^2 r_+^2 B_+^{r2} \approx 10^{47} \left(\frac{a}{M}\right)^2 M_9^2 B_4^2 \, \text{erg/s}, \qquad (213)$$

where M_9 is the mass of the black hole expressed in units of $10^9 M_\odot$, and the magnetic field is given in units of 10^4 G.

As expressed, and often found in the literature, the value of 10^4 G is artificial. The magnetic field threading the ergosphere can be scaled to the Eddington luminosity

$$L_{Edd} = 1.26 \times 10^{47} M_9 \, \text{erg/s} \qquad (214)$$

by the expression

$$\frac{B^2}{8\pi} = \epsilon_B \ell_{Edd} \frac{L_{Edd}}{4\pi r_+^2 c}, \qquad (215)$$

where the parameter ℓ_{Edd} reflects not only different accretion rates but different radiative efficiencies, including the reduction in radiative efficiencies in the low Eddington ratio, advection-dominated regime. Thus

$$\frac{\frac{d\mathcal{E}}{dt}|_{BZ}}{L_{Edd}} \approx \epsilon_B \ell_{Edd} \ll 1. \qquad (216)$$

Jets and the Extreme Universe

The time-averaged luminosities of FSRQs extend to values in excess of $L_\gamma \approx 10^{49} \, \text{erg s}^{-1}$ (Fig. 7). In the extraordinary 2010 November flare of 3C 454.3 [9], L_γ reached apparent isotropic luminosity of $(2.1 \pm 0.2) \times 10^{50} \, \text{erg s}^{-1}$ over a period of a few hours, making it the most luminous blazar yet observed. Black-hole mass estimates for 3C 454.3 are in the range $0.5 \lesssim M_9 \lesssim 4$ [98], where $10^9 M_9 M_\odot$ is the mass of the black hole powering this AGN. For this range of masses, the Eddington luminosity therefore ranges from $\approx 6 \times 10^{46} \, \text{erg s}^{-1}$ to $\approx 5 \times 10^{47} \, \text{erg s}^{-1}$. During this extreme outburst, the apparent luminosity of 3C 454.3 was more than a factor of ≈ 400 greater than its Eddington luminosity. Even its time-averaged luminosity of $L_\gamma \approx 10^{49} \, \text{erg s}^{-1}$ is super-Eddington by a factor of ≈ 20.

Assuming that the Eddington condition *does* limit accretion flow onto the black hole, which is likely to be the case for the long-term average luminosity if not for the flaring luminosity, then the absolute radiant luminosity is limited to a value of $L_{abs} \lesssim 5 \times 10^{47} \, \text{erg s}^{-1}$. This is consistent with the large apparent luminosities if the emission is highly beamed. For a simple top-hat jet beaming factor, a jet opening angle θ_j implies a beaming factor $f_b = \theta_j^2/2$ for a two-sided jet with $\theta_j \ll 1$, from Eq. (11). A mechanism for collimation is, however, required. Should this arise from the Blandford-Znajek process, then we are still restricted to values of the absolute Blandford-Znajek power, Eq. (213). As we have seen, by scaling the energy density of the magnetic field to the energy density of accreted matter near the event horizon shows that the Blandford-Znajek power is likewise Eddington-limited. Making the
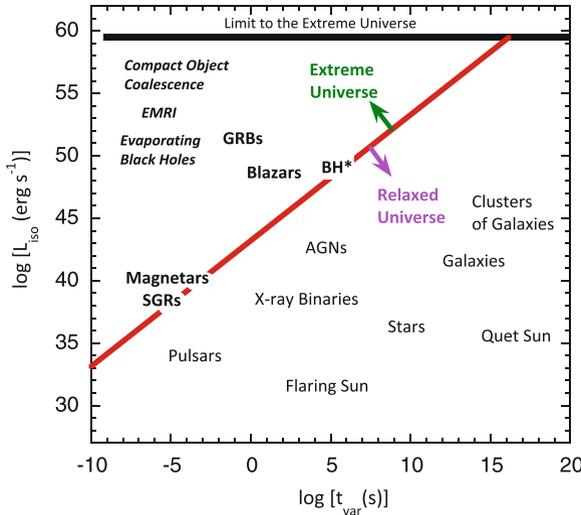
Fig. 31 Luminosity versus characteristic timescales for variability or duration for various types of astronomical sources and cosmic events. Upper limit is from Eq. (17) and the horizontal line separating the extreme and moderate universe is defined by Eq. (14)

hypothesis [383] that the extraction of energy through black-hole rotation collimates the jet outflow with $\cos\theta_j \approx a/M$, then $f_b \cong 1 - (a/M)$ and $\theta_j \cong \sqrt{2(1-a/M)}$, implying $a/M > 1 - (L_{\rm Edd}/L_\gamma)$. If the jet opening angle is a consequence of the bulk Lorentz factor Γ of the outflow, then $\Gamma \gtrsim \sqrt{L_\gamma/2L_{\rm Edd}}$.

For the case of 3C 454.3 in its flaring state, when $L_\gamma/L_{\rm Edd} \gtrsim 10^3$, this hypothesis then implies that $a/M > 0.999$ and $\Gamma \gtrsim 23$, which is consistent with the value $\Gamma_{min} \approx 14$ from $\gamma\gamma$ opacity arguments [9]. In the most conservative case with $M_9 = 4$, $a/M \cong 0.998$ and $\Gamma \gtrsim 15$. This is marginally consistent with the limiting maximum value $a/M \cong 0.996$ suggested by Aschenbach from analyses of microquasars and the Galactic Center black hole [377]. The smaller black-hole mass estimate, $M_9 = 0.5$ [98], implies a value of a/M that violates this limit by a large margin. Much work, both numerical and theoretical, has been devoted to jet formation from the Blandford-Znajek process, and it is unclear if jet collimation can be described by the guess that $\cos\theta_j \approx a/M$, but it is interesting to suggest this possibility, which leads to values of Γ consistent with separate inferences regarding the outflow Lorentz factor.

Figure 31 summarizes some various source classes and their residence in a diagram of L versus time. The extreme universe, as contrasted with the moderate, or relaxed universe, is defined by Eq. (14), when the ratio of the apparent luminosity and variability timescale exceeds the Eddington luminosity divided by the light-crossing time for the Schwarzschild radius, $L_{\rm Edd}/t_{\rm S} = 1.26 \times 10^{43}$ erg/s^2. "BH*" stands for events where normal main sequence stars are tidally captured by a black hole, as in the 28 March 2011 event [379], and "EMRI" stands for extreme mass-ratio

inspiral events [385]. The luminosity limit of the extreme universe is given by Eq. (17), namely $c^5/G = 3.63 \times 10^{59}$ erg/s. Through observations of blazars and GRBs, γ-ray astronomy is pushing toward this limit. A coordinated assault involving gravitational wave observatories, the Fermi γ-ray telescope, neutrino observatories and multiwavelength campaigns might over the next decade reach this limit of the extreme universe.

References

1. Abdo et al. 2010, Science, 329, 817
2. Cheung, C. C., Donato, D., Wallace, E., Corbet, R., Dubus, G., Sokolovsky, K., & Takahashi, H. 2010, The Astronomer's Telegram, 2487, 1
3. CTA Consortium, Actis, M., et al. 2011, *Experimental Astronomy*, 32, 193 (arXiv:1008.3703)
4. Dermer, C. D., & Menon, G. 2009, *High Energy Radiation from Black Holes: Gamma Rays, Cosmic Rays, and Neutrinos* (Princeton University Press, Princeton)
5. Sandoval, A., et al. 2009, arXiv:0912.3329
6. The Pierre AUGER Collaboration, Abreu, P., Aglietta, M., et al. 2010, Astroparticle Physics, 34, 314

    Lecture 1: GeV Instrumentation and the GeV Sky with the Fermi Gamma-ray Space Telescope

7. Abdo, A. A., et al. 2010, *Phys. Rev. Lett.*, 104, 101101
8. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2009, Astroparticle Physics, 32, 193
9. Abdo, A. A., et al. 2011, ApJ, 733, L26
10. Atwood, W. B., et al. 2009, ApJ, 697, 1071
11. Casandjian, J.-M., & Grenier, I. A. 2008, A&A, 489, 849
12. Dermer, C. D. 2007, ApJ, 659, 958
13. Elliot, J. L., & Shapiro, S. L. 1974, ApJ, 192, L3
14. Feroci, M. et al. 2010, A&A, 510, A9
15. Fichtel, C. E., Kniffen, D. A., & Hartman, R. C. 1973, ApJ, 186, L99
16. Fichtel, C. E., Hartman, R. C., Kniffen, D. A., Thompson, D. J., Ogelman, H., Ozel, M. E., Tumer, T., & Bignami, G. F. 1975, ApJ, 198, 163
17. Hartman, R. C., et al. 1999, ApJS, 123, 79
18. Hermsen, W., et al. 1977, Nature, 269, 494
19. Kraushaar, W. L., Clark, G. W., Garmire, G. P., Borken, R., Higbie, P., Leong, V., & Thorsos, T. 1972, ApJ, 177, 341
20. Kurfess, J. D., Bertsch, D. L., Fishman, G. J., & Schönfelder, V. 1997, Proceedings of the Fourth Compton Symposium, 410, 509
21. Mattox, J. R., et al. 1996, ApJ, 461, 396
22. Michelson, P. F., Atwood, W. B., & Ritz, S. 2010, *Reports on Progress in Physics*, 73, 074901
23. Misner, C. W., Thorne, K. S., & Wheeler, J. A. 1973, *Gravitation* (San Francisco: W.H. Freeman and Co.)
24. Pelassa, V., Preece, R., Piron, F., Omodei, N., and Guiriec, S. 2010, arXiv:1002.2617
25. Pittori, C., et al. 2009, A&A, 506, 1563
26. Sreekumar, P., et al. 1998, ApJ, 494, 523
27. Stecker, F. W., Hunter, S. D., & Kniffen, D. A. 2008, Astroparticle Physics, 29, 25
28. Strong, A. W., Moskalenko, I. V., & Reimer, O. 2004, ApJ, 613, 956
29. Swanenburg, B. N., et al. 1981, ApJ, 243, L69
30. Thompson, T. A., & Lacki, B. C. 2011, arXiv:1103.1886
31. Thompson, D. J. 1986, *Nucl. Instrum. Methods Phys. Res.* A, 251, 390
32. Thompson, D. J., et al. 1993, ApJS, 86, 629
33. Thompson, D. J., & Fichtel, C. E. 1982, A&A, 109, 352

Lecture 2: Fermi Gamma-ray Source Catalogs and Fermi Pulsars

34. Abdo, A. A., et al. 2008, Science, 322, 1218
35. Abdo, A. A., et al. 2009, ApJS, 183, 46
36. Abdo, A. A., et al. 2009, ApJ, 700, 597
37. Abdo, A. A., et al. 2009, Science, 325, 840
38. Abdo, A. A., et al. 2009, Science, 325, 845
39. Abdo, A. A., et al. 2009, Science, 325, 848
40. Abdo, A. A., et al. 2009, ApJ, 699, 1171
41. Abdo, A. A., Ackermann, M., Atwood, W. B., et al. 2009, ApJ, 696, 1084
42. Abdo, A. A., et al. 2009, Phys. Rev. Lett., 103, 251101
43. Abdo, A. A., et al. 2010, ApJS, 188, 405
44. Abdo, A. A., et al. 2010, ApJS, 187, 460; (e) 2011, 193, 22A
45. Abdo, A. A., et al. 2010, ApJ, 715, 429
46. Abdo, A. A., et al. 2010, Science, 328, 725
47. Abdo, A. A., et al. 2010, A&A, 524, A75
48. Abdo, A. A., et al. 2010, ApJ, 708, 1254
49. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 723, 649
50. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 713, 154
51. Abdo, A. A., Ajello, M., Antolini, E., et al. 2010, ApJ, 720, 26
52. Abdo, A. A., et al. 2011, *Science*, 331, 739
53. Ackermann, M., et al. 2011, ApJ, 726, 35
54. Atoyan, A. M., & Aharonian, F. A. 1996, MNRAS, 278, 525
55. Atwood, W. B., Ziegler, M., Johnson, R. P., & Baughman, B. M. 2006, ApJ, 652, L49
56. Baring, M. G. 2001, astro-ph/0106161
57. Caraveo, P. A., Bignami, G. F., Mignani, R., & Taff, L. G. 1996, ApJ, 461, L91
58. Demorest, P. B., Pennucci, T., Ransom, S. M., Roberts, M. S. E., & Hessels, J. W. T. 2010, Nature, 467, 1081
59. de Jager, O. C., Harding, A. K., Michelson, P. F., Nel, H. I., Nolan, P. L., Sreekumar, P., & Thompson, D. J. 1996, ApJ, 457, 253
60. Dyks, J., Harding, A. K., & Rudak, B. 2004, ApJ, 606, 1125
61. Fermi-LAT Collaboration, Freiere, P. C. C., et al. 2011, *Science*, 334, 1107
62. Goldreich, P., & Julian, W. H. 1969, ApJ, 157, 869
63. Gould, R. J. 1965, Physical Review Letters, 15, 577
64. Halpern, J. P., & Holt, S. S. 1992, Nature, 357, 222
65. Harding, A. K., Strickman, M. S., Gwinn, C., et al. 2002, ApJ, 576, 376
66. Komissarov, S. S., & Lyutikov, M. 2011, MNRAS, 414, 2017
67. Kuiper, L., Hermsen, W., Verbunt, F., et al. 2000, A&A, 359, 615
68. Kulkarni, S. R., Clifton, T. C., Backer, D. C., Foster, R. S., & Fruchter, A. S. 1988, Nature, 331, 50
69. Lyutikov, M., Balsara, D., & Matthews, C. 2012, MNRAS, 422, 3118
70. Nolan, P. L., Abdo, A. A., Ackermann, M., et al. 2012, ApJS, 199, 31
71. Ramanamurthy, P. V., Bertsch, D. L., Dingus, B. L., et al. 1995, ApJ, 447, L109
72. Ray, P. S., Kerr, M., Parent, D., et al. 2011, ApJS, 194, 17
73. Ray, P. S., & Parkinson, P. M. S. 2011, High-Energy Emission from Pulsars and their Systems, 37 (arXiv:1007.2183)
74. Romani, R. W. 2011, American Institute of Physics Conference Series, 1357, 205
75. Romani, R. W., Kargaltsev, O., & Pavlov, G. G. 2005, ApJ, 627, 383
76. Romani, R. W., & Watters, K. P. 2010, ApJ, 714, 810
77. Shapiro, S. L., & Teukolsky, S. A. 1983, *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects* (New York, Wiley-Interscience)
78. Tavani, M., Striani, E., Bulgarelli, A., et al. 2010, The Astronomer's Telegram, 2855, 1
79. Tavani, M., Bulgarelli, A., Vittorini, V., et al. 2011, Science, 331, 736
80. VERITAS Collaboration, Aliu, E., Arlen, T., et al. 2011, Science, 334, 69

81. Weinberg, S. 1972, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley: New York)
82. Wilson-Hodge, C. A., Cherry, M. L., Case, G. L., et al. 2011, ApJ, 727, L40
83. Yuan, Q., Yin, P.-F., Wu, X.-F., et al. 2011, ApJ, 730, L15

Lecture 3: Fermi AGN Catalogs

84. Abdo, A. A., et al. 2009, ApJ, 707, L142
85. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2009, ApJ, 707, 1310
86. Abdo, A. A., et al. 2010, ApJ, 716, 30
87. Abdo, A. A., et al. 2010, ApJ, 721, 1425
88. Abdo, A. A., et al. 2010d, ApJ, 710, 1271
89. Abdo, A. A., et al. 2010e, ApJ, 720, 912
90. Abdo, A. A., et al. 2011, ApJ, 727, 129
91. Ackermann, M., et al. 2010, ApJ, 721, 1383
92. Ackermann, M., Ajello, M., Allafort, A., et al. 2011, ApJ, 743, 171; arXiv:1108.1420
93. Ackermann, M., Ajello, M., Allafort, A., et al. 2011, ApJ, 741, 30
94. Ackermann, M., Ajello, M., Allafort, A., et al. 2012, ApJ, 747, 104
95. Aharonian, F., et al. 2007, ApJ, 664, L71
96. Aharonian, F., et al. 2009, ApJ, 696, L150
97. Aleksić, J., Antonelli, L. A., Antoranz, P., et al. 2010, ApJ, 723, L207
98. Bonnoli, G., Ghisellini, G., Foschini, L., Tavecchio, F., & Ghirlanda, G. 2010, MNRAS, 410, 368
99. Ghisellini, G., Maraschi, L., & Dondi, L. 1996, A&A Suppl., 120, 503
100. Ghisellini, G., Maraschi, L., & Tavecchio, F. 2009, MNRAS, 396, L105
101. Fanaroff, B. L., & Riley, J. M. 1974, MNRAS, 167, 31P
102. Finke, J. D., Dermer, C. D., Böttcher, M. 2008, ApJ, 686, 181
103. Finke, J. D., & Dermer, C. D. 2010, ApJ, 714, L303
104. Healey, S. E., Romani, R. W., Taylor, G. B., Sadler, E. M., Ricci, R., Murphy, T., Ulvestad, J. S., & Winn, J. N. 2007, ApJS, 171, 61
105. Kirk, J. G., Rieger, F. M., & Mastichiadis, A. 1998, A&A, 333, 452
106. Lenain, J.-P., Ricci, C., Türler, M., Dorner, D., & Walter, R. 2010, A&A, 524, A72
107. Lister, M. L., & Homan, D. C. 2005, AJ, 130, 1389
108. Mattox, J. R., Wagner, S. J., Malkan, M., McGlynn, T. A., Schachter, J. F., Grove, J. E., Johnson, W. N., & Kurfess, J. D. 1997, ApJ, 476, 692
109. Massaro, E., Giommi, P., Leto, C., Marchegiani, P., Maselli, A., Perri, M., Piranomonte, S., & Sclavi, S. 2009, A&A, 495, 691
110. Neronov, A., Semikoz, D., & Vovk, I. 2010, A&A, 519, L6
111. Pogge, R. W. 2000, *New Astronomy Reviews*, 44, 381
112. Poutanen, J., & Stern, B. 2010, ApJ, 717, L118
113. Reimer, A. 2007, ApJ, 665, 1023
114. Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988, ApJ, 325, 74
115. Striani, E., et al. 2010, ApJ, 718, 455
116. Tavecchio, F., Maraschi, L., & Ghisellini, G. 1998, ApJ, 509, 608
117. Urry, C. M., & Padovani, P. 1995, PASP, 107, 803
118. Vestrand, W. T., Stacy, J. G., & Sreekumar, P. 1995, ApJ, 454, L93

Lecture 4: Relativistic Jet Physics

119. Abdo, A. A., et al. 2009, ApJ, 699, 817
120. Abdo, A., et al., 2009, Science, 323, 1688
121. Abdo, A. A., et al. 2010, *Nature*, 463, 919
122. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 722, 520
123. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2011, ApJ, 736, 131
124. Ackermann, M. et al., ApJ, Ackermann, M., et al. 2010, ApJ, 716, 1178
125. Ackermann, M., et al. 2011, ApJ, 729, 114
126. Aharonian, F. A., Coppi, P. S., & Völk, H. J. 1994, ApJ, 423, L5

127. Aharonian, F. A. 2000, *New Astronomy*, 5, 377
128. Aleksić, J., Antonelli, L. A., Antoranz, P., et al. 2011, ApJ, 730, L8
129. Ando, S., & Kusenko, A. 2010, ApJ, 722, L39
130. Asano, K., Guiriec, S., & Mészáros, P. 2009, ApJ, 705, L191
131. Atoyan, A. M., & Dermer, C. D. 2003, ApJ, 586, 79
132. Berezinsky, V., Gazizov, A. Z., & Grigorieva, S. I. 2005, *Physics Letters* B, 612, 147
133. Berezinsky, V. 2008, *Advances in Space Research*, 41, 2071
134. Blumenthal, G. R., & Gould, R. J. 1970, *Reviews of Modern Physics*, 42, 237
135. Bromberg, O., & Levinson, A. 2009, ApJ, 699, 1274
136. Celotti, A., & Fabian, A. C. 1993, MNRAS, 264, 228
137. Cenko, S. B., et al. 2010, ApJ, 711, 641
138. D'Avezac, P., Dubus, G., & Giebels, B. 2007, A&A, 469, 857
139. Dermer, C. D., & Mitman, K. E. 1999, ApJ, 513, L5
140. Dermer, C. D. 2007, ApJ, 664, 384
141. Dermer, C. D., Ramirez-Ruiz, E., & Le, T. 2007, ApJ, 664, L67
142. Dermer, C.D., Cavadini, M., Razzaque, S., Finke, J.D., Chiang, J. & Lott, B., 2011 ApJ, 733, L21
143. Dolag, K., Kachelriess, M., Ostapchenko, S., & Tomàs, R. 2011, ApJ, 727, L4
144. Elyiv, A., Neronov, A., & Semikoz, D. V. 2009, Phys. Rev. D, 80, 023010
145. Finke, J. D., Razzaque, S., & Dermer, C. D. 2010, ApJ, 712, 238
146. Gaidos, J. A., et al. 1996, Nature, 383, 319
147. Georganopoulos, M., Kirk, J. G., & Mastichiadis, A. 2001, ApJ, 561, 111; (e) 2004, ApJ, 604, 479
148. Ghisellini, G., Tavecchio, F., Foschini, L., et al. 2010, MNRAS, 402, 497
149. Granot, J., Cohen-Tanugi, J., & do Couto e Silva, E. 2008, ApJ, 677, 92
150. Gu, M., Cao, X., & Jiang, D. R. 2001, MNRAS, 327, 1111
151. Guetta, D., Piran, T., & Waxman, E. 2005, ApJ, 619, 412
152. Hillas, A. M. 1984, *Ann. Rev. of Astronomy and Astrophysics*, 22, 425
153. Ichiki, K., Inoue, S., & Takahashi, K. 2008, ApJ, 682, 127
154. Jones, F. C. 1968, *Physical Review*, 167, 1159
155. Jorstad, S. G., et al. 2005, AJ, 130, 1418
156. Kneiske, T. M., & Dole, H. 2010, A&A, 515, A19
157. Kotera, K., Allard, D., & Lemoine, M. 2011, A&A, 527, A54
158. Le, T., Dermer, C. D. 2007, ApJ, 661, 394
159. Li, Z. 2010, ApJ, 709, 525
160. Massaro, F., Tramacere, A., Cavaliere, A., Perri, M., & Giommi, P. 2008, A&A, 478, 395
161. Mücke, A., & Protheroe, R. J. 2001, Astroparticle Physics, 15, 121
162. Murase, K., & Nagataki, S. 2006, Phys. Rev. D, 73, 063002
163. Murase, K., Zhang, B., Takahashi, K., & Nagataki, S. 2009, MNRAS, 396, 1825
164. Murase, K., Takahashi, K., Inoue, S., Ichiki, K., & Nagataki, S. 2008, ApJ, 686, L67
165. Neronov, A., & Semikoz, D. V. 2009, Phys. Rev. D, 80, 123012
166. Neronov, A., & Vovk, I. 2010, Science, 328, 73
167. Neronov, A., Semikoz, D. V., Tinyakov, P. G., & Tkachev, I. I. 2011, A&A, 526, A90
168. Plaga, R. 1995, Nature, 374, 430
169. Razzaque, S., Mészáros, P., & Zhang, B. 2004, ApJ, 613, 1072
170. Razzaque, S. 2010, ApJ, 724, L109
171. Razzaque, S., Dermer, C. D., & Finke, J. D. 2010, *Open, Astronomy*, 3, 150 (arXiv:0908.0513)
172. Sikora, M., Begelman, M. C., & Rees, M. J. 1994, ApJ, 421, 153
173. Soderberg, A. M., et al. 2006, Nature, 442, 1014
174. Takahashi, K., Murase, K., Ichiki, K., Inoue, S., & Nagataki, S. 2008, ApJ, 687, L5
175. Tanaka, Y. T., Stawarz, Ł., Thompson, D. J., et al. 2011, ApJ, 733, 19
176. Tavecchio, F., Ghisellini, G., Bonnoli, G., & Ghirlanda, G. 2010, MNRAS, 405, L94
177. Tavecchio, F., Ghisellini, G., Foschini, L., Bonnoli, G., Ghirlanda, G., & Coppi, P. 2010, MNRAS, 406, L70

178. Tavecchio, F., Ghisellini, G., Bonnoli, G., & Foschini, L. 2011, MNRAS, 414, 3566
179. Wang, X.-Y., Li, Z., Dai, Z.-G., & Mészáros, P. 2009, ApJ, 698, L98

Lecture 5: γ Rays from Cosmic Rays in the Galaxy

180. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2009, Phys. Rev. D, 80, 122004
181. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2009, ApJ, 706, L1
182. Abdo, A., et al. 2009 ApJ, 703, 1249
183. Abdo, A. A., et al. 2009, Phys. Rev. Lett., 102, 181101
184. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 710, 133
185. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 712, 459
186. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, Science, 327, 1103
187. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 710, L92
188. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 714, 927
189. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2011, ApJ, 734, 116
190. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2011, ApJ, 734, 28
191. Achterberg, A. 1979, A&A, 76, 276
192. Ackermann, M., et al. 2010, Phys. Rev. D, 82, 092004
193. Ackermann, M., et al. 2011, ApJ, 726, 81
194. Ackermann, M., Ajello, M., Allafort, A., et al. 2011, Science, 334, 1103
195. Ackermann, M., Ajello, M., Allafort, A., et al. 2012, ApJ, 745, 144
196. Ackermann, M., Ajello, M., Allafort, A., et al. 2012, *Physical Review Letters*, 108, 011103
197. Adriani, O., et al. 2009 Nature, 458, 607
198. Aharonian, F. A., & Atoyan, A. M. 1999, A&A, 351, 330
199. Aharonian, F. A., Akhperjanian, A. G., Aye, K.-M., et al. 2004, Nature, 432, 75
200. Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2006, A&A, 449, 223
201. Aharonian, F., et al. 2008, Phys. Rev. Lett., 101, 261104
202. Atoyan, A., & Dermer, C. D. 2012, ApJ, 749, L26
203. Bell, A. R. 1978, MNRAS, 182, 147
204. Bell, A. R. 1978, MNRAS, 182, 443
205. Berezinskii, V. S., Bulanov, S. V., Dogiel, V. A., & Ptuskin, V. S. 1990, *Astrophysics of Cosmic Rays* (Amsterdam: North-Holland)
206. Blandford, R. D., and Eichler, D. 1987, *Physics Reports*, 154, 1
207. Bouchet, L., Strong, A. W., Porter, T. A., Moskalenko, I. V., Jourdain, E., & Roques, J. 2011, ApJ, 739, 29
208. Chang, J., et al. 2008, Nature, 456, 362
209. Cordes, J. M., & Lazio, T. J. W. 2002, arXiv:astro-ph/0207156
210. Crocker, R. M., & Aharonian, F. 2011, Phys. Rev. Lett., 106, 101102
211. Dermer, C. D. 1986, A&A, 157, 223
212. Dobler, G., Finkbeiner, D. P., Cholis, I., Slatyer, T., & Weiner, N. 2010, ApJ, 717, 825
213. Drury, L. O'C. 1983, *Reports on Progress in Physics*, 46, 973
214. Esposito, J. A., Hunter, S. D., Kanbach, G., & Sreekumar, P. 1996, ApJ, 461, 820
215. Finkbeiner, D. P. 2004, ApJ, 614, 186
216. Fishman, G. J., Bhat, P. N., Mallozzi, R., et al. 1994, Science, 264, 1313
217. Frail, D. A. 2011, arXiv:1108.4137
218. Funk, S. 2010, http://meetings.aps.org/Meeting/APR10/Event/116093, APS April Meeting Abstracts, 3002
219. Gaisser, T. K. 1990, *Cosmic Rays and Particle Physics* (Cambridge: Cambridge University Press)
220. Giglietto, N., for the Fermi Large Area Telescope Collaboration 2009, arXiv:0912.3734
221. Ginzburg, V. L., & Syrovatskii, S. I. 1964, *The Origin of Cosmic Rays* (New York: Macmillan)
222. Giordano, F., Naumann-Godo, M., Ballet, J., et al. 2011, 2012, ApJ, 744, L2
223. Green, D. A., & Stephenson, F. R. 2003, Supernovae and Gamma-Ray Bursters, 598, 7
224. Grenier, I. A., Casandjian, J.-M., & Terrier, R. 2005, Science, 307, 1292
225. Hayakawa, S. 1969, *Cosmic ray physics. Nuclear and astrophysical aspects*, Interscience Monographs and Texts in Physics and Astronomy, (New York: Wiley-Interscience)

226. Hewitt, J. W., Yusef-Zadeh, F., & Wardle, M. 2009, ApJ, 706, L270
227. Hunter, S. D., et al. 1997, ApJ, 481, 205
228. Katagiri, H., Tibaldo, L., Ballet, J., et al. 2011, ApJ, 741, 44
229. Lagage, P. O., & Cesarsky, C. J. 1983, A&A, 118, 223
230. Longair, M. S. 1994, *High Energy Astrophysics* (Cambridge: Cambridge University Press), 2nd ed.
231. Lopez, L. A., Ramirez-Ruiz, E., Badenes, C., et al. 2009, ApJ, 706, L106
232. Lopez, L. A., Ramirez-Ruiz, E., Huppenkothen, D., Badenes, C., & Pooley, D. A. 2011, ApJ, 732, 114
233. Lebrun, F., Bennett, K., Bignami, G. F., et al. 1983, ApJ, 274, 231; (e) 1984, ApJ, 282, 359
234. Mertsch, P., & Sarkar, S. 2010, JCAP, 10, 19
235. Mertsch, P., & Sarkar, S. 2011, Phys. Rev. Lett., 107, 091101
236. Morlino, G., & Caprioli, D. 2012, A&A, 538, A81
237. Moskalenko, I. V., Porter, T. A., & Digel, S. W. 2006, ApJ, 652, L65; (e) 2007 ApJ, 664, L143
238. Moskalenko, I. V., Porter, T. A., Digel, S. W., Michelson, P. F., & Ormes, J. F. 2008, ApJ, 681, 1708
239. Moskalenko, I. V., & Porter, T. A. 2009, ApJ, 692, L54
240. Orlando, E., & Strong, A. W. 2007, *Astrophysics & Space Science*, 309, 359 (astro-ph/0607563)
241. Orlando, E., & Strong, A. W. 2008, A&A, 480, 847
242. Porter, T. A., Moskalenko, I. V., Strong, A. W., Orlando, E., & Bouchet, L. 2008, ApJ, 682, 400
243. Porter, T. A., Johnson, R. P., & Graham, P. W. 2011, ARA&A, 49, 155
244. Reynolds, S. P. 1988, IAU Colloq. 101: Supernova Remnants and the Interstellar Medium, 331
245. Schlickeiser, R. 1984, A&A, 136, 227
246. Schlickeiser, R. 1985, A&A, 143, 431
247. Seckel, D., Stanev, T., & Gaisser, T. K. 1991, ApJ, 382, 652
248. Strong, A. W., & Moskalenko, I. V. 1998, ApJ, 509, 212
249. Strong, A. W., Moskalenko, I. V., & Reimer, O. 2000, ApJ, 537, 763
250. Strong, A. W., Moskalenko, I. V., & Reimer, O. 2004, ApJ, 613, 962
251. Strong, A. W., Moskalenko, I. V., & Ptuskin, V. S. 2007, *Ann. Rev. of Nuclear and Particle Science*, 57, 285
252. Strong, A. W., Porter, T. A., Digel, S. W., Jóhannesson, G., Martin, P., Moskalenko, I. V., Murphy, E. J., & Orlando, E. 2010, ApJ, 722, L58
253. Sturner, S. J., & Dermer, C. D. 1995, A&A, 293, L17
254. Su, M., Slatyer, T. R., & Finkbeiner, D. P. 2010, ApJ, 724, 1044
255. Tanaka, T., Uchiyama, Y., Aharonian, F. A., et al. 2008, ApJ, 685, 988
256. Thompson, D. J., Bertsch, D. L., Morris, D. J., & Mukherjee, R. 1997, JGR, 1021, 14735
257. Torres, D. F., Romero, G. E., Dame, T. M., Combi, J. A., & Butt, Y. M. 2003, *Phys. Reports*, 382, 303
258. Uchiyama, Y., 2011, arXiv:1104.1197
259. Yüksel, H., Kistler, M. D., & Stanev, T. 2009, Phys. Rev. Lett., 103, 051101

Lecture 6: γ Rays from Star-Forming Galaxies and Clusters of Galaxies, and the Diffuse γ-Ray Background

260. Abdo, A. A., et al. 2010, A&A, 512, A7
261. Abdo, A. A., et al. 2010, A&A, 523, A46
262. Abdo, A. A., et al. 2010, ApJ, 709, L152
263. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 720, 435
264. Abdo, A. A., et al. 2009, ApJ, 699, 31
265. Abdo, A. A., et al. 2010, A&A, 523, L2
266. Acciari, V. A., et al. 2009, Nature, 462, 770
267. Acero, F., et al. 2009, Science, 326, 1080

268. Ackermann, M., Ajello, M., Allafort, A., et al. 2010, ApJ, 717, L71
269. Ajello, M., Shaw, M. S., Romani, R. W., et al. 2012, ApJ, 751, 108
270. Berrington, R. C., & Dermer, C. D. 2003, ApJ, 594, 709
271. Berezinsky, V. S., Blasi, P., & Ptuskin, V. S. 1997, ApJ, 487, 529
272. Brunetti, G., Blasi, P., Cassano, R., & Gabici, S. 2004, MNRAS, 350, 1174
273. Brunetti, G., & Blasi, P. 2005, MNRAS, 363, 1173
274. Brunetti, G., & Lazarian, A. 2011, MNRAS, 410, 127
275. Brunetti, G. 2011, *Journal of Astrophysics and Astronomy*, 32, 437
276. Fields, B. D., Pavlidou, V., & Prodanović, T. 2010, ApJ, 722, L199
277. Gabici, S., & Blasi, P. 2003, Astroparticle Physics, 19, 679
278. Gendre, M. A., Best, P. N., & Wall, J. V. 2010, MNRAS, 404, 1719
279. Goodenough, L., & Hooper, D. 2009, arXiv:0910.2998
280. Inoue, Y., & Totani, T. 2009, ApJ, 702, 523
281. Kushnir, D., & Waxman, E. 2009, JCAP, 8, 2
282. Loeb, A., & Waxman, E. 2000, Nature, 405, 156
283. Miniati, F., Ryu, D., Kang, H., et al. 2000, ApJ, 542, 608
284. Pavlidou, V., & Fields, B. D. 2001, ApJ, 558, 63
285. Stecker, F. W., & Venters, T. M. 2011, ApJ, 736, 40
286. Tammann, G. A., Loeffler, W., & Schroeder, A. 1994, ApJS, 92, 487
287. Torres, D. F. 2004, ApJ, 617, 966

Lecture 7: Microquasars, Radio Galaxies, and the EBL

288. Abdo, A. A., et al. 2009, ApJ, 707, 55
289. Abdo, A., et al. 2009, Science, 326, 1512
290. Abdo, A. A., et al. 2009, ApJ, 706, L56
291. Abdo, A. A., et al. 2009, ApJ, 701, L123
292. Abdo, A. A., et al. 2010, ApJ, 719, 1433
293. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 723, 1082
294. Abdo, A. A., Ackermann, M., Ajello, M., et al. 2011, ApJ, 736, L11
295. Aharonian, F., et al. 2005, A&A, 442, 1
296. Böttcher, M., Dermer, C. D., & Finke, J. D. 2008, ApJ, 679, L9
297. Bosch-Ramon, V., Romero, G. E., & Paredes, J. M. 2006, A&A, 447, 263
298. Chen, A., Reyes, L. C., & Ritz, S. 2004, ApJ, 608, 686
299. Chiaberge, M., Celotti, A., Capetti, A., & Ghisellini, G. 2000, A&A, 358, 104
300. Corbet, R.H.D., et al. 2011, The Astronomer's Telegram #: 3221
301. Dermer, C. D. 1995, ApJ, 446, L63
302. Dermer, C. D., & Böttcher, M. 2006, ApJ, 643, 1081
303. Dermer, C. D., & Razzaque, S. 2010, ApJ, 724, 1366
304. Domínguez, A., Primack, J. R., Rosario, D. J., et al. 2011, MNRAS, 410, 2556
305. Dubus, G. 2006, A&A, 456, 801
306. Essey, W., & Kusenko, A. 2010, Astroparticle Physics, 33, 81
307. Essey, W., Kalashev, O. E., Kusenko, A., & Beacom, J. F. 2010, Phys. Rev. Lett., 104, 141102
308. Essey, W., Kalashev, O., Kusenko, A., & Beacom, J. F. 2011, ApJ, 731, 51
309. Franceschini, A., Rodighiero, G., & Vaccari, M. 2008, A&A, 487, 837
310. Finke, J. D., & Razzaque, S. 2009, ApJ, 698, 1761
311. Georganopoulos, M., Kirk, J. G., & Mastichiadis, A. 2001, ApJ, 561, 111
312. Georganopoulos, M., & Kazanas, D. 2003, ApJ, 594, L27
313. Georganopoulos, M., Sambruna, R. M., Kazanas, D., Cillis, A. N., Cheung, C. C., Perlman, E. S., Blundell, K. M., & Davis, D. S., 2008, ApJ, 686, L5
314. Georganopoulos, M., Finke, J. D., & Reyes, L. C. 2010, ApJ, 714, L157
315. Ghisellini, G., Tavecchio, F., & Chiaberge, M. 2005, A&A, 432, 401
316. Gilmore, R. C., Madau, P., Primack, J. R., Somerville, R. S., & Haardt, F. 2009, MNRAS, 399, 1694
317. Hill, A. B., Szostek, A., Corbel, S., et al. 2011, MNRAS, 415, 235

318. Johnston, S., Manchester, R. N., Lyne, A. G., D'Amico, N., Bailes, M., Gaensler, B. M., & Nicastro, L. 1996, MNRAS, 279, 1026
319. Kneiske, T. M., Mannheim, K., & Hartmann, D. H. 2002, A&A, 386, 1
320. Kneiske, T. M., Bretz, T., Mannheim, K., & Hartmann, D. H. 2004, A&A, 413, 807
321. Kotera, K.K., Allard, D., & Lemoine, M. 2011, A&A, 527, A54
322. Lewin, W., van Paradijs, J., & van den Heuvel, E. P. J., eds., 1995, *X-ray Binaries* (Cambridge Univ. Press, Cambridge)
323. Maraschi, L., & Treves, A. 1981, MNRAS, 194, 1P
324. Mazin, D., & Raue, M. 2007, A&A, 471, 439
325. Migliori, G., Grandi, P., Torresi, E., et al. 2011, A&A, 533, A72
326. Mirabel, I. F., Rodriguez, L. F., Cordier, B., Paul, J., & Lebrun, F. 1992, Nature, 358, 215
327. Mirabel, I. F., & Rodríguez, L. F. 1994, Nature, 371, 46
328. Mirabel, I. F., & Rodríguez, L. F. 1999, ARA&A, 37, 409
329. Murase, K., Dermer, C. D., Takami, H., & Migliori, G. 2012, ApJ, 749, 63
330. Orellana, M., & Romero, G. E. 2007, *Astrophysics and Space Science*, 309, 333
331. Paredes, J. M. 2005, *Chinese Journal of Astronomy and Astrophysics Supplement*, 5, 121
332. Paredes, J. M., Martí, J., Ribó, M., & Massi, M. 2000, Science, 288, 2340
333. Pavlov, G. G., Misanovic, Z., Kargaltsev, O., & Garmire, G. P. 2011, The Astronomer's Telegram, 3228, 1
334. Primack, J. R., Bullock, J. S., & Somerville, R. S. 2005, *High Energy Gamma-Ray Astronomy*, 745, 23
335. Razzaque, S., Dermer, C. D., & Finke, J. D. 2012, ApJ, 745, 196
336. Salamon, M. H., & Stecker, F. W. 1998, ApJ, 493, 547
337. Schroedter, M. 2005, ApJ, 628, 617
338. Stecker, F. W., Malkan, M. A., & Scully, S. T. 2006, ApJ, 648, 774
339. Stecker, F. W., Baring, M. G., & Summerlin, E. J. 2007, ApJ, 667, L29
340. Tavani, M., et al. 1996, *Astronomy & Astrophysics Supp.*, 120, 221
341. Tavani, M., et al. 2009, Nature, 462, 620

Lecture 8: Fermi Observations of Gamma Ray Bursts

342. Abdo, A. A., et al. 2009, ApJ, 706, L138
343. Abdo, A. A., et al. 2009, Nature, 462, 331
344. Abdo, A. A., et al. 2009, ApJ, 707, 580
345. Abdo, A. A., et al. 2010, ApJ, 712, 558
346. Ackermann, M., et al. 2010, ApJ, 717, L127
347. Band, D. L. 2002, ApJ, 578, 806
348. Bregeon, J., McEnery, J., & Ohno, M. 2011, GRB Coordinates Network, 12218, 1
349. Briggs, M. S., et al. 1999, ApJ, 524, 82
350. Chakraborti, S., Ray, A., Soderberg, A. M., Loeb, A., & Chandra, P. 2011, *Nature Communications*, 2
351. De Pasquale, M., et al. 2010, ApJ, 709, L146
352. Dingus, B. L. 1995, *Astrophys. Space Sci.*, 231, 187
353. Eichler, D., Guetta, D., & Pohl, M. 2010, ApJ, 722, 543
354. Ghisellini, G., Ghirlanda, G., Nava, L., & Celotti, A. 2010, MNRAS, 403, 926
355. González, M. M., Dingus, B. L., Kaneko, Y., Preece, R. D., Dermer, C. D., & Briggs, M. S. 2003, Nature, 424, 749
356. Guetta, D., Pian, E., & Waxman, E. 2011, A&A, 525, A53
357. Hopkins, A. M., & Beacom, J. F. 2006, ApJ, 651, 142
358. Hurley, K., et al. 1994, Nature, 372, 652
359. Kumar, P., & Barniol Duran, R. 2009, MNRAS, 400, L75
360. Le, T., & Dermer, C. D. 2009, ApJ, 700, 1026
361. Liang, E., Zhang, B., Virgili, F., & Dai, Z. G. 2007, ApJ, 662, 1111
362. Murase, K., Ioka, K., Nagataki, S., & Nakamura, T. 2006, ApJ, 651, L5
363. Murase, K., & Takami, H. 2009, ApJ, 690, L14

364. Oates, S. R., Beardmore, A. P., Holland, S. T., et al. 2011, GRB Coordinates Network, 12215, 1
365. Porciani, C., & Madau, P. 2001, ApJ, 548, 522
366. Sari, R., Piran, T., & Narayan, R. 1998, ApJ, 497, L17
367. Sanders, D. B. 2004, *Advances in Space Research*, 34, 535
368. Schmidt, M. 2001, ApJ, 559, L79
369. Tierney, D., & von Kienlin, A. 2011, GRB Coordinates Network, 12187, 1
370. Toma, K., Wu, X.-F., & Mészáros, P. 2011, MNRAS, 415, 1663
371. Vasileiou, V., Piron, F., Tierney, D., et al. 2011, GRB Coordinates Network, 12188, 1
372. Wanderman, D., & Piran, T. 2010, MNRAS, 406, 1944
373. Wang, X.-Y., Razzaque, S., & Mészáros, P. 2008, ApJ, 677, 432
374. Wick, S. D., Dermer, C. D., & Atoyan, A. 2004, Astroparticle Physics, 21, 125
375. Yüksel, H., Kistler, M. D., Beacom, J. F., & Hopkins, A. M. 2008, ApJ, 683, L5

Lecture 9: Fermi Acceleration and Ultra-High Energy Cosmic Rays

376. Arons, J. 2003, ApJ, 589, 871
377. Aschenbach, B. 2004, A&A, 425, 1075
378. Blandford, R. D., & Znajek, R. L. 1977, MNRAS, 179, 433
379. Bloom, J. S., Giannios, D., Metzger, B. D., et al. 2011, Science, 333, 203
380. Calvez, A., Kusenko, A., & Nagataki, S. 2010, Phys. Rev. Lett., 105, 091101
381. Dermer, C. D. 2007, arXiv:0711.2804
382. Dermer, C. D. 2011, American, Institute of Physics Conference Series, 1358, 355, arXiv:1103.4834
383. Dermer, C., & Lott, B. 2012, Journal of Physics Conference Series, 355, 012010, arXiv:1110.3739
384. Farrar, G. R., & Gruzinov, A. 2009, ApJ, 693, 329
385. Gair, J. R. 2009, *Classical and Quantum Gravity*, 26, 094034
386. Ghisellini, G., Ghirlanda, G., Tavecchio, F., Fraternali, F., & Pareschi, G. 2008, MNRAS, 390, L88
387. Jackson, J. D. 1962, *Classical Electrodynamics* (New York: Wiley)
388. Komissarov, S. S. 2004, MNRAS, 350, 427
389. Levinson, A. 2006, Trends in Black Hole Research, 119 (arXiv:astro-ph/0502346)
390. Levinson, A. 2006, *International Journal of Modern Physics A*, 21, 6015
391. Menon, G., & Dermer, C. D. 2005, ApJ, 635, 1197
392. Menon, G., & Dermer, C. D. 2007, *General Relativity and Gravitation*, 39, 785
393. Menon, G., & Dermer, C. D. 2009, *Physical Review D*, 79, 123005
394. Menon, G., & Dermer, C. D. 2011, MNRAS, 417, 1098
395. Murase, K., Inoue, S., & Nagataki, S. 2008, ApJ, 689, L105
396. Ryde, F., & Pe'er, A. 2009, ApJ, 702, 1211
397. Wald, R. M. *General Relativity*, 1984 (Chicago, University of Chicago Press)
398. Waxman, E., & Bahcall, J. 1999, Phys. Rev. D, 59, 023002
399. Waxman, E. 2004, *New Journal of Physics*, 6, 140
400. Znajek, R. L. 1977, MNRAS, 179, 457

Index