International Journal of Modern Physics A
Vol. 19, No. 15 (2004) 2385–2472

World Scientific
www.worldscientific.com

GAMMA-RAY BURSTS: PROGRESS, PROBLEMS & PROSPECTS

BING ZHANG*,† and PETER MÉSZÁROS‡
Department of Astronomy & Astrophysics, The Pennsylvania State University, 525 Davey Lab, University Park, PA 16802, USA
**bzhang@astro.psu.edu*

Received 16 November 2003

The cosmological gamma-ray burst (GRB) phenomenon is reviewed. The broad observational facts and empirical phenomenological relations of the GRB prompt emission and afterglow are outlined. A well-tested, successful fireball shock model is introduced in a pedagogical manner. Several important uncertainties in the current understanding of the phenomenon are reviewed, and prospects of how future experiments and extensive observational and theoretical efforts may address these problems are discussed.

Keywords: Gamma-ray bursts; high energy astrophysics; cosmology.

1. Introduction

The study of astrophysical objects is overwhelmingly done by astronomers using the temporal and spectral information contained in the electromagnetic signals that these objects emit. Gamma-ray bursts (GRB's) are, by definition, electromagnetic signals in the gamma-ray band (in the spectral domain) with short durations (in the temporal domain). They are, however, unusual in having most of their electromagnetic output in gamma-rays, typically at sub-MeV energies, and having most of it concentrated into a brief episode, typically lasting tens of seconds.

The road leading to understanding the nature of these objects has been bumpy, mainly due to the limited information contained in these abrupt gamma-ray episodes. For comparison, almost all the other astrophysical objects are observed in a broader spectral band during a much longer observation time. Historically, the lack (until recently) of observational breakthroughs on GRB's had left ample freedom for modelers to play around, so that the cumulative list of models championed to interpret GRB's has been more numerous than for any other astrophysical phenomenon. As in other fields in science, experimental (or observational) progress

†Also Canadian Institute of Theoretical Astrophysics.
‡Also Department of Physics, Pennsylvania State University; currently on leave at The Institute for Advanced Study, Princeton NJ 08540.

eventually constrains and eliminates most predictive models. At a certain stage of the development of a field, it is useful to ask oneself questions such as what is it that we know, what is it that we still do not know, and how could we find the answers for the unknowns. This is the main emphasis of this review. There exist already a number of comprehensive reviews on GRB observations,[1–3] on theories of GRB[4,5] and on some general or selected topics.[6–15] Our purpose here is to provide a review which differs from others in several aspects. For example, we summarize the current observational progress in an organized, itemized form, making a distinction between solid facts and empirical laws. The introduction of the standard theoretical model is pedagogical, the content of which overlaps other reviews. However, we also dedicate a lot of space to discuss the uncertainties in the current models and highlight some active debates that are going on in the community. We also attempt to reflect in an unbiased manner the work from various groups, rather than focusing on our own work, although some "selection effects" must still exist due to our incomplete survey of the literature.

The structure of the review is as follows. We start with a brief summary of the GRB research history as well as the unique role of GRB study in general astrophysics. We then compile the observational facts and empirical laws (Sec. 2), and introduce the standard theoretical framework, i.e. the fireball shock model (Sec. 3). In Sec. 4 we outline some outstanding issues in the GRB field, which comprises most of the hot topics (and sometimes hot debates) in the current GRB scene. These issues include whether the fireball is dominated by the conventional (hadronic) kinetic energy content or by the less-studied magnetic energy; whether the GRB prompt emission occurs before the fireball deceleration (internal) or at the deceleration radius (external); whether the GRB jets are uniform or with a quasi-universal angular structure; whether the GRB progenitors give rise to a one-step or to a (long delay) two-step collapse, etc. After outlining these issues, we attempt to foresee in Sec. 5 developments in the upcoming new era for GRB study, led by future missions such as Swift and GLAST, complemented by other space- and ground-based detectors for both electromagnetic and non-electromagnetic signals. We review some of the theoretical predictions related to these observational capabilities and discuss how future observational campaigns could constrain and possibly settle some of the outstanding issues listed in Sec. 4.

1.1. *Major landmarks in the study of GRB*

As a framework for the following discussions, we list below some of the most important events in the development of the understanding of GRB's. There may be many other important events which are not included in this list, and we apologize for any omission. Most of the papers selected below are marked by their high citation counts in the ADS archives, an indication that the works are significant and widely accepted. For the theoretical papers, we only include those that survived the later observational data and correctly contributed to the current knowledge of the GRB

phenomenon, or those which, although not fully tested yet, represent "popular" views and draw broad attention from the GRB community. The papers extend up to October 2003, when this review is completed. They include both observational and theoretical developments, which have mutually influenced, and have benefited from, each other.

A list of key advances on the observational side is:[a]

- GRB's were discovered serendipitously in the late 1960s, and the data were reported several years later.[16,17]
- The *Compton Gamma-ray Observatory* (*CGRO*) was launched in 1991, with the *Burst and Transient Experiment* (*BATSE*) on board; *BATSE* provided evidence for an isotropic spatial distribution of GRB's, giving significant support to a cosmological origin interpretation.[18]
- Two categories of GRB's, "long" ($t_\gamma \gtrsim 2$ s) and "short," ($t_\gamma \lesssim 2$ s) were identified.[19]
- Systematic analyses of GRB spectral data indicate that a so-called "GRB function" (or "Band function") fits reasonably well most of the GRB spectra (for both classes of bursts).[20]
- Another detector on board of *CGRO*, the Energetic Gamma Ray Experiment Telescope (*EGRET*), detected some bursts in the hard gamma-ray band; one burst, GRB 940217, was detected to have long-lived GeV emission extending for 1.5 hrs.[21]
- A signature consistent with cosmological time dilation was detected.[22]
- In 1997, the Italian–Dutch satellite *Beppo-SAX* pinpointed the first GRB low energy afterglow (for GRB 970228) in the X-ray band,[23] facilitating optical[24] and radio[25] detections.
- The first measurements of the GRB redshifts were obtained (for GRB 970508[26] and GRB 971214)[27] giving a solid proof that GRB's are at cosmological distances.
- A likely GRB-supernova association (GRB 980425 vs. SN 1998bw) was discovered.[28,29]
- A bright and prompt optical flash and a radio flare were discovered to accompany the energetic burst GRB 990123.[30,31]
- An achromatic steepening break in the afterglow lightcurves was found in several bursts, hinting that at least some GRB fireballs are likely to be collimated.[32,33]
- X-ray spectral features with moderate significance were discovered in several GRB X-ray afterglows.[34,35]
- Assuming the lightcurve steepening is due to the jet break, a collection of the burst jet data revealed that GRB's have a quasi-standard energy reservoir.[36]
- A class of so-called "X-ray flashes" (XRF's) were identified,[37,38] which appear to be closely related to GRB.

[a]Notice that all the afterglow-related breakthroughs are for the so-called "long" GRB's (see Sec. 2 for definition). No afterglow from a "short" GRB has so far been firmly identified.

- Prompt optical afterglows from two more GRB's, GRB 021004 and GRB 021211, were detected by ground-based robotic telescopes.[39–41]
- The prompt gamma-ray emission was reported to be strongly polarized in GRB 021206[42] (cf. Refs. 43 and 44).
- The first clear GRB-SN association was identified: an unambiguous supernova light-curve bump and spectrum in the GRB 030329's optical afterglow lightcurve was discovered.[45,46]
- A distinct high energy spectral component which appears to evolve separately from the usual sub-MeV component was identified in GRB 941017.[47]
- Broadband observations of GRB 030329 suggest the there might exist a multi-component jet structure in some bursts, the total energy output being standard.[48,49]

On the theoretical side, some of the major steps in the development towards the contemporary "standard" model are listed below. This list does not include many neutron-star-based Galactic GRB models (see Ref. 50 for a review), many of which were proposed before the proofs of a cosmological origin of GRB's became solid.

- A generic argument about the "compactness problem" was raised[51] in 1975.
- A self-similar solution of the relativistic blast waves was found in 1976.[52] Although this study was not aimed at GRB's, it laid a mathematical foundation for relativistic GRB blast wave models, and has been extensively applied since the discovery of GRB afterglows.
- Some generic features of a gamma-ray fireball were discussed in 1978.[53]
- A cosmological model of GRB's was formally proposed in 1986,[54–56] involving neutron star–neutron star (NS–NS) mergers as the power source. A pure photon-pair fireball was found to expand relativistically, leaving a quasi-thermal photospheric emission.
- It was found that by including a small amount of baryon contamination, the fireball thermal energy is essentially converted into kinetic energy of the expanding gas.[57]
- It was pointed out that the kinetic energy is re-converted to heat and radiation via shocks.[58,59] This suggestion laid the foundation for the current fireball-shock models. The specific scenario proposed in Refs. 58 and 59 involves shock forming when the fireball is decelerated by the ambient interstellar medium (ISM), which is known as the "external shock scenario." Synchrotron radiation was proposed as the mechanism of GRB emission, noting that this would soften and fade as a power law in time.
- Around the same time, several major GRB progenitor models were further developed or proposed: the NS–NS merger model was further studied in some detail;[60,61] a "failed" supernova model was proposed,[62] which is now known as the "collapsar" model; and a millisecond magnetar (neutron stars with surface magnetic field of order 10^{15} G) model was proposed.[63,140]

- Another fireball shock scenario involving collisions among the individual shells in the fireball outflow, known as the "internal shock scenario" was proposed,[64,65] as a way to resolve the very short time variability problem. An alternative "internal" energy dissipation mechanism involving strong magnetic fields was also proposed.[374]
- The hydrodynamical evolution of a fireball was extensively modeled both numerically and analytically.[66–73]
- The detailed reverse and forward shock synchrotron and inverse Compton spectral components of external shocks were explored,[74,75] indicating the presence of a reverse shock optical synchrotron component (now identified as responsible for prompt optical flashes) and a GeV component due to inverse Compton.
- Based on the synchrotron radiation assumption, GRB radio and optical afterglows were suggested.[76,77] A detailed self-consistent afterglow calculation[78] stressing also the X-ray afterglow was performed shortly before the observational discovery of afterglows. The predictions include the power-law decay of the optical lightcurves as well as prompt optical flashes arising from the reverse external shock during the early stage of the afterglows. All these were verified by the later observations.[24,30]
- Prompted by the discovery of the first GRB afterglows[23–25] the fireball model was studied and tested extensively, and found to be successful in interpreting the observations.[79–83]
- It was suggested[84] that GRB's are likely originated from "hypernovae," and are associated with star forming regions.
- The simplest standard afterglow model was set up,[85,86] and was further developed by incorporating additional effects.[85,87–90]
- Stimulated by the increasing evidence that GRB's are associated with supernovae, the GRB progenitor models involving collapses of massive stars were further developed. The "collapsar model" was extensively studied numerically[91–93] and analytically.[94,95] A variant of the massive-star-collapse model, invoking a delayed GRB with respect to the SN explosion, was proposed,[96] known as the "supranova" model.
- The scenario involving GRB prompt gamma-rays from internal shocks and an afterglow from the external shock as a unified sequence requiring a joint fit was analyzed.[97,98] In parallel to this, extensive efforts were made to model the GRB prompt emission both within the internal shock models[99–103] and the external shock models.[104–107]
- The consequences of GRB's originating from collimated jets was proposed and studied,[108,109] leading to an extensive campaign of theoretical jet modeling.[110–113]
- Prompted by the detection of the optical flashes associated with GRB 990123,[30] GRB 021004[39] and GRB 021211,[40,41] the reverse shock emission was studied in greater detail.[114–123] The recipes to use early afterglow data to systematically infer fireball parameters were recently proposed.[120,122,123]

- The possibility of producing Fe X-ray lines in supranova and in collapsar stellar funnel models was investigated.[124–126,94,127–129]
- Broadband modeling of GRB afterglows was carried out, leading to constraints on some unknown model parameters.[130–134]
- Prompted by the finding of a standard energy reservoir in long GRB's,[36] it was proposed that long GRB's may have a quasi-universal structured jet configuration.[135–137]

1.2. *The multi-disciplinary nature of the GRB field*

The GRB field is almost unique in astrophysics in its multi-disciplinary nature. Involving stellar-scale events located at cosmological distances, GRB's straddle the traditional distance scales, and are a high energy phenomenon which emits a broad-band electromagnetic spectrum extending over at least fifteen decades, as well some possible non-electromagnetic signals (such as cosmic rays, neutrinos and gravitational waves). This makes the GRB field an intersection of many branches in astrophysics.

GRB's were discovered at about the same time when quasars (or broadly speaking active galactic nuclei, AGN's) and pulsars, two other major developments in the 1960's, were discovered. The nature of those two classes of objects was agreed upon soon after their discoveries. AGN's are extra-galactic sources believed to be powered by gigantic black holes, while pulsars are compact neutron stars located in our Galaxy. Since the lack of GRB observational breakthroughs hampered the progress in this latter field, scientists from both the AGN and pulsar community brought to bear their collective wisdom from their own fields to tackle the GRB problem. The GRB neutron star models were developed to fairly sophisticated levels,[50] especially motivated by the reported detection of absorption and emission features in some GRB spectra.[138,139] It turned out that the "classical" GRB's are of cosmological origin, so that the wisdom borrowed from the AGN community (especially about blazars, the most energetic type of AGN's) finally bore the most rewarding fruit. Nonetheless, neutron star models eventually turned out to be partially correct, i.e. one sub-class of such models that invokes ultra-strong magnetic fields[140] turned out to be useful in interpreting the class of soft, repeating GRB's (which appear to observationally distinct from the "classical" cosmological GRB's). These sources, known as "soft gamma-ray repeaters" (SGR's), together with another type of objects called "anomalous X-ray pulsars" (AXPs), are now believed to be "magnetars," neutron stars with super-strong ($\sim 10^{15}$ G at surface) magnetic fields.[141,142] Even within the current paradigm, pulsar models can still provide helpful insights for the understanding of cosmological GRB's. For example, the recent claim of strong polarization of gamma-ray prompt emission in GRB 021206[42] (cf. Refs. 43 and 44) suggests a possible strongly magnetized central engine. A similar conclusion of a strong but less extreme magnetization was reached through a combined reverse-forward shock emission analysis for GRB 990123[120] and GRB 021211.[120,143] This

hints that magnetic fields and/or Poynting flux may play an essential role in GRB's, and that some insights from pulsar wind nebula theories may turn out to be important to unveil the GRB mystery. We discuss this in more detail in Subsec. 4.1.

Besides its close intimacy with the AGN and the pulsar fields, the GRB field is also broadly related to several other branches of astrophysics. First, in the stellar context, since long GRB's are related to the deaths of massive stars, as supported by many observational facts, GRB study is closely related to the fields of stellar structure and evolution,[144] supernovae,[145,146] and supernova remnants (e.g. the external shock invoked in GRB models is the relativistic version of the supernova remnant shock).[147–149] Progenitor studies have stimulated stellar population studies.[150] Central engine studies, on the other hand, have greatly promoted studies about mechanisms for extracting energy from accretion disks or spinning black holes.[151–156] Second, within the galactic context, GRB afterglow lightcurves[90,157,158] and spectral features[34,35,159] probe the properties (e.g. density profile and chemical abundance) of the ambient interstellar medium (ISM) or the prestellar wind. Studying the properties of the GRB host galaxies as well as the GRB locations within the host galaxies brings valuable information about the global star forming history of the universe and the nature of GRB progenitors.[160–162] Third, GRB's are also important objects within the cosmological context. Not only can they play a role similar to AGN's in probing the observed low-redshift universe, but they may also form and can be detectable at much higher redshifts.[163–166] Thus, detecting high-z GRB's would allow us to see into deeper and earlier epochs of the universe, and to probe how the universe ends its dark age and gets re-ionized.[167] Fourth, in high energy astrophysics, the origin of ultrahigh energy cosmic rays (UHECR's) has remained a mystery for decades. Among many other models, a GRB origin[168,169] is a leading model for the so-called "bottom-up models." In this scenario, cosmic ray protons are believed to be accelerated in GRB shocks. The same processes can also generate high energy neutrinos,[65,170–172] so that GRB's are currently regarded by many as the top potential high energy neutrino sources for large area neutrino telescopes being built. Finally, several GRB progenitor scenarios are believed to generate gravitational wave signals,[173,174] and GRB's are also one of the major targets for large gravitational wave detectors. On balance, it would appear that hardly any other field has such a broad range of interactions with other branches of astrophysics.

2. Observational Progress: Facts and Empirical Relations

GRB observational facts have been extensively reviewed, and here we summarize them in an organized, itemized manner. We will try to separate relatively solid observational facts from the empirical statistical laws which are derived from these facts. We divide our discussion into prompt emission and afterglow emission.

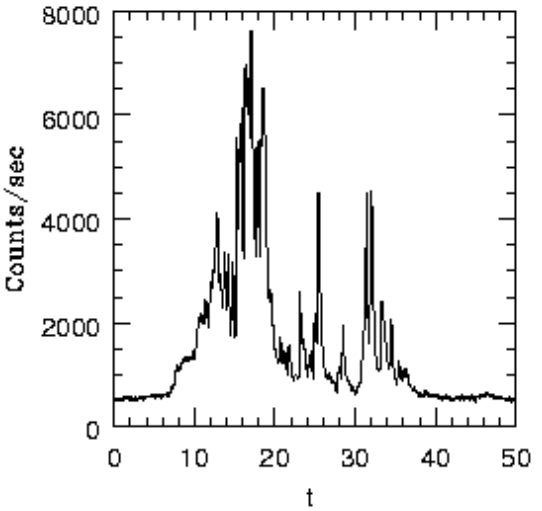

Fig. 1. Typical lightcurve of a BATSE GRB, showing photon count as a function of time (in unit of second), from Ref. 1.

2.1. *Prompt emission*

2.1.1. *Solid facts*

The main characteristics of GRB prompt emission have been collected in Ref. 1 (and references cited in that paper). Below we list the solid facts in itemized form, including also some recent discoveries in the field.[b]

- Temporal properties:
  - (a) Durations (T, technically defined according to different criteria, e.g. T_{90} or T_{50} being the time interval within which 90% or 50% of the burst fluence is detected): they span 5 orders of magnitude, i.e. from $\gtrsim 10^{-2}$ s to 10^3 s, typical values: ~ 20 s for long bursts and ~ 0.2 s for short bursts.
  - (b) Lightcurves: very irregular. Some bursts consist of very erratic, spiky components, while others are smooth with one or a few components. Some bursts contain distinct, well-separated emission episodes. Figure 1 is a typical GRB lightcurve. A compilation of different types of burst profiles can be found in Ref. 1.
  - (c) Widths of individual pulses (δt) vary in a wide range. The shortest spikes have millisecond or even sub-millisecond widths, and $\delta t/T$ could reach as low as $10^{-3} - 10^{-4}$.

[b]We discuss here only the prompt gamma-ray (and X-ray) emission. Prompt optical flashes have been detected in some bursts accompanied with the prompt gamma-ray emission. However, since they are generally interpreted as due to emission from the external reverse shock, we discuss the optical flashes in the early afterglow sub-section below.

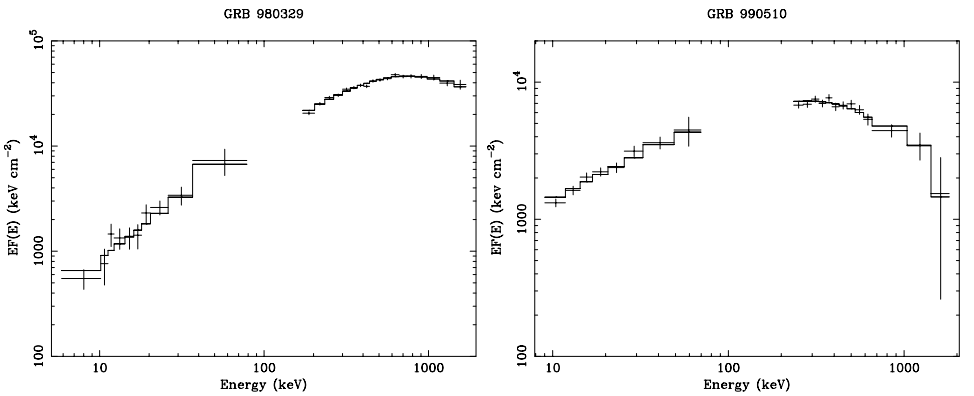

Fig. 2. Examples for the GRB prompt emission spectrum (GRB 980329 and GRB 990510, from Ref. 215).

(d) The vast majority of individual pulses are asymmetric, with leading edges steeper than the trailing edges, although only a small fraction can be visually descerned. By integral means GRB's are asymmetric on all timescales,[175] and on average.[176] Smooth single peak bursts typically have a fast-rise–exponential-decay (FRED)-type lightcurve. This comprises about 7% of all bursts.[177]

- Spectral properties:

(a) The continuum spectrum is non-thermal. Thermal (Planck-like) spectra are ruled out for the great majority of bursts. For most, the spectrum is well described by a smoothly-joining broken power law, known as a "Band-function"[20] (see Fig. 2). Three independent spectral parameters are involved, i.e. a low energy photon spectral index (α), a high energy photon spectral index (β), and the transition energy (E_0) or peak of the energy spectrum for $\beta < -2$ (E_p). This spectral shape is valid both for the integrated emission over the whole burst duration, and for the emission during a certain temporal segment of the burst.

(b) For a sample of 156 bright BATSE bursts with 5500 total spectra,[178] it is found that $\alpha \sim -1 \pm 1$, $\beta \sim -2^{+1}_{-2}$, and the E_p distribution is *lognormal*, centered around ~ 250 keV with a full-width at half-maximum less than a decade in energy. The distributions of α and β are also generally suitable for describing fainter and softer bursts. The "narrow" E_p distribution among different bursts, however, is likely to be influenced by selection effects. Various investigations indicate that the lack of high E_p bursts is likely intrinsic.[179] In the low energy regime, however, the narrowness of the distribution function is mainly due to the "bright" flux-truncation in the sample. Lately, a group of X-ray transient events, which resemble normal GRB's in many aspects but

with E_p around or below 40 keV has been identified.[37,38] These bursts, named "X-ray flashes" (XRF's), extend the E_p distribution to the softer regime and are typically fainter. Globally, it seems that a "narrow" distribution of E_p may be obtained only when one is dealing with a sample of bursts within a narrow peak-flux range.[180,181] On the other hand, since the bright BATSE sample[178] contains many spectra for each burst (at least eight), the results point to an intriguing conclusion, i.e. within a particular burst, the E_p distribution is indeed very narrow. Also bursts tend to soft in time,[182,183] see more discussions in Subsec. 2.1.2.

(c) High energy spectral components: Dozens of BATSE GRB's have also been detected at higher energies, e.g. by EGRET[184] and *Solar Maximum Mission* (SMM).[179] Most of these detections are consistent with a Band spectrum extended to high energies without further breaks. Recently, however, a distinct high energy component was reported in the time-dependent spectra of GRB 941017.[47] This component sticks out at the high energy end of the conventional sub-MeV component, and extends to ~ 200 MeV (due to instrumental upper cutoff) with a photon number index of -1 that cannot be attributed to an extrapolation of the MeV spectrum. More intriguingly, this component does not soften in time together with the low energy component, hinting an independent origin, i.e. perhaps a different emission site. At even higher energies (TeV), an excess of events coincident in time and space with GRB 970417 has been reported by the Milagrito group.[185] In the temporal domain, GeV emission was detected in GRB 940217 lasting 1.5 hours.[21] However, this may be categorized as a high energy afterglow[186] rather than prompt emission.

(d) Spectral features: Absorption and emission features in GRB prompt emission spectra were reported by the Soviet satellites *Venera* 11 and *Venera* 12,[138] and by the Japanese mission *Ginga*.[139] The significance of these features was in the $2-3\sigma$ range, with one claimed at 4σ. However, these were not confirmed by the BATSE team.[1] The only spectral feature in prompt emission reported in recent years was a 3.8 keV absorption feature in GRB 990705[187] detected by BeppoSAX.

- Polarization properties:
  Another important piece of information for electromagnetic radiation is its polarization, which is difficult to measure, especially in the gamma-ray band. Recently, a breakthrough in this direction was reported with the *Ramaty High Energy Spectroscopic Imager* (RHESSI). It was found that the prompt emission of the bright GRB 021206 was strongly polarized, and the claimed polarization degree is $80\% \pm 20\%$[42] (cf. Refs. 43 and 44). Unfortunately, its close angular distance from the Sun (which actually facilitated the detection by RHESSI) prevented the detection of the optical afterglow. A typical radio afterglow was however detected for this source.[188]

- Global properties:
  (a) Angular distribution: Isotropic for all bursts, or for either long or short bursts, respectively.[18,189] There might be a small group of the so-called "long-lag" bursts[176] whose distribution is not isotropic but follows the super-galactic plane. Also, a sub-group intermediate between long and short bursts may show departures from isotropy.[190]
  (b) Intensity distribution: Generally there are two ways to quantify the intensity distribution. One way is through the peak flux distribution (i.e. the so-called $\log N - \log \mathrm{PF}$ plot).[1] This results in a $-3/2$ power law slope at high fluxes, as expected in a homogeneous Euclidean model, and a shallower distribution at lower flux regimes that deviates from this simple homogeneous model. A second way is to find the average value of $V/V_{\max}$, where V and $V_{\max}$ are the volume of space enclosed by the distance of the source and the maximum volume of space enclosed by the distance at which the source could be still detected, respectively. It is found that $\langle V/V_{\max}\rangle$ is less than 1/2 (the value expected for the homogeneous Euclidean model), in the BATSE data.[1] Such an intensity distribution, along with the isotropic angular distribution, are completely consistent with GRB's being cosmological events.[191]
  (c) Event rate: BATSE detected GRB's at a rate of about 1 per day. By correcting for sky coverage and other factors, the actual event rate is ~ 600 per year. Averaging over the whole Hubble volume (for $H_0 \sim 70$ km s^{-1} Mpc^{-1}), this corresponds to an average birth rate of ~ 7.5 Gpc^{-1} yr^{-1} or ~ 0.4 Myr^{-1} galaxy^{-1}, where an average galaxy number density $n_g \sim 0.02$ Mpc^{-3} is adopted. The local GRB birth rate is presumably smaller due to the drastic decrease in the star forming rate at low redshifts, and a widely quoted local ($z = 0$) rate is[191] ~ 0.5 Gpc^{-1} yr^{-1}, or ~ 0.025 Myr^{-1} galaxy^{-1}. We now know that there are related cosmic events such as X-ray flashes (XRF's, see next) that may increase the total event rate by a factor of 2 or 3. This would also increase the GRB-XRF birth rate by a factor 2 or 3. We also know that GRB's are very likely collimated (see Subsec. 2.2). The beaming factor is uncertain dependent on the possible jet structure. Assuming uniform jets, the true-to-observed beaming correction is[36,192] ~ 500 for GRB's. For structured jets, this number is generally smaller. For a quasi-universal Gaussian-type structured jet which fits reasonably well the current GRB-XRF data,[137] the true-to-observed beaming correction factor is 14 for the combined GRB-XRF sample. The real GRB-XRF birth rate (which corresponds to the birth rate of their progenitor) should be increased by a same factor. *Swift* will help to determine both the observed GRB-XRF event rate and the geometric correction factor more precisely.

- Taxonomy:
  (a) The duration distribution is clearly bimodal (Fig. 3), leading to a classification of GRB's into two types, i.e. "long" bursts with $T_{90} > 2$ s, and

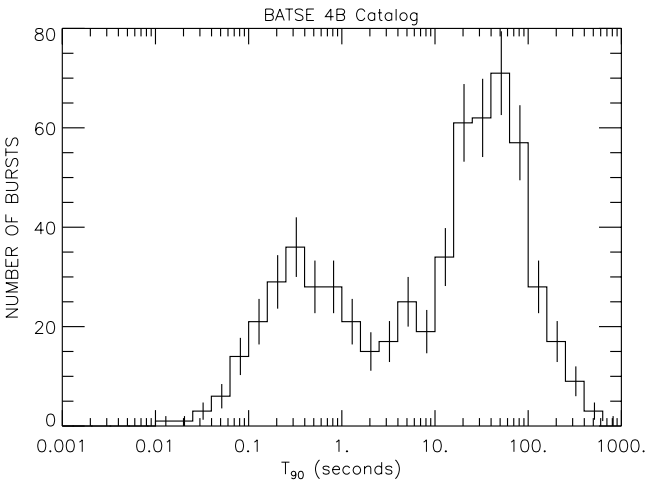

Fig. 3. The bimodal distribution of durations of the BATSE GRB's, from Ref. 189.

"short" bursts with $T_{90} < 2$ s.[19] Such a classification is confirmed in the "hardness" domain, with long bursts typically being soft and short bursts typically being hard.[19] The long and short categories roughly consist of 75% and 25% of the total GRB population. It is commonly speculated that the two types of GRB's may have different origins. All our current knowledge of GRB counterparts (afterglows and host galaxies, etc.) are for long bursts. The nature of short bursts is still a mystery. There have been extensive searches for a third type of GRB's with intermediate durations,[193–195] but the case is not conclusive.

(b) Based on temporal characteristics (lightcurves), it is difficult to categorize GRB's into well-defined sub-types. Nonetheless, some phenomenological classes have been suggested.[1] These classifications are however not fundamental. The differences among different types are likely caused by different behaviors of the central engine, different emission sites, or different viewing angles, not necessarily reflecting different types of progenitor.

(c) Based on the spectral hardness, recently the so-called "X-ray rich GRB's" and "X-ray flashes" (XRF's)[37,38] have been widely discussed as forming an apparently new type of transient event with respect to the conventional GRB's (Fig. 4). Whether they are the product of intrinsically different mechanisms or are simply a natural extension of GRB's towards softer and fainter regimes[196] is still unclear, but evidence is mounting that XRF's and GRB's are related events.

(d) The "long-lag" bursts such as GRB 980425/SN 1998bw may belong to a subtype of long bursts at closer distances (associated with the super-galactic plane).[176]

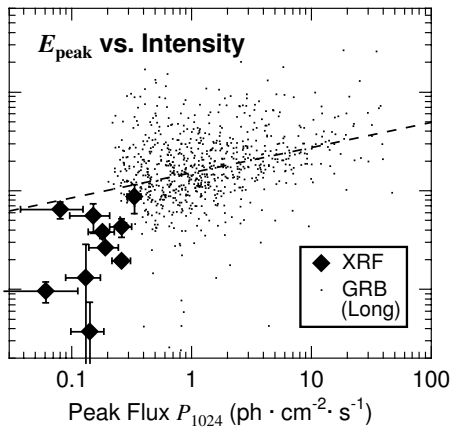

Fig. 4. The E_p-fluence distribution of the conventional long GRB's as well as the recently identified X-ray flashes, from Ref. 38.

2.1.2. *Empirical laws*

Besides the above relatively solid observational facts, some "secondary" empirical relations have been derived through various statistical analyses. These relations are potentially useful, posing important constraints on the theoretical models. Some are even helpful for deriving some important but unknown parameters. Below is a non-exhaustive list. Items (1)–(4) are for the temporal information alone, while items (5)–(8) include the combined temporal and spectral information.

- Power density spectra (PDS's) are powerful tools to quantify GRB temporal characteristics. It is found that the averaged PDS for many long bursts follows a power law of index $-5/3$ over almost two decades in Fourier frequency, with a break around ~ 1 Hz[197,198] at the higher end. This indicates that variabilities shorter than ~ 1 s are smeared out, for reasons as yet unclear.
- By tracking individual pulses, the pulse width distribution for long bursts is found to be lognormal.[199,200,574,575] with the peak variability timescale $\delta t_{\rm pk} \sim 1$ s. In some GRB's, there are so-called "quiescent times" where no gamma-rays are emitted, and the durations of such quiescent episodes are positively correlated with the duration of the emission episode immediately following the quiescence.[201] By subtracting the quiescent times, the intervals between pulses also have a lognormal distribution.[199,200,574,575]
- The temporal behavior can also be quantified through a "variability" parameter, defined by some appropriate statistical definitions.[202,203] An intriguing empirical relation is that the more variable bursts (with larger V parameters) tend to have higher intrinsic isotropic luminosities L (derived from afterglow spectroscopic measurements), i.e. $L \sim V^{3.3}$. Although such a correlation is derived from a small parent sample, it exhibits a "Cepheid"-like correlation and makes "variability" a possible distance indicator of GRB's (Fig. 5, bottom left).

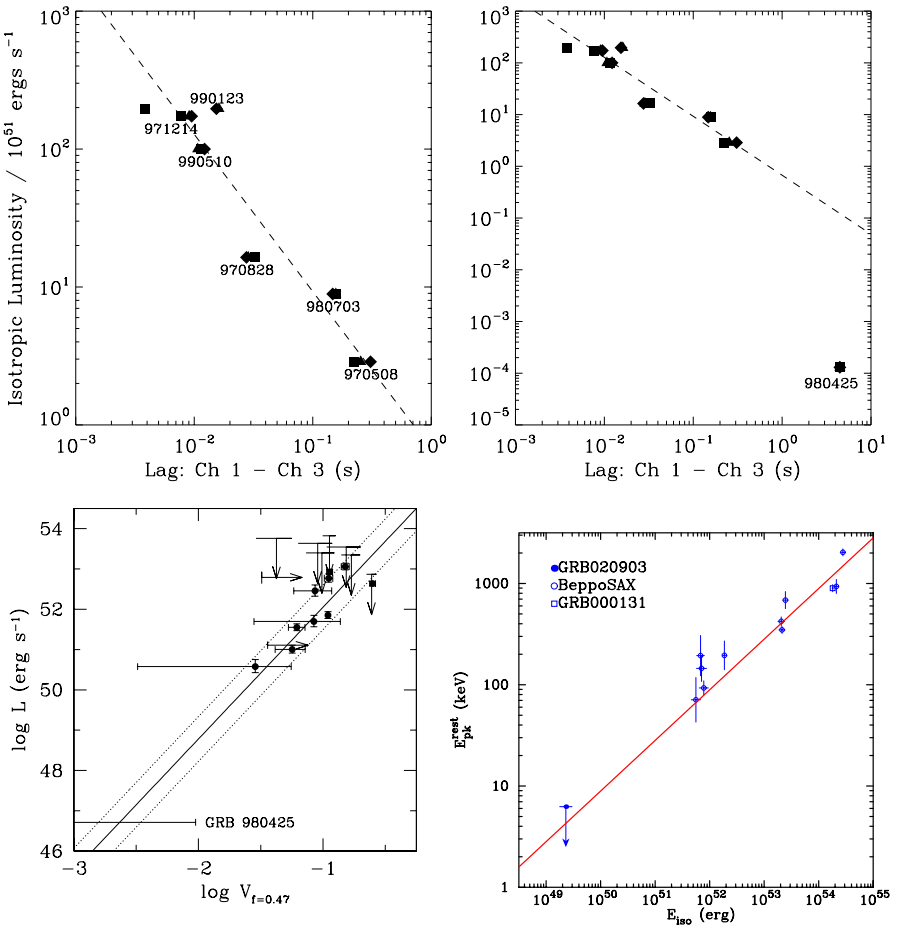

Fig. 5. Three "Cepheid"-like correlations found in GRB data, which could be adopted as rough distance indicators and redshift estimators. (1) Top: the spectral lag-luminosity correlation (from Ref. 210); (2) Bottom left: the variability-luminosity correlation[202] (from Ref. 203); and (3) Bottom right: the E_p-luminosity correlation[215] which extends to XRF's (from Ref. 217).

- The cumulative lightcurves can reveal the global energy release rate of the central engine. For most bursts, the cumulative lightcurve has a constant slope as a function of time.[204] There are two smaller groups (comprising $\sim 4.8\%$ and $\sim 2.8\%$ of the whole population studied) whose cumulative lightcurves increase with time more rapidly or more slowly than the linear increase, respectively.[205]
- In many bursts (although not exclusively) there is a clear trend of spectral evolution. There are two types of evolution. One is "hard-to-soft,"[182,183] in which hard emission leads the soft emission, and another is "tracking,"[206] for which the spectral hardness tracks the intensity. In either case, the E_p (derived by fixing both α and β) decay is found to be exponential with photon fluence, and within a burst, the decaying constant is invariant from pulse to pulse.[207]

- For asymmetric pulses, the pulse peak times migrate towards later times at lower energies, and the pulse width becomes wider.[208,209] For 6 GRB's with spectroscopic redshift measurements, it is found that the spectral lag is correlated with the isotropic luminosity as $L_{53} \approx 1.3 \times (\tau/0.01\ \mathrm{s})^{-1.14}$, making spectral lag another possible "Cepheid-like" luminosity indicator[210] (Fig. 5, top).
- The low energy spectral index α is found to correlate with the E_p of time-resolved GRB spectra, although the trend is still controversial.[211,212]
- The E_p's are found to be positively correlated with the isotropic luminosity.[213,214] With the spectroscopic redshifts of 12 BeppoSAX bursts, a firm correlation, i.e. $E_p \propto E_{\rm rad}^{0.52\pm0.06}$ (where $E_{\rm rad}$ is the total isotropic energy radiated in the X-ray and gamma-ray range, i.e. 1 keV–10 MeV) is found,[215] which may be taken as yet another "Cepheid-like" luminosity indicator[216] (Fig. 5, bottom right). The relation is found to extend from the hard GRB to the soft XRF regime.[217,218] However, due to the small sample effect, the issue is not conclusive, yet. Furthermore, GRB 980425 had an $E_p \sim 70$ keV and $E_{\rm iso} \sim 10^{48}$ ergs, which apparently does not fit into the simple correlation.
- Another redshift indicator is found in the gamma-ray spectral data by taking into account of excesses from the exact power-law dependence.[219]
- Luminosity function: Although the present afterglow spectroscopic redshift measurements provide too small a sample to perform a direct measurement, several attempts have been made to derive the luminosity function of GRB's. Although without full consensus, it is likely that the luminosity is a power-law, with a possible steepening at high luminosities.[191,220,221,176,222]
- Using the $V - L$ correlation and the spectroscopic redshift data,[215] it is found that there might exist a cosmological evolution effect for GRB properties, such as luminosity (or even luminosity function), total energy, E_p, duration, etc., although more spectroscopic redshift measurements are needed and selection effects need to be further clarified.[221,223]

2.2. *Afterglow emission*

2.2.1. *Observational facts*

The afterglow observations have been extensively reviewed in Refs. 2, 3, 9 and references therein. Here we outline some of the main results in itemized forms. These include

- Global properties:
  (a) Afterglows are (quite) broad-band, having been detected in the X-ray, the optical/infrared and the radio bands. In each band, the lightcurve generally displays a power-law decay behavior. X-ray afterglows are exclusively decaying when they are detected. Optical afterglows are generally decaying, with an initial early rising lightcurve having been caught in only a few bursts (e.g. GRB 990123).[30] In contrast, an initial rising lightcurve in the radio band

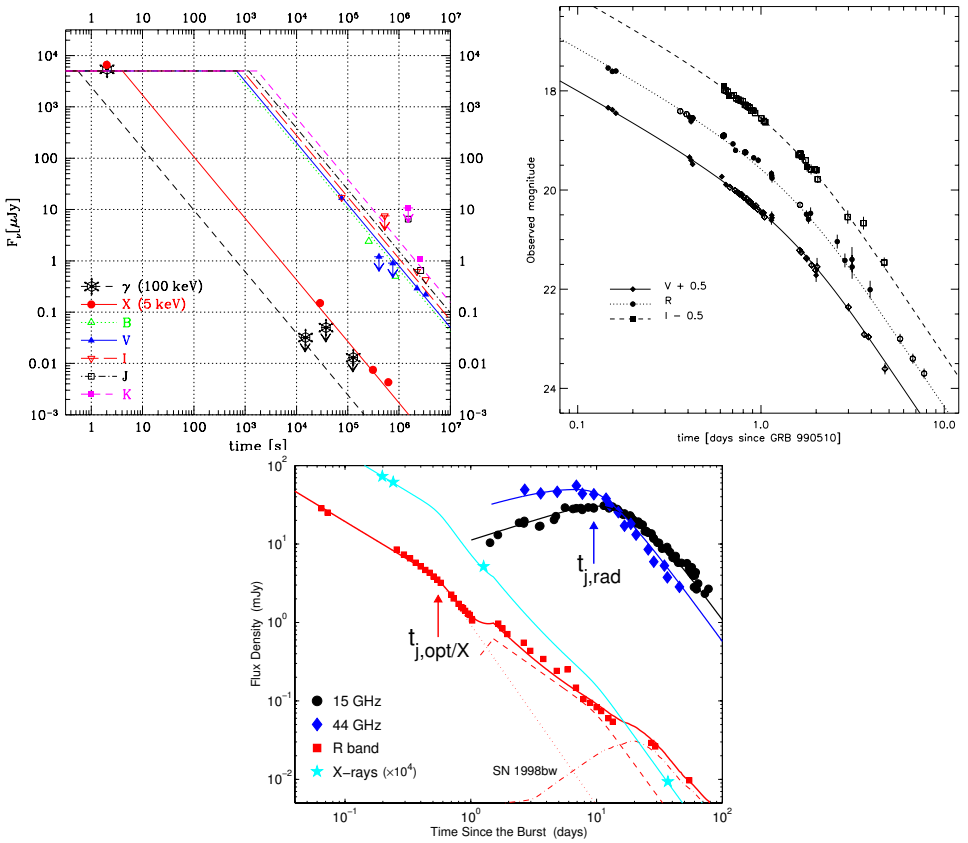

Fig. 6. Several well-monitored broad-band afterglow lightcurves. Upper left: GRB 970228 (Ref. 79 and references therein); Upper right: GRB 990510 (from Ref. 33), notice the achromatic jet break; Bottom: GRB 030329. Reprinted by permission from Nature (Ref. 48) copyright (2003) Macmillan Publishers Ltd. Notice irregularities in the lightcurves and the association of a supernova bump.

has been detected in many GRB's (followed by a canonical decay), and the typical timescale for the radio afterglow to reach the peak is $\sim$10 days.[224] Figure 6 gives several examples of broadband afterglow lightcurves.

(b) Very often there are various types of deviations from the simple power law decay. These include steepenings, bumps and wiggles (e.g. GRB 021004; GRB 030329).

(c) Not all bursts have afterglows detected in all of the three main bands. X-ray afterglows are the most commonly detected. About 60% of BeppoSAX bursts with X-ray afterglow detections are detected in the optical band. The other $\sim$40% of GRB afterglows are optically dark, i.e. there are no optical transients identified with sufficient exposure. These are dubbed as "dark bursts." In the HETE era, the fraction of dark bursts gets smaller, e.g. 10%.[225] Radio afterglows are detected in about half of all GRB afterglows.

(d) Essentially every GRB with an afterglow detection has an underlying host galaxy. The GRB host galaxy properties (magnitude, redshift distribution, morphologies, etc) are typical of normal, faint, star-forming galaxies.[13] The GRB afterglow's positional offsets with respect to the host galaxy are consistent with GRB's being associated with the star-forming regions in the galaxies.[160]

(e) GRB's are at cosmological distances. Their redshifts are usually measured from the emission features of the host galaxies or the absorption features imposed on the afterglow continuum. As of October 2003, there are about 33 redshift measurements, and detected redshifts range from 0.168 for GRB 030329 (or 0.0085 for GRB 980425) to 4.5 for GRB 000131. See Ref. 226 for a compilation of the afterglow data, including redshifts.

(f) At least some GRB's are associated with supernova explosions. A famous example was GRB980425/SN1998bw association.[28,29] SN 1998bw was a peculiar, energetic Type Ib/c supernova. Using it as a template, other possible associations have been claimed by identifying a reddened bump in the optical afterglow lightcurves of GRB 980326,[227] GRB 970228,[228,229] GRB 000911,[230] GRB 991208,[231] GRB 990712,[232] GRB 011121,[233] and GRB 020405,[234] which may be attributed to supernova remnant contribution. Very recently, an unambiguous supernova signature has been detected in the $z = 0.168$ GRB 030329, firmly establishing the GRB-SN association in this object[45,46] (Fig. 7).

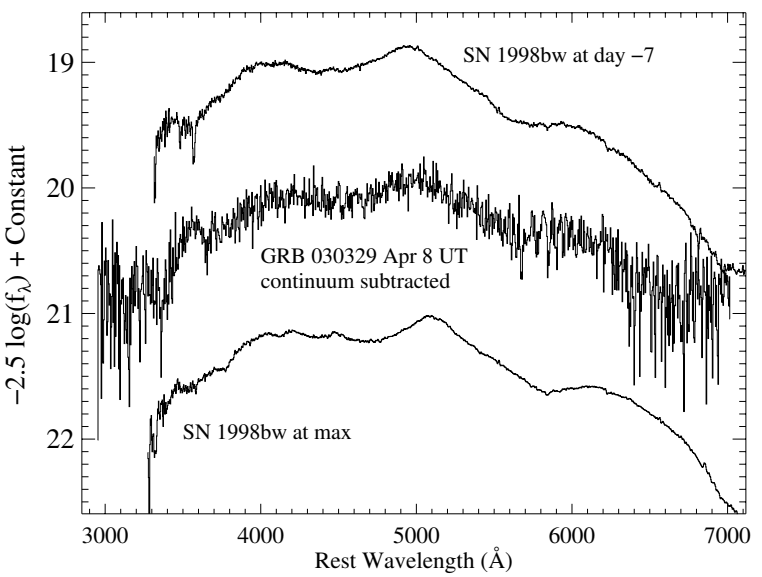

Fig. 7. The optical afterglow spectrum of GRB 030329 clearly reveals the spectral signature of Type-Ib/c supernovae (e.g. SN 1998bw) at day-10 since the burst trigger. This establishes a firm association between long GRB's and SNs (from Ref. 45).

- X-ray afterglows:
  (a) The continuum spectrum is essentially a power law. By writing $F_{\rm x}(t,\nu) \propto t^{\alpha}\nu^{\beta}$, one has[235] $\alpha \sim -0.9$ and $\beta \sim -1.4$, both with a range of scatter. Late time lightcurve flattening was observed in a few bursts such as GRB 000926.[133]
  (b) As of October 2003, X-ray emission line features have been claimed to be detected in the X-ray afterglows of 6 bursts with moderate significance ($<4.5\sigma$, typically 3σ), i.e. GRB 970501,[236] GRB 970828,[237] GRB 991216,[34] GRB 000214,[238] GRB 011211,[35] and GRB 020813.[239]
  (c) Analyses of BATSE data reveal soft tails in some GRB's which are likely to be early X-ray afterglows.[240]
- Optical afterglows:
  (a) Expressing the optical flux as $F_{\rm opt}(t,\nu) \propto t^{\alpha}\nu^{\beta}$, one has[226] $\alpha \sim -1$ (at early times) and $\beta \sim -0.7$, both with a range of scatter.
  (b) In some bursts, a clear lightcurve steepening is seen after some time $t_{\rm break}$ of order of days. The break is achromatic over different bands, and the temporal index after the break is typically -2 with some scatter. This break is typically attributed to the presence of a jet, and is termed the "jet break." At later times, the decay rate gradually slows down until finally reaching a constant level due to the contribution of the host galaxy.
  (c) Other irregular temporal features are occasionally seen in various bursts.[241] These include,[226] for example, a substantial re-brightening in GRB 970508, an achromatic bump signature in GRB 000301C, wiggles in GRB 021004 and step-like features in GRB 030329.
  (d) Polarization: As of October 2003, 8 GRB optical afterglows have been detected to be polarized,[242–244] i.e. GRB 980425, GRB 990510, GRB 990712, GRB 010222, GRB 020405, GRB 020813, GRB 021004 and GRB 030329. The degree of polarization is small, typically several per cent. Large polarization angle changes (by nearly 90 degrees) were found in GRB 021004 and GRB 030329.[245,244]
  (e) Early optical flashes: Optical afterglow observations typically have started hours after the burst trigger due to technical reasons. However, early optical afterglows of two GRB's, i.e. GRB 990123[30] and GRB 021211,[41,40] were detected within 100 seconds after the GRB trigger. In both cases, the early afterglows indicate a steep decay with index ~ 2. For GRB 990123, a sharp rising lightcurve was also detected, and the V-magnitude is ~ 9 at the peak.[30] The early optical afterglow of a third burst, GRB 021004, was detected about 3 minutes after the trigger, which displays a very shallow initial lightcurve decay.[39]
- Radio afterglows:
  (a) In the radio band, the spectral index is generally positive for the observations which are typically in the 5 and 8.5 GHz bands. The lightcurves usually do

not follow a simple power law decline.[246] Some sources can be observed on timescales of years, and a late-time flattening (with respect to the standard fireball model) is often observed.[247]

(b) Prompt, short-lived radio flares have been detected in several bursts.[248,249]

(c) At early times, radio afterglows show strong fluctuations which are suppressed at later times.[250] This can be interpreted as being due to interstellar scintillation effects.[251] The detection of such an effect in GRB 970508 clearly suggested that the source was expanding super-luminally, which gave a solid observational proof for the fireball shock model.[252]

- Taxonomy:
  According to their afterglow data, long GRB's may be further classified into several sub-types. Below are two such suggested classifications. They do not necessarily reflect intrinsically different groups (e.g. with different progenitors), but might be caused by environmental effects. So far there is no attempt to categorize GRB sub-types using combined prompt emission and afterglow data, but such efforts should be useful for identifying real sub-types.

  (a) Optically dark bursts: A fraction of GRB's with precise localizations do not have bright enough optical afterglows to be detectable. Possible reasons include dust extinction, high redshift, or intrinsically faint nature. Evidence for dust extinction has been collected for some bursts.[253] On the other hand, observations for GRB 020124 and GRB 021211 indicate that at least some optically dark bursts are simply bursts with intrinsically faint afterglows.[254,41] Indirect evidence for high redshift is also available for the recent dark burst GRB 031026.[255]

  (b) Fast-fading GRB's: Several bursts (e.g. GRB 980519, GRB 980326) show a steep afterglow decay ($\alpha \sim -2$) in their early phase. They do not fit into the "standard energy reservoir" scenario,[36] and may constitute a peculiar class of GRB's.[256]

2.2.2. *Empirical laws*

As in the prompt emission studies, some secondary empirical laws have been discovered in the afterglow data. Below is a non-exhaustive list.

- Perhaps the most intriguing finding is that long GRB's seem to have a standard energy reservoir. This conclusion is based on the model that GRB's are collimated and that lightcurve steepenings are due to a jet. From a simple toy model, one can derive the so-called jet opening angle θ_j using the jet break time. For those bursts whose redshifts, and hence the total gamma-ray energies (assuming isotropic radiation) in the prompt phase, $E_{\gamma,\rm iso}$, are measured, it is found that $E_{\gamma,\rm iso}\theta_j^2$ is essentially a constant, which means that the geometrically corrected gamma-ray energy for different bursts is essentially an invariant[36,256] (Fig. 8).

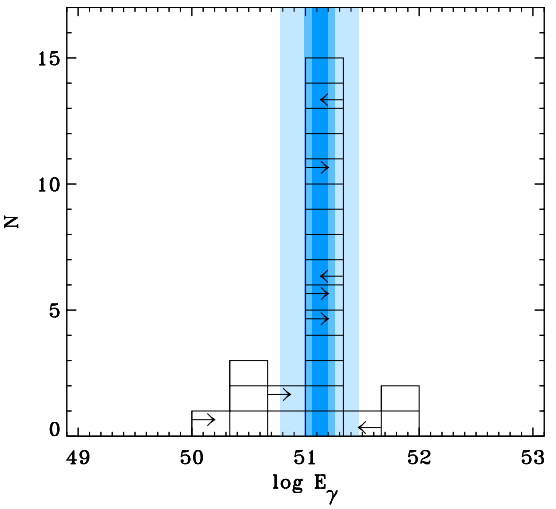

Fig. 8. The geometry corrected gamma-ray energy (i.e. $E_\gamma \sim E_{\gamma,\rm iso}\theta_j^2/2$, where $E_{\gamma,\rm iso}$ is the total energy emitted in gamma-rays assuming isotropic radiation, and θ_j is the jet opening angle inferred from afterglow lightcurves) is found to be a constant in many bursts, referring to a standard energy reservoir of long GRB's.[36] Shown is the distribution of E_γ with the latest data (from Ref. 256).

- The total GRB fireball energy should be $E_{\rm tot} \geq E_\gamma + E_K$, where E_γ is the energy released as gamma-rays in the prompt phase, and E_K is the kinetic energy which is left over after the prompt phase and is dissipated in the afterglow phase. The latter component (E_K) may be directly measured using broadband afterglow fits[131,132] or via analyses of the X-ray afterglow data for different GRB's at the same afterglow epoch.[257–259] Both approaches lead to the same conclusion that E_K is also a standard value for different bursts. Combining this and the previous argument suggests that long GRB's have a standard energy reservoir.
- An anti-correlation has been derived between the GRB spectral lag and the jet opening angle.[260] This relation, when coupled with the Cepheid-like correlation for the spectral lag,[210] is another manifestation of the anti-correlation between the isotropic luminosity and the jet opening angle, which is the direct consequence of the standard energy reservoir relation.[36]
- Other model parameters can also be inferred from broadband afterglow fits. Consensus is emerging about the values of some of these parameters. For example, the environment of most GRB's appears to be an interstellar medium (ISM) with an approximately constant density, although the afterglows of a small group of GRB's are consistent with a stratified wind-type medium (with a density profile $n \propto r^{-2}$).[131,132] The ISM density appears not to vary greatly among bursts.[132] In the shock region, the portion of the energy which goes into the electrons (denoted as ϵ_e) is typically ~ 0.1 with some scatter, while the energy portion that goes into the magnetic fields (denoted as ϵ_B) is typically ~ 0.01 or less.[131,132]

3. Theoretical Progress: Standard Fireball Shock Model

In this section, we briefly introduce the key ingredients of a generic standard GRB fireball shock model in a pedagogical manner. This theoretical framework is the most widely used for interpreting the current GRB and especially afterglow observations. Its "standard" nature is the product of its predictive power and success in passing various observational tests. This section focuses on those aspects of the model which are robust and have the least uncertainties. Discussion of some of the less certain aspects of the fireball model are deferred to Sec. 4. The contents of this section overlap to some significant degree several other reviews found in Refs. 4, 5 and 9.

3.1. *Relativistic bulk motion*

A first ingredient of the standard model is that the emitting material responsible for GRB's and afterglows must be moving relativistically. This is a consensus of *all* cosmological GRB models (even if these models are non-standard and differ considerably in technical details) and even for the old galactic halo models. The arguments leading to this conclusion are straightforward,[51,54–56] as discussed by, e.g. Refs. 4, 5, 7 and 14. For a typical GRB gamma-ray fluence $F \sim 10^{-6}$ erg cm^{-2} and distance $D \sim 3$ Gpc, the total isotropic gamma-ray energy released is typically $E = 4\pi D^2 F \sim 10^{51}$ ergs. Naively (without relativistic motion), the scale of the emission area is $c\delta t = 3 \times 10^8$ cm ($\delta t/10$ ms). Assuming that a fraction f_p of photons is above the two-photon pair production ($\gamma\gamma \to e^+e^-$) threshold [$\epsilon_1\epsilon_2(1 - \cos\theta_{12}) \geq 2(m_e c^2)^2$, where m_e is the electron mass, ϵ_1 and ϵ_2 are the energies of two photons, and θ_{12} is the angle between the momenta of the two photons], and using an approximate pair production cross-section of the order of the Thomson cross section $\sigma_T = 6.25 \times 10^{-25}$ cm^2, the pair-production optical depth is huge, i.e. $\tau_{\gamma\gamma} = f_p \sigma_T F D^2/(c\delta t)^2 m_e c^2 \sim 10^{15} f_p (F/10^{-6}\ \mathrm{erg\ cm^{-2}})(D/3\ \mathrm{Gpc})^2\ (\delta t/10\ \mathrm{ms})^{-2}$. Thus, the gamma-rays should have been attenuated in the source before traveling through the universe and reaching the earth. The only way to get rid of this apparent paradox is by invoking relativistic bulk motion, i.e. the GRB emitting region as a whole moves towards us observer with a high Lorentz factor.

The relativistic motion eases this "compactness problem" in two ways. Suppose that the emitting region (e.g. an ejected shell of relativistic material) flies towards the observer with a bulk Lorentz factor Γ. First, the photon energy is blue-shifted by a factor of Γ, so that in the shell comoving frame, the bulk of gamma-rays as observed are actually X-rays. This greatly decreases the number of photons above the pair production threshold, i.e. f_p drops by a factor of $\Gamma^{2(\alpha-1)}$, where $\alpha \sim 2$ is the observed photon number spectral index, i.e. $N(E_{\rm ph})dE_{\rm ph} \propto E_{\rm ph}^{-\alpha} dE_{\rm ph}$. Here the factor $(\alpha - 1)$ arises from the integration leading to the above-threshold photon numbers, and the factor 2 takes into account the contributions of both the test photons and the target photons for pair production.[261] A second effect introduced by considering relativistic motion is that the real physical scale of the emission

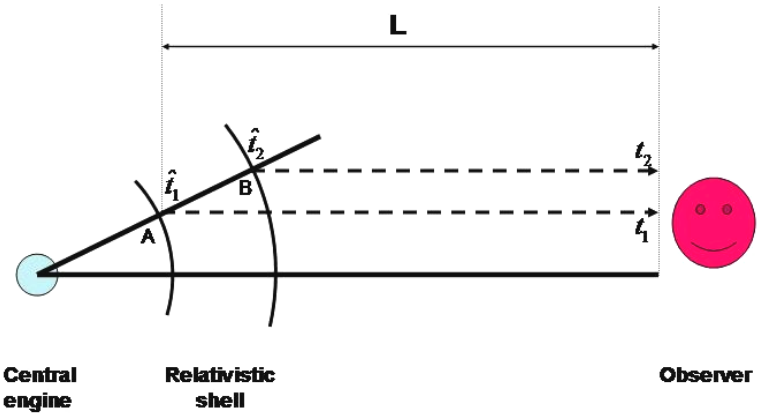

Fig. 9. The geometric configuration among the GRB central engine, relativistic emitting shells and the observer.

region is $\Gamma^2 c\delta t$ for an observed timescale of δt (see Subsec. 3.2 below). So altogether the pair optical depth drops by a factor of[c] $\Gamma^{2+2\alpha} \sim \Gamma^6$. For typical bursts, $\Gamma \geq 100$ is required to have $\tau_{\gamma\gamma} < 1$. This is only a rough estimate; more sophisticated analyses result in various lower limits of Γ in different bursts.[262–265,261] The typical Γ required to satisfy the observations is of order ~ 100. Hence, GRB's involve the fastest bulk motions known so far in the universe.

There are other pieces of evidence of relativitic motion of the fireball. One is from the radio afterglow data that initially shows large interstellar scintillation but gets supressed later on. This presents a clear evidence of superluminal expansion of the fireball caused by relativistic motion.[252] Another is from the analyses of early afterglow reverse shock data.[114,120] These analyses directly point to a large initial Lorentz factor of the fireball (see Subsec. 5.1 for more discussions).

3.2. *Two reference frames, three timescales*

A typical GRB problem involves three major physical elements (Fig. 9): a central engine, a relativistically moving shell (ejected by the central engine) which produces the GRB emission, and an observer. There are essentially only *two* inertial frames. One is the rest frame of both the central engine and the observer (aside from a cosmological redshift factor, which is small compared to special relativistic effects), and the other is the rest (comoving) frame of the relativistic shell or ejecta. The physical quantities (e.g. scale length and time) as viewed in the two inertial frames are different, and are related through special relativistic Lorentz transformations. For example, a length scale Δ' in the comoving frame is converted to $\Delta = \Delta'/\Gamma$ along the shell's moving direction in the rest frame of the central engine.

[c]Notice that the expressions of the optical depth in some of the previous reviews[4] is incorrect,[261] given the above definition of α.

Similarly there are only two sets of clocks attached in both inertia frames, so that $dt' = d\hat{t}/\Gamma$, where $d\hat{t}$ and dt' are the time intervals elapsed for *the same pair of events* in the central engine/observer frame and the comoving frame, respectively. However, in the GRB problem, there is a third relevant time scale involved. The complication comes from the propagation effect. In general, we can consider a shell emitter moving with a dimensionless speed β, at an angle θ with respect to the line of sight of the observer (Fig. 9). In the rest frame of the central engine/observer, the shell emits a first photon towards the observer at the time $\hat{t}_1$ at the location A (the radius r), and emits a second photon towards the observer at time $\hat{t}_2$ at the location B (the radius $r + dr$), as recorded by clocks precisely adjusted in this inertia frame. The time interval for emitting these two successive photons is $d\hat{t} = \hat{t}_2 - \hat{t}_1 \simeq dr/c$. Let us assume that the distance between the observer and the location A is L. The first photon arrives at the observer at $t_1 = \hat{t}_1 + L/c$, while the second photon arrives at the observer at $t_2 = \hat{t}_2 + (L/c - \beta\mu\, d\hat{t})$ where $\mu = \cos\theta$. These two times are also measured by clocks precisely adjusted in the same inertial frame. The time interval for the observer to receive the two adjacent photon signals is simply

$$dt = (1 - \beta\mu)d\hat{t} \simeq d\hat{t}/2\Gamma^2 = dr/(2\Gamma^2 c)\,, \tag{1}$$

where in deriving the second part of the equation, we have assumed $\Gamma = (1 - \beta^2)^{-1/2} \gg 1$ and $\theta \ll 1$ (so that $\mu \sim 1$). Notice that Eq. (1) is a pure propagation effect, which is also valid for the non-relativistic case. It is not noticeable in the Newtonian case, simply because $d\hat{t} \simeq dt$ when $\beta \ll 1$. In the relativistic regime, this effect becomes very important, and Eq. (1) is one of the fundamental arguments used to solve the "compactness problem" as discussed above in Subsec. 3.1.

It is worth noticing that $d\hat{t}$ and dt are two different times in *the same inertial frame*, which describe two different pairs of events. $d\hat{t}$ describes the time for the shell to *emit* two photons, while dt describes the time for the observer to *receive* the same two photons. Since $d\hat{t}$ also describes the actual time of the shell behavior (e.g. when the shell moves to the location A or location B), the best way to refer to these two times is to call $d\hat{t}$ the "time in the rest frame of the central engine" or "the time in the fixed (or lab) frame," while calling dt the "observer's time." In the literature, the latter is sometimes called "the time in the observer's frame," which is in principle right, but this is not the exact distinction between dt and $d\hat{t}$. Sometimes in the literature $d\hat{t}$ is called "the time in the burster's frame." This can be confusing to some readers, since the "burster" may be understood either as the central engine powering the burst (in which case the definition is correct), or it may be understood as the flying shell which emits the burst (in which case the definition is incorrect). Also notice that in some articles, the $d\hat{t}$ discussed here is denoted as dt, while the dt discussed here is denoted as dT, $dt_{\rm obs}$ or $dt_{\oplus}$. Since the observer's time is the most relevant one to describe the phenomenon, it is more convenient to define it as dt, the most straightforward notation. In the rest of this review, we will adhere to such a notation system.

A variant of dt which is also widely discussed in GRB problems is the so-called "angular time" $dt_{\rm ang}$. Instead of fixing θ and varying r (or $\hat{t}$), in some problems one needs to fix r (and hence $\hat{t}$) but vary θ. This is relevant for the problem in which a shell-like emitter (with a fixed radius r) is illuminated instantaneously at the same engine-frame $\hat{t}$, while the observer sees a long-duration emission caused by the time delay of radiation coming from higher latitudes. In a similar manner as for the derivation of dt, one has $dt_{\rm ang} = (r/c)\sin\theta d\theta$, or $t_{\rm ang} = (r/c)(1-\cos\theta)$. For relativistic motion ($\Gamma \sim 1/\theta \gg 1$), one has $\theta \sim \sin\theta \sim 1/\Gamma$, so that $t_{\rm ang} \simeq (r/c\Gamma^2)$, which is of the same order as t [Eq. (1)].[266,97] Notice again that $t_{\rm ang}$ is also a pure propagation effect and is also applicable for the Newtonian case.

The third time scale is the comoving time of the shell, measured by another set of clocks. It is related to the engine-frame time $d\hat{t}$ through

$$dt' = d\hat{t}/\Gamma = dt/\Gamma(1-\beta\mu) = \mathcal{D}dt \simeq 2\Gamma dt\,, \tag{2}$$

where $\mathcal{D} = [\Gamma(1-\beta\mu)]^{-1}$ is the Doppler factor (since the observed radiation frequency is boosted by the same factor with respect to the frequency in the comoving frame, i.e. $\nu = \mathcal{D}\nu'$). The final approximation in Eq. (2) is again for $\Gamma \gg 1$ and $\mu \sim 1$.

3.3. *Fireball evolution, characteristic radii and times*

The evolution of generic fireballs expanding into an ambient medium has been extensively studied,[53,54,56,55,66,68,70–73] assuming a fireball composition of photons, electron/positron pairs and a small amount of baryons (but negligible magnetic fields). In a simplest toy model, we assume the following input parameters: (a) an average constant luminosity of the central engine, L; (b) the duration of the central engine energy injection, T, so that the total energy of the fireball is $E = LT$ and the initial width of the whole fireball shell in the fixed frame is $\Delta_0 = cT$; (c) the variability time scale, $t_v \sim 1$ ms $\ll T$, which is due to the intermittent nature of the fireball central engine, as reflected by the spiky, irregular GRB lightcurves (Subsec. 2.1.1). The fireball shell therefore actually consists of many mini-shells; (d) the average mass loading rate, $\dot{M}$, so that the so-called dimensionless entropy is $\eta = L/\dot{M}c^2$; and (e) the particle (usually hydrogen) number density of the ambient medium, n (in the simple toy model here, n is taken as a constant, which is typical for an interstellar medium, ISM). The evolution is characterized by several characteristic radii (see Fig. 10 for a cartoon picture), which we will discuss in turn below.

- Initial state ($r = r_0$): The base of the fireball flow is connected to the GRB central engine, a black hole–torus system or a rapidly rotating magnetar. Assuming a $10M_\odot$ black hole, the initial length scale (taken 3 Schwarzschild radii)[72] is $r_0 \sim 6GM/c^2 \sim 10^7$ cm. This radius is also the typical width of the mini-shells, i.e. $\delta_0 = ct_v = (3\times10^7\ {\rm cm})\ (t_v/1\ {\rm ms})$. The initial state of the fireball is hot, with photons and pairs in equilibrium with a temperature $\mathcal{T}_0 = (L/4\pi r_0^2\sigma)^{1/4} \sim (10^{10}\ {\rm K})\ L_{51}^{1/4} r_{0,7}^{-1/2}$. Hereafter the convention $Q = 10^n Q_n$ will be adopted in c.g.s.

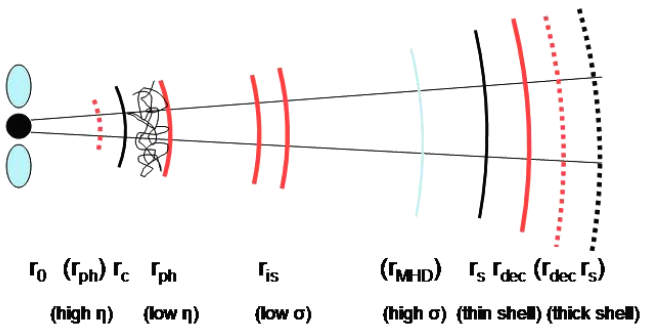

Fig. 10. Various characteristic radii in a generic relativistic fireball.

units (e.g. L_{51} means luminosity in unit of 10^{51} erg s^{-1}). In the initial state, the baryons are essentially at rest with respect to the central engine.

- End of ejection ($r = \Delta_0$): After the central engine constantly ejects energy (with luminosity L) for a time T, the fireball radius is $r = \Delta_0 = cT = (3 \times 10^{10}$ cm$)\ (T/1$ s).
- Coasting radius ($r = r_c$): As the fireball shell expands, the baryons will be accelerated by radiation pressure. The fireball bulk Lorentz factor increases linearly with radius, until reaching the maximum Lorentz factor Γ_0, which is usually the dimensionless entropy η, but sometimes lower than η due to the limited radiation pressure.[72,73] The fireball finally coasts with a constant Lorentz factor Γ_0 at the coasting radius,[66,68,71] $r_c = \Gamma_0 \Delta_0 \sim (10^{13}$ cm$)\ (T/1$ s$)(\Gamma_0/300)$ (this is the radius at which the entire mass of the fireball has achieved the coasting Γ). During acceleration, since all the materials essentially move with speed of light, the shell width in the fixed frame remains constant, $\Delta = \Delta_0$. The comoving shell width $\Delta' = \Gamma\Delta_0$, on the other hand, increases with radius linearly. The above conclusions hold assuming that the fireball shell evolves as a whole. However, if the mini-shells evolve separately, i.e. for well-separated GRB pulses, these mini-shells are likely to evolve independently, and coast at a smaller radius $r_c = \Gamma_{\rm sh}\delta_0$, where $\Gamma_{\rm sh}$ is the Lorentz factor of that particular mini-shell, and $\delta_0 = ct_v$ is the width of the mini-shells.
- Photospheric radius ($r = r_{\rm ph}$): As the fireball shell expands, the photon number density and the typical photon energy drop. At a certain radius, the photons become optically thin to both pair production and to Compton scattering off the free electrons associated with baryons entrained in the fireball. At this radius, although much of the initial energy is converted to the kinetic energy of the shell, some energy will be radiated away as emission from $r_{\rm ph}$ with an approximately blackbody spectrum (the fireball in the Big Bang has similarly a blackbody spectrum, seen now as the cosmic microwave background emission from the last scattering surface). This is the first electromagnetic signal detectable from the fireball. This photosphere radius $r_{\rm ph}$ is usually above the coasting radius r_c, with a temperature $\mathcal{T} = \mathcal{T}_0(r_{\rm ph}/r_c)^{-2/3}$, but could be below r_c if the initial fireball is clean

enough (i.e. a large enough η), in which case $\mathcal{T} = \mathcal{T}_0$. The typical value of the photosphere is $r_{\rm ph} \sim (10^{12} - 10^{13})$ cm for $\eta \sim (100 - 1000)$. A full discussion about various regimes of this "baryonic" photosphere is presented in Ref. 73.

- Internal shock radius ($r = r_{is}$): For an intermittent central engine with typical variability timescale of t_v, the typical distance between adjacent mini-shells is usually also characterized as $d = ct_v$. Suppose that a rear shell moves faster than a leading shell, i.e. $\Gamma_{(2)} \gg \Gamma_{(1)} \sim \Gamma_0$, the (fixed frame) time at which the fast shell catches up with the slow shell is $\hat{t}_{is} = d/(v_{(2)} - v_{(1)}) \simeq 2\Gamma^2_{(1)} d/c$, and the distance is $r_{is} \simeq \hat{t}_{is} c \simeq 2\Gamma^2_{(1)} d \simeq 2\Gamma^2_0 ct_v \simeq (6 \times 10^{13}\ {\rm cm})\ (\Gamma_0/100)^2 (t_v/0.1\ {\rm s})$. At such distances, mini-shells collide with each other and typically form strong "internal" shocks.[64,99,100]
- Shell spreading radius ($r = r_s$): For a shell with initial width Δ_0, assuming the Lorentz factor scatter $\Delta\Gamma$ within the shell is of order Γ_0, the shell starts to spread at a radius $r_s \simeq \Gamma^2_0 \Delta_0$, based on a similar argument as used to derive the internal shock radius.[66] The spreading radius of the mini-shells is roughly the internal shock radius. The spreading radius of the whole GRB shell is however much longer, typically $r_s = (3 \times 10^{15}\ {\rm cm})\ (\Gamma_0/100)^2 (T/10\ {\rm s})$. Beyond the spreading radius r_s, the shell width starts to spread in the fixed frame, i.e. $\Delta = r/\Gamma^2$ for $r > r_s$.
- Deceleration radius ($r = r_{\rm dec}$): The fireball shell is eventually decelerated by the ambient medium (e.g. ISM). During the intial fireball-medium interaction, a reverse shock propagates into the fireball to stop it. Usually a deceleration radius ($r_{\rm dec}$) is defined as the radius where the reverse shock crosses the fireball shell. For a prompt fireball or a fireball whose duration is short enough, we have $r_{\rm dec} = r_\Gamma$, i.e. the radius where the ISM mass collected by the fireball is equal to $(1/\Gamma_0)$ of the fireball rest mass, i.e. $M_{\rm ISM} \simeq \Delta M/\Gamma_0$, where $\Delta M = \dot{M}T$ is the total baryon loading of the fireball. For constant density medium, this radius is $r_\Gamma = (3E_{\rm iso}/4\pi n m_p c^2 \Gamma^2_0)^{1/3} \simeq (2.6 \times 10^{16}\ {\rm cm})\ (E_{\rm iso,52}/n)^{1/3} (\Gamma_0/300)^{-2/3}$. As long as the shell spreading radius r_s is less than r_Γ (or $t_\Gamma = r_\Gamma/c\Gamma^2_0 \Rightarrow T$), the fireball decelerates at $r_{\rm dec} = r_\Gamma$. This occurs at an observer-frame time[58] $t_{\rm dec} \equiv t_\Gamma \simeq 5\ (E_{\rm iso,52}/n)^{1/3} (\Gamma_0/300)^{-8/3} (1+z)$ s. In literature this is usually termed the "thin shell" case.[70,71] Alternatively, if the shell is thick enough (e.g. for a long duration of fireball ejection so that $T > t_\Gamma$), the deceleration radius moves further out to $r_\Gamma < r_{\rm dec} < r_s$. This is the "thick shell" case.[70,71] As the fireball starts to decelerate, a strong external shock also forms and propagates into the medium. So the deceleration radius is essentially the initial external shock radius.

Since the fireball moves essentially at the speed of light, the relevant fixed-frame (or central-engine frame) times corresponding to the fireball reaching various radii are simply derived by dividing the relevant distance with c. For example, the time for internal shocks to happen is about several hours while the deceleration time is about 10 days. If we imagine a GRB occurring in our neighborhood, say 10 pc from us, whose relativistic jet beam is exactly perpendicular to our line of sight, we would be

able to follow the real time advance of the GRB emission and trace when the jet head reaches various radii. [For $\theta = \pi/2$ and $\mu = 0$, we have $d\hat{t} = dt$ in Eq. (1)]. However, when we see a GRB from cosmological distances, by definition, the jet is beamed towards us. The relativistic propagation effect Eq. (1) squeezes all the timescales to within seconds. For example, the time delay between the onset of the internal shock emission and the launch of the fireball (which may be due to the collapse event whose trigger time may be recorded by future gravitational wave detectors) is only $t_{is} \sim (r_{is}/c)/2\Gamma_0^2 \sim (18\ \mathrm{ms})\ r_{is,14}(\Gamma_0/300)^{-2}(1+z)$, while the delay for the external shock emission is only $t_{\rm dec} = t_\Gamma \simeq (r_\Gamma/c)/2\Gamma_0^2 \sim (5\ \mathrm{s})(E_{\rm iso,52}/n)^{1/3}(\Gamma_0/300)^{-8/3}(1+z)$ (for the thin shell case). In both expressions, the factor $(1+z)$ takes into account the cosmological time dilation effect.

3.3.1. *How common is the thick shell case?*

Although the thick shell case has been widely discussed,[70,71,118,119,122] it is worth questioning whether a thick shell description is indeed relevant in reality. In all the current discussions, it is conventionally assumed that the fireball shell width is defined by the duration of the GRB itself, i.e. $\Delta_0 = cT/(1+z)$ (where the factor $(1+z)$ is again to correct the cosmological time dilation). By comparing $r_s = \Gamma_0^2\Delta_0$ with r_Γ (or alternatively comparing T with t_Γ), one can define a critical Lorentz factor[118−120] $\Gamma_c = 125(E_{\rm iso,52}/n)^{1/8}(T/100\ \mathrm{s})^{-3/8}[(1+z)/2]^{3/8}$. For $\Gamma_0 > \Gamma_c$, which is not difficult to satisfy, the fireball is in the thick shell regime. Hydrodynamically, the reverse shock becomes relativistic in the thick shell regime, while in the thin shell case, the reverse shock keeps Newtonian until reaching mildly relativistic at the deceleration radius.

However, the above analysis is based on the assumption that the central engine has kept an essential constant luminosity through out T. In reality, the GRB energy injection is likely to be intermittent, sometimes with a broad gap between emission episodes. In such cases, it makes no more sense to take $\Delta_0 = cT$. Instead, one should separate the whole duration into several well-defined emission episodes, i.e. broad pulses, and treat these sub-shells to decelerate independently. One therefore has several consecutive thin shells, but rarely a thick shell. This situation has actually been indicated in GRB 990123, the only burst from which the rising lightcurve of the optical flash was caught.[30] In the standard afterglow model the optical flash is attributed to the reverse shock emission (see Subsec. 5.1). The optical flash peak time (which corresponds to the time when the reverse shock cross the shell, $t_\times = t_{\rm dec}$) is at 50 s after the burst trigger. The GRB duration, however, was 63 s, which exceeds $t_\times$. This already is incompatible with the thick shell regime, since in theory, $t_\times$ should not be less than T. Furthermore, the rising lightcurve is very steep, which is consistent with a thin shell model, but inconsistent with a thick shell model which predict a much shallower rising slope.[118] Inspecting the GRB lightcurve in detail, we find that the gamma-ray lightcurve of GRB 990123 is well represented as consisting of two components, a first more intense one which lasts

less than 50 s, and a second less intense one with even shorter duration. Thus, it is reasonable to regard 50 s as the deceleration time of the first shell in the thin shell regime. More generally, we suspect that in most cases when a GRB runs into an ISM, the thin shell case is very common. The thick shell description, may be more relevant in the case that a GRB runs into a pre-stellar wind or a constant dense medium. In such a case, the critical time for the thick shell case is much shorter due to the high density of the wind at the deceleration radius.[122,123]

3.4. *Relativistic shocks*

The collision between the relativistic fireball and the ISM leads to a relativistic forward shock and possibly a relativistic reverse shock as well. The internal shocks are usually at least mildly relativistic due to the large Lorentz factor contrasts between the colliding shells.

For a relativistic shock, when the upstream matter is cold, the shock jump condition gives[52,70]

$$\begin{aligned} n_2 &= (4\gamma_{21}+3)n_1 \simeq 4\gamma_{21}n_1\,, \\ e_2 &= (\gamma_{21}-1)n_2 m_p c^2 \simeq \gamma_{21} n_2 m_p c^2 \simeq 4\gamma_{21}^2 n_1 m_p c^2\,. \end{aligned} \tag{3}$$

The subscripts "1" and "2" denote the unshocked (up stream) and shocked (down stream) materials, respectively; n is number density, e is internal energy, both measured in the comoving frames of the fluids, and γ_{21} is the relative Lorentz factor between the fluids 2 and 1, while the Lorentz factor of the shock front itself is $\gamma_s = \sqrt{2}\gamma_{21}$.

In discussing the interaction between two fluids, two shocks are formed simultaneously at the instant of contact, which propagate into the two fluids, respectively. There are four regions for the two fluids separated by the forward shock, the contact discontinuity, and the reverse shock. For the GRB external shock case, these four regions are (1) unshocked ISM, (2) shocked ISM, (3) shocked shell, and (4) unshocked shell.[70] The shock jump condition [Eq. (3)] also applies at the reverse shock. A final condition $e_2 = e_3$ (since $p = e/3$ for relativistic fluids, and equal pressure is required at the contact discontinuity) and the fact that the fluids move at the same speed across the contact discontinuity finally close the problem, and a simple analytical description about the system is available.[70] Similar analyses can be applied to collisions between the mini-shells. Each collision is accompanied with a pair of internal shocks propagating into the two colliders.

There can also be more complicated cases, e.g. involving three (or more) shell interactions. This happens in the phase when a fireball shell is already decelerated by the ISM, while a trailing shell (which could be ejected by the central engine at a later time, or be ejected by the engine at essentially a same time as the leading shell but with a much lower Lorentz factor) catches up with the decelerated leading shell and collides with it. This is a typical problem in the study of GRB afterglow evolution when there is additional energy injection (see Subsec. 3.7.3), and the recent observed "step-like" optical afterglow lightcurve of GRB 030329[241] provides

a strengthened motivation to study such a theoretical problem. We have performed a detailed analysis of the three-shell-interaction hydrodynamics.[267] In this case, there are altogether six regions separated by three shocks and two contact discontinuities. By applying shock jump conditions to all three shocks (and noticing that the leading shell is hot), one can derive a self-consistent solution only when the relative Lorentz factor between the trailing and the leading shell exceeds a critical value defined by the energy ratio between the two shells. Otherwise, the injection is only mild, with the injected material connecting to the initial shell as a single long-lasting thick shell.

3.5. *Synchrotron emission*

The GRB prompt emission is clearly non-thermal, and so is the afterglow emission. The most natural mechanism for non-thermal emission is synchrotron emission, i.e. emission from relativistic electrons gyrating in random magnetic fields. The question of whether the GRB prompt emission (e.g. the γ-rays) is definitely due to synchrotron is still subject to debate (see Subsec. 4.3). However, the synchrotron shock model is widely accepted as the major radiation mechanism in the external shock, which is thought to be responsible for the observed broad-band afterglows.

There are three major assumptions that are adopted in almost all the current GRB afterglow models. Firstly, electrons are assumed to be "Fermi" accelerated at the relativistic shocks and to have a power-law distribution with a power-law index p upon acceleration, i.e. $N(E_e)dE_e \propto E_e^{-p}dE_e$. This is consistent with current shock acceleration numerical simulations.[147–149] Secondly, a fraction ξ_e (generally taken to be $\lesssim 1$) of the total electrons associated with the ISM baryons are accelerated, and the total electron energy is a fraction ϵ_e of the total internal energy in the shocked region. Thirdly, the strength of the magnetic fields in the shocked region is unknown, but its energy density ($B^2/8\pi$) is assumed to be a fraction ϵ_B of the internal energy. These "micro-physics" parameters, p, ϵ_e (ξ_e) and ϵ_B, reflect our inability of tackling the problem, whose values are usually fitted from the data.[130–134]

There are several critical energies in the power-law distribution of the electrons. For $p > 2$ (which is consistent with numerical simulations, and seems to be consistent with most of the observational data), a lower limit is set by the requirement that the average energy density in the shock-heated region is $\gamma_{21} n_2 m_p c^2$ Eq. (3), which reads $\gamma_m = g(p)(m_p/m_e)(\epsilon_e/\xi_e)\Gamma \sim 310[g(p)/(1/6)](\epsilon_e/\xi_e)\Gamma$, where $g(p) = (p-2)/(p-1)$ with $p = 2.2$ adopted, and we have simply redefined γ_{21} as Γ. According to the standard synchrotron emission theory,[268] the radiation power of an electron is $P_e = (4/3)\sigma_T c \gamma_e^2 (B^2/8\pi) \propto \gamma_e^2$, so that high energy electrons "cool" more rapidly. For a continuous injection, as is the case in an afterglow (i.e. the forward shock keeps plowing into the ISM), there is a break in the electron spectrum at $\gamma_e = \gamma_c$, above which the electron energy spectrum is steepened due to cooling, i.e. $N(E_e)dE_e \propto E_e^{-p-1}dE_e$. This energy is time-dependent, which is defined by equating the comoving dynamical timescale of the blastwave ($t' \sim \Gamma t$, Eq. (2)) to

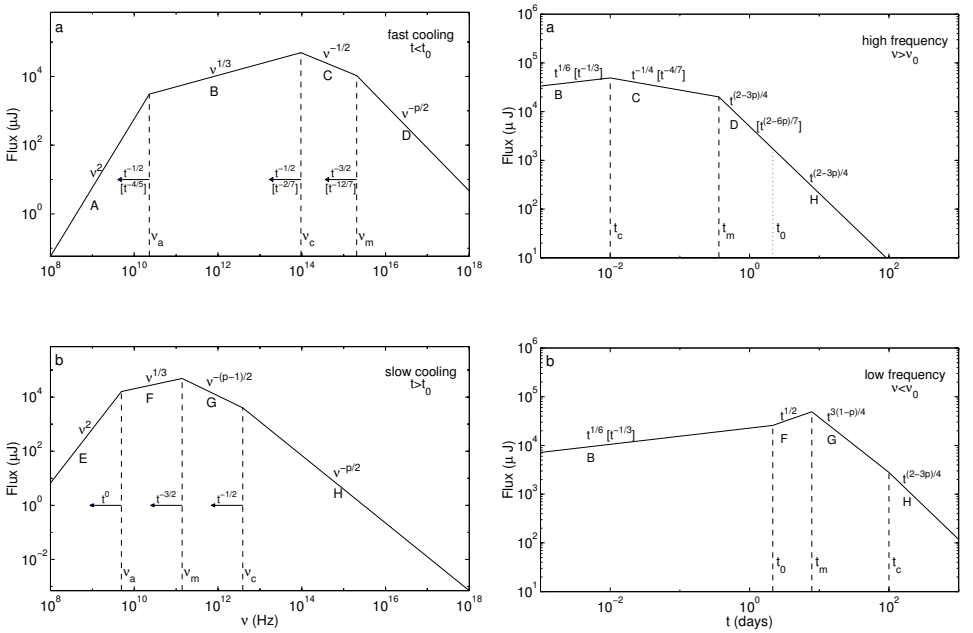

Fig. 11. The afterglow synchrotron spectra (left) and lightcurves (right) for the simplest fireball blastwave model (from Ref. 86).

the cooling timescale of the electron ($t'_c = \gamma_e m_e c^2/P_e$).[85,86] Finally, the maximum energy of the electrons (γ_M) is defined by equating the typical comoving acceleration timescale, $t'_{\rm acc} \sim 2\pi r_L/c$ (where r_L is the Larmor radius), with the shorter one of the dynamical timescale and the cooling scale. For electrons, the latter is relevant, which results in $\gamma_M \sim (3e/\sigma_T B)^{1/2} \sim 5\times 10^7 (B/1\ {\rm G})^{-1/2}$, where e is electron charge here.

The typical observed emission frequency from an electron with (comoving) energy $\gamma_e m_e c^2$ and with a bulk Lorentz factor Γ is $\nu = \Gamma\gamma_e^2 (eB/2\pi m_e c)$. Thus, three critical frequencies are defined by the three characteristic electron energies. These are ν_m (the injection frequency), ν_c (the cooling frequency), and ν_M (the maximum synchrotron frequency). In the afterglow problem, there is one more frequency, ν_a, which is defined by synchrotron self-absorption at lower frequencies. So the final GRB afterglow synchrotron spectrum is a four-segment broken power law[86,85] separated by the typical frequencies ν_a, ν_m, and ν_c. Depending on the order between ν_m and ν_c, there are two types of spectra[86] (Fig. 11). For $\nu_m < \nu_c$, which is called the "slow cooling case," the spectrum is

$$F = F_{\nu,m} \begin{cases} (\nu_a/\nu_m)^{1/3}(\nu/\nu_a)^2 & \nu < \nu_a\,, \\ (\nu/\nu_m)^{1/3} & \nu_a \le \nu < \nu_m\,, \\ (\nu/\nu_m)^{-(p-1)/2} & \nu_m \le \nu < \nu_c\,, \\ (\nu_c/\nu_m)^{-(p-1)/2}(\nu/\nu_c)^{-p/2} & \nu_c \le \nu \le \nu_M\,. \end{cases} \tag{4}$$

For $\nu_m > \nu_c$, which is called the "fast cooling case," the spectrum is

$$F = F_{\nu,m} \begin{cases} (\nu_a/\nu_c)^{1/3}(\nu/\nu_a)^2 & \nu < \nu_a\,, \\ (\nu/\nu_c)^{1/3} & \nu_a \le \nu < \nu_c\,, \\ (\nu/\nu_c)^{-1/2} & \nu_c \le \nu < \nu_m\,, \\ (\nu_m/\nu_c)^{-1/2}(\nu/\nu_m)^{-p/2} & \nu_m \le \nu \le \nu_M\,. \end{cases} \tag{5}$$

There are several more complicated regimes, involving self-absorption.[269,123] In the above expressions, the normalization factor is calculated by multiplying the total number of radiating electrons $4\pi r^3 n_1/3$ by the peak flux from a single electron,[86] which is only the function of B and is independent of the energy (γ_e) of the electron.[86,130]

One direct consequence of synchrotron emission is that the emission from an individual particle is polarized.[268] Theoretical predictions about the polarization degree of afterglow emission have been made.[270−272] Due to the random nature of the post-shock magnetic fields, the polarization is likely to be largely averaged out, and only a small degree of polarization is left. The most likely situation which can result in detectable linear polarization in an afterglow is thought to be around the jet-break time,[273,274] in which case a collimated jet and an off-axis line of sight conspire to produce an asymmetry which can lead to net polarization. In particular, for a large enough offset of the line-of-sight, a 90° change of polarization angle is predicted[273,274] around the jet break time for the simplest uniform conical jet model. Current afterglow (optical) polarization observations detect low-degree ($< 5\%$) of polarization emission,[242,243] but the data are too sparse to adequately test the concrete models.

3.6. *Simplest afterglow model*

The simplest afterglow model[78,79,81,82,85,86] is based on several assumptions, which minimize the complications. These assumptions include (a) isotropic fireball; (b) constant ambient density (ISM); (c) impulsive injection in the fireball; (d) relativistic fireball; (e) synchrotron emission of the electrons; and (f) constraints on the microphysics parameters (e.g. no evolution, $p > 2$, etc). Under such conditions, the fireball Lorentz factor evolves with radius r (or $\hat{t}$) and with observer's time t as

$$\begin{aligned} &\Gamma \propto r^{-3/2} \propto t^{-3/8}\,, \quad && r \propto t^{1/4}\,, \\ &\Gamma \propto r^{-3} \propto t^{-3/7}\,, && r \propto t^{1/7}\,. \end{aligned} \tag{6}$$

The first scaling is valid for an adiabatic evolution of the fireball in which the energy $E \propto nr^3\gamma^2$ is constant, which is generally valid at late epochs (later than hours) in all afterglows, and is also valid at earlier epochs for many afterglows. The second scaling is valid for the evolution of a "radiative" fireball, in which the total energy in the fireball decreases prominently due to radiation loss,[85,275,276,278] while the momentum $\propto nr^3\gamma$ conserves. This extreme radiative regime is only

valid for the fast-cooling regime when $\epsilon_e \sim 1$. This last condition appears unlikely, according to the current afterglow fits.[130–134] More generally, a blastwave may not be strictly adiabatic or radiative. A reasonable treatment is to adopt a quasi-adiabatic evolution with small radiative correction.[279]

With a certain dynamics (e.g. adiabatic evolution), one can quantify the time evolutions of the critical frequencies, ν_m, ν_c, ν_a, as well as the normalization $F_{\nu,m}$, parameterized in terms of the burst properties (i.e. $E_{\rm iso}$, n and luminosity distance d_L) and the micro-physics parameters (i.e. p, ξ_e, ϵ_e, and ϵ_B). Although the scalings on all these parameters are the same for different works, different treatments of normalization easily cause a factor of a few difference in the coefficients. Here we take the latest value of the coefficients,[9] in which a typical value $p = 2.2$ has been adopted. The dependence of unknown ξ_e is also incorporated.

$$\nu_m = (6 \times 10^{15}\ {\rm Hz})\ (1+z)^{1/2} E_{52}^{1/2} (\epsilon_e/\xi_e)^2 \epsilon_B^{1/2} (t/1\ {\rm day})^{-3/2}\,, \tag{7}$$

$$\nu_c = (9 \times 10^{12}\ {\rm Hz})\ (1+z)^{-1/2} \epsilon_B^{-3/2} n^{-1} E_{52}^{-1/2} (t/1\ {\rm day})^{-1/2}\,, \tag{8}$$

$$\nu_a = (2 \times 10^{9}\ {\rm Hz})\ (1+z)^{-1} (\epsilon_e/\xi_e)^{-1} \epsilon_B^{1/5} n^{3/5} E_{52}^{1/5}\,, \tag{9}$$

$$F_{\nu,m} = (20\ {\rm mJy})\ (1+z) \epsilon_B^{1/2} n^{1/2} E_{52} d_{L,28}^{-2}\,. \tag{10}$$

At a given time a comparison of ν_m and ν_c shows whether the flow is in the slow cooling or fast cooling regime, and using Eqs. (4) and (5), one can calculate afterglow lightcurves for a particular band (i.e. fixing a particular observation frequency). The different segments on the lightcurves have different temporal decay indices that correspond to different spectral regimes.[86] The commonly used temporal indices include the following. In X-rays, shortly after the trigger, the temporal index is essentially constant, i.e. $(2-3p)/4$, which is steeper than -1 for typical p values. This is consistent with the observations. In the optical band, the lightcurve first rises as $\propto t^{1/2}$, peaks around hours, and then decays with an index $-3(p-1)/4 \sim -1$. Although the rising lightcurve has not been firmly detected (but has been inferred)[119] due to the technical limitations, the falling part of the lightcurve is generally consistent with the observational data. The rising lightcurve may be buried by the contribution of the reverse shock emission at earlier times (as might have been the case of GRB 990123 and GRB 021211), but there is a large region of parameter space for which this part is not expected to be buried and should be observable.[120]

A convenient way to test various model regimes is to perform simultaneous measurements of both the temporal and the spectral indices. Writing $F_\nu(t,\nu) \propto t^\alpha \nu^\beta$, the relations of α and β in various regimes are listed in Table 1. Also listed in the table are the jet, wind, and $p < 2$ models, discussed below.

A more sophisticated model takes into account the radiation of the whole bulk of the blast wave, which is usually modeled as a Blandford–McKee[52] self-similar profile, as well as the emission contributions from the so-called equal-arrival-time surface, which is typically is egg-shaped or pear-shaped.[280–283] These effects result in

Table 1. Temporal index α and spectral index β in various afterglow models, the convention $F_\nu \propto t^\alpha \nu^\beta$ is adopted, from Refs. 85, 86, 110, 289 and 326. The assumption $\nu_a < \min(\nu_m, \nu_c)$ is made. (Under certain conditions, e.g. for the wind fast cooling case in some limited regime, the higher ν_a case is relevant,[289] so that the values collected here are no longer valid). The jet model applies for the sideways expanding phase, which is valid for both ISM and wind cases and is usually in the slow cooling regime.

	β	α ($p > 2$, $p \sim 2.3$)	$\alpha(\beta)$	α ($1 < p < 2$, $p \sim 1.5$)	$\alpha(\beta)$
ISM, slow cooling					
$\nu < \nu_a$	2	$\frac{1}{2}$		$\frac{17p-26}{16(p-1)} \sim -0.06$	
$\nu_a < \nu < \nu_m$	$\frac{1}{3}$	$\frac{1}{2}$	$\alpha = \frac{3\beta}{2}$	$\frac{p+2}{8(p-1)} \sim 0.9$	
$\nu_m < \nu < \nu_c$	$-\frac{p-1}{2}$	$\frac{3(1-p)}{4} \sim -1.0$	$\alpha = \frac{3\beta}{2}$	$-\frac{3(p+2)}{16} \sim -0.7$	$\alpha = \frac{3(2\beta-3)}{16}$
$\nu > \nu_c$	$-\frac{p}{2}$	$\frac{2-3p}{4} \sim -1.2$	$\alpha = \frac{3\beta+1}{2}$	$-\frac{3p+10}{16} \sim -0.9$	$\alpha = \frac{3\beta-5}{8}$
ISM, fast cooling					
$\nu < \nu_a$	2	1		1	
$\nu_a < \nu < \nu_c$	$\frac{1}{3}$	$\frac{1}{6}$	$\alpha = \frac{\beta}{2}$	$\frac{1}{6}$	$\alpha = \frac{\beta}{2}$
$\nu_c < \nu < \nu_m$	$-\frac{1}{2}$	$-\frac{1}{4}$	$\alpha = \frac{\beta}{2}$	$-\frac{1}{4}$	$\alpha = \frac{\beta}{2}$
$\nu > \nu_m$	$-\frac{p}{2}$	$\frac{2-3p}{4} \sim -1.2$	$\alpha = \frac{3\beta+1}{2}$	$-\frac{3p+10}{16} \sim -0.9$	$\alpha = \frac{3\beta-5}{8}$
Wind, slow cooling					
$\nu < \nu_a$	2	1		$\frac{13p-18}{8(p-1)} \sim 0.4$	
$\nu_a < \nu < \nu_m$	$\frac{1}{3}$	0	$\alpha = \frac{3\beta-1}{2}$	$\frac{5(2-p)}{12(p-1)} \sim 0.4$	
$\nu_m < \nu < \nu_c$	$-\frac{p-1}{2}$	$\frac{1-3p}{4} \sim -1.5$	$\alpha = \frac{3\beta-1}{2}$	$-\frac{p+8}{8} \sim -1.2$	$\alpha = \frac{2\beta-9}{8}$
$\nu > \nu_c$	$-\frac{p}{2}$	$\frac{2-3p}{4} \sim -1.2$	$\alpha = \frac{3\beta+1}{2}$	$-\frac{p+6}{8} \sim -0.9$	$\alpha = \frac{\beta-3}{4}$
Wind, fast cooling					
$\nu < \nu_a$	2	2		2	
$\nu_a < \nu < \nu_c$	$\frac{1}{3}$	$-\frac{2}{3}$	$\alpha = -\frac{\beta+1}{2}$	$-\frac{2}{3}$	$\alpha = -\frac{\beta+1}{2}$
$\nu_c < \nu < \nu_m$	$-\frac{1}{2}$	$-\frac{1}{4}$	$\alpha = -\frac{\beta+1}{2}$	$-\frac{1}{4}$	$\alpha = -\frac{\beta+1}{2}$
$\nu > \nu_m$	$-\frac{p}{2}$	$\frac{2-3p}{4} \sim -1.2$	$\alpha = \frac{3\beta+1}{2}$	$-\frac{p+6}{8} \sim -0.9$	$\alpha = \frac{\beta-3}{4}$
Jet, slow cooling					
$\nu < \nu_a$	2	0		$\frac{3(p-2)}{4(p-1)} \sim -0.8$	
$\nu_a < \nu < \nu_m$	$\frac{1}{3}$	$-\frac{1}{3}$	$\alpha = 2\beta - 1$	$\frac{8-5p}{6(p-1)} \sim 0.2$	
$\nu_m < \nu < \nu_c$	$-\frac{p-1}{2}$	$-p \sim -2.3$	$\alpha = 2\beta - 1$	$-\frac{p+6}{4} \sim -1.9$	$\alpha = \frac{2\beta-7}{4}$
$\nu > \nu_c$	$-\frac{p}{2}$	$-p \sim -2.3$	$\alpha = 2\beta$	$-\frac{p+6}{4} \sim -1.9$	$\alpha = \frac{\beta-3}{2}$

smooth transitions around the spectral and temporal breaks, which provide a better model fit to the afterglow data. A complete description is presented in Ref. 284.

Another important ingredient of the standard afterglow model is the emission component from the reverse external shock. During the early phase of the fireball

shell — ISM interaction, a reverse shock propagates into the fireball shell itself while the forward shock propagates into the ISM.[78,74] The reverse shock heats up the shell and accelerates electrons which emit synchrotron radiation as well. The shock is short-lived and ends when it crosses the shell. After this shock crossing time the electrons cool, leaving a rapidly decaying emission component. Due to the high particle density in the shell compared to that in the external medium (e.g. the ISM), the peak frequency of the reverse shock is in the optical/IR band at the crossing time $t_\times$. Thus the reverse shock emission usually leads to an optical flash.[67,78,114–116] The early optical afterglow data of GRB 990123, GRB 021004 and GRB 021211 are consistent with the reverse shock interpretation.[114,116,117,119,120,40,285] We discuss in more detail the reverse shock and early afterglow signatures in Subsec. 5.1.

3.7. *Additional features of realistic afterglow models*

The six assumptions made to define the simplest afterglow model are introduced mainly to simplify the problem. In reality, they are not necessarily satisfied, or not in all cases (although in many cases, they work surprisingly well). Violations of one or more of these assumptions lead to second-generation or modified fireball models. Such effects are in fact naturally expected, and they can describe the physical problem more realistically without the need for other ad-hoc assumptions. In this sense, they are also part of the standard fireball paradigm. The studies of the following effects comprise a majority of the recent GRB afterglow theoretical modeling effort. Because of space limitations, we will not go into the details of each of them, but rather refer the readers to the original papers that discuss these effects.

3.7.1. *Jets*

The "isotropic" assumption is one of the likeliest to be unrealistic, and investigations of the consequences of assuming that GRB's are collimated relativistic flows, i.e. jets,[108,109] have been fruitful. There are two quantities colloquially (and confusingly) referred to as the "beaming factor" in the GRB problem. One is the geometric beaming factor, which in the simplest jet model is just the opening angle of the jet, θ_j. The second is the relativistic beaming factor, i.e. the emission of a particle or an object that moves with a Lorentz factor Γ is beamed into a cone with opening angle of $1/\Gamma$. More accurately, it is useful to refer to the first one as the collimation factor (or angle), and to the second one as the relativistic beaming factor. Jets give rise to an interesting interplay between the two effects. Initially the jet is ultra-relativistic, with $1/\Gamma < \theta_j$. An observer on the beam only receives information from within the relativistic light cone and has no knowledge about whether outside this cone the emitter is radiating or not. The description of the dynamical evolution is therefore equivalent to the isotropic case. As the jet slows down, eventually the relativistic beam becomes wider than the geometric beam or collimation angle, i.e. $1/\Gamma > \theta_j$. Two effects come into play. First is the edge effect, i.e. the observer starts to feel a deficit of energy per solid angle (which is the

key parameter in afterglow theories, denoted as $E_{\rm iso}/4\pi$ in the above notations). Second, the causally connected region starts to extend to the whole jet cone around the same time, and can keep expanding sideways. This means the jet can start to expand sideways.[109] The times for the two effects to take effect are close to each other,[111] or may coincide,[110] depending on the assumption of the unknown expansion speed. In the asymptotic regime for times much longer than the jet break time, the dynamics is

$$\Gamma \propto \exp(-r/l) \propto t^{-1/2}\,, \qquad r \propto t^{0}\,, \tag{11}$$

(where $l = [E_j/(4\pi/3)nm_pc^2]^{1/3}$ is the so-called Sedov length at which the collected ISM rest mass energy is equal to the jet energy E_j itself), so that the temporal dependence of various critical frequencies as well as the temporal decay indices all differ from the isotropic case (Table 1). In the asymptotic regime the post-jet-break optical lightcurve should have $\propto t^{-p} \sim t^{-2}$, much steeper than the isotropic case ($\propto t^{-1}$). Therefore a jet break is the natural interpretation for the apparent steepening observed in many GRB optical afterglows.

Although the behavior in the asymptotic regimes are well known, the detailed behavior near the jet break involves various complex effects including the jet spreading hydrodynamics. Many efforts have been made to model the smoothness of the jet breaks,[111–113,286] but consensus is yet to be achieved even for the simplest uniform jet case. Hydrodynamical numerical simulations are needed, but so far such studies are still preliminary.[287,288] Furthermore, GRB jets may not be simply uniform, but may be structured.[135,136] This could bring more interesting effects in modeling (see Subsec. 4.4 for more discussions).

3.7.2. *Non-uniform external medium*: *winds and bumps*

The GRB ambient density may not be uniform. A well-discussed scenario involves an external medium produced by a stellar wind from the massive star progenitor, for example, a Wolf–Rayet star. The wind can be modeled to first approximation as ejecting material with a constant mass loss rate ($\dot{M}$) and velocity (v_w). Conservation of mass gives a $\rho = Ar^{-2}$ density profile,[90] with $A = \dot{M}/4\pi v_w = 5\times10^{11}A_*$ g cm^{-1}. Such a density profile changes the fireball dynamics into

$$\Gamma \propto r^{-1/2} \propto t^{-1/4}\,, \qquad r \propto t^{1/2}\,, \tag{12}$$

so that the lightcurves for the homogeneous fireball,[85,88,90,289] for the collimated jets,[290,291] as well as for the early afterglow involving reverse shock emission,[121,122] are all modified accordingly (Table 1). It is noticeable that most of the current GRB afterglow data are consistent with a constant density external medium,[131,132,36] although a handful of bursts could be well modeled by the wind model.[289,292,293]

An external density jump would be expected at the interface of the stellar wind and the ISM outside of it,[294] which causes a distinct afterglow signature.[158,295] The ISM itself may also have density fluctuations, which would also add imprints on the

afterglow lightcurves.[157,296–298] The consequence of a sudden drop in the density profile has also been discussed.[299]

3.7.3. *Post-injection and variable injection*

The fireball deceleration time depends on the ambient density, but the duration of the energy injection into the fireball is determined by the central engine behavior. These two timescales are independent of each other. Thus it is possible, or likely, that the central engine is still active while the afterglow starts. This gives rise to a post-injection of energy into the fireball, causing "refreshed shocks,"[87] so called because the kinetic energy of the late-arriving material (catching up with the decelerated initial material) revitalizes the external shock. The injection could be either in a continuous form[87,300–303] with modulations superimposed, or discrete, e.g. through disrupted, collisional events.[304,267] The injection energy, both early and late, could be either in the form of a Poynting-flux-dominated flow or kinetic-energy-dominated shells.[267] However, the main characteristic of the post injection signature is that the flux level is systematically increased after each injection event without resuming the pre-injection level,[267] a signature that has been seen in GRB 030329.[305]

Post-injection effect changes the energy per solid angle along the line of sight, and influences the fireball dynamics. If the fireball is not uniform in the angular sense but is patchy, the fireball deceleration will cause bright or dim spots to enter the relativistic beam, which equivalently changes the energy per solid angle of the fireball along the line of sight.[306] Both this model, and the varying density profile model, were proposed to interpret the recent observed optical lightcurve wiggles in GRB 021004.[297,298] X-ray data could disentangle the two scenarios, since above the cooling frequency ν_c emission is in the density-independent regime. For GRB 021004, variability was also found in X-rays,[39] which lends supports to the patchy beam model.

3.7.4. *Relativistic to Newtonian transition*

A decelerating relativistic fireball eventually becomes non-relativistic (Newtonian) at the Sedov radius $l = (3E/4\pi n m_p c^2)^{1/3} = (1.2 \times 10^{18}\ \mathrm{cm})\ (E_{52}/n)^{1/3}$, when the collected ISM rest mass energy is equal to the fireball energy. In the Newtonian phase, the fireball dynamics evolves as[79,89,307–310]

$$v \propto r^{-3/2} \propto t^{-3/5}\,, \qquad r \propto t^{2/5}\,. \tag{13}$$

The temporal decay indices are[89] $-(15p-21)/10$ for $\nu_m < \nu < \nu_c$, and $-(3p-4)/2$ for $\nu > \nu_c$, which for typical values of p is steeper than -1 as expected for the relativistic isotropic fireball, but is flatter than -2 as expected for the post-jet-break case. Therefore the evolution of a collimated jet involves a steepening around the jet break (typically days to weeks) and a later flattening as the fireball transmits into the Newtonian phase.[309] This later transition time is

$t_N \sim l/c = (450 \text{ days})(E_{52}/n)^{1/3}$, which is longer than a year for typical parameters. Usually this is of observational interest in the radio band.[308] In order to interpret some shallow lightcurve steepenings with relativistic-to-Newtonian transition effect (rather than the jet effect), a very dense medium ($n \sim 10^6$) is needed.[89] A self-consistent analytical description of the relativistic-Newtonian transition is proposed in Ref. 307.

3.7.5. *High energy spectral components*

Besides synchrotron emission, there are other mechanisms giving rise to high energy spectral components which must be operative at some level, and which may have interesting observational consequences. A straightforward and widely-discussed component is the synchrotron self-inverse Compton (IC) component.[67,74,75,311,312,81,313–319,186,320] The IC effect plays two roles in studying GRB afterglows. First, electrons cool both via synchrotron and IC, so that IC potentially influences the value and evolution of ν_c. Defining $Y = L_{IC}/L_{\rm syn}$, it is found[319] that $Y = [-1 + (1 + 4\eta\epsilon_e/\epsilon_B)^{1/2}]/2 \simeq (\eta\epsilon_e/\epsilon_B)^{1/2}$ for $Y \gg 1$, where η is the overall (synchrotron plus IC) radiation efficiency. The condition for IC cooling to be important is essentially $\eta\epsilon_e > \epsilon_B$. The second role of IC is that it forms a second spectral component, extending beyond the high end of the synchrotron spectrum. To first order, it can also be approximated as a four-segment broken power law, separated by three critical frequencies $\nu_a^{\rm IC} = \gamma_m^2\nu_a$, $\nu_m^{\rm IC} = \gamma_m^2\nu_m$, and $\nu_c^{\rm IC} = \gamma_c^2\nu_c$. In reality the spectrum is more rounded and the power-law breaks are not as sharp.[319] The high energy end of the spectrum is usually the Klein–Nishina limit.[186] The flux normalization can be derived through $F_{\nu,m}^{\rm IC}/F_{\nu,m} \sim (16/3)\sigma_T\xi_e nr$, which is $< 10^{-6}$ for typical parameters and most times of interest.[186] The condition for the IC component to stick out above the synchrotron component is defined by $F_\nu^{\rm IC}(\nu_c^{\rm IC}) > F_\nu(\nu_c^{\rm IC})$ for slow-cooling, and $F_\nu^{\rm IC}(\nu_m^{\rm IC}) > F_\nu(\nu_m^{\rm IC})$ for fast-cooling, both relating to the same physical condition.[186] Both this IC-emission-dominated condition and the above IC-cooling-dominated condition require $\epsilon_B \ll \epsilon_e$, although they are in principle different from each other. It is worth noticing that current afterglow modeling suggests that the shock condition $\epsilon_B \ll \epsilon_e$ is common among various bursts,[131–134] so that IC is potentially important in afterglow physics.

When observing in a particular fixed band, the IC component peak would sweep across the band at a later time than the synchrotron. This leads to a distinct bump signature in the lightcurve. The IC peak time is[186] $t_{\rm IC} = (3.4 \text{ days})(\epsilon_e/0.5)^{0.89}\epsilon_{B,-2}^{0.08}$ $\xi_e^{1.63}E_{52}^{-0.06}n^{-0.66}(1+z)^{0.32}\nu_{18}^{-0.68}$. The higher the energy band, the earlier the bump shows up, and a denser medium helps to ease the bump condition needed to detect the IC component. A GeV afterglow bump should be common for the currently favored shock parameters.[186] In the X-ray band, an IC bump is expected to emerge on the power-law decaying lightcurve around a couple of days if the medium density is moderately dense (say, $n > 5$),[317,319,186] and such a bump has already been detected in GRB 000926.[133]

Besides the IC component, several other high energy spectral components are also expected, which are related to hadronic processes. It is believed that the relativistic shocks also accelerate protons besides accelerating electrons. The protons also radiate via the synchrotron mechanism, and can also interact with photons (e.g. synchrotron photons from the electrons) to produce pions and muons.[321–323] The neutral pions decay into gamma-rays directly, and charged pions and muons also emit gamma-rays via synchrotron radiation.[322] However, these components may not contribute significantly to the afterglow emission for the currently-favored shock parameters, and the parameter space regime for these components to be important is small.[186]

3.7.6. *Microphysics*

Although an electron distribution index $p > 2$ is consistent with numerical simulations of shock acceleration,[147–149] and is generally consistent with the observational data,[131,132] in some bursts (e.g. GRB 010222) the temporal decaying slopes both before and after the "jet break" are too shallow, so that one seems to require $1 < p < 2$.[324,325] In such a regime, the maximum energy power is distributed towards the high energy end of the electron spectrum, bringing some new features for afterglow emission. A possible mechanism for producing such distributions is given in Ref. 387, and a study of this afterglow regime is presented in Ref. 326, the relevant conclusions being also collected in Table 1.

Essentially all current afterglow model fits assumed non-evolution of all the shock parameters, p, ϵ_e, ϵ_B. In principle, these may change, but this is hard to quantify, and the effect may be degenerate with other effects so that one might never be able to disentangle them. Detailed afterglow fitting seems to be compatible with the model that one or more such parameters evolve with time.[134,327] There might also be a gradient of magnetic fields behind the shock, but the data suggest that fields do not decay very rapidly.[328]

3.7.7. *Pair formation, neutrons, grains and other effects*

There are a number of other physical processes that would modify the simplest afterglow model.

First, the prompt gamma-ray form a radiation front which moves ahead of the blast wave and interacts with the ISM before the blastwave starts to decelerate.[329,330] The gamma-rays back-scattered by the medium can interact with outgoing gamma-rays and generate electron–positron pairs. The pairs enhance the opacity, and for a moderately dense external medium, the process leads to a run-away pair-loading process that modifies the blastwave dynamics considerably.[330–333]

Second, it is likely that there are free neutrons entrained in the fireball.[334–336] If neutrons and protons coast at essentially a same speed (which is not the case

under certain conditions),[171,337] the neutron shell would lead the proton shell when the latter is decelerated by the ISM. The comoving neutron decay time scale is $\tau'_n \sim 900$ s, so in the fixed frame, the distance where neutrons decay is $R_\beta = c\tau'_n\Gamma_n \simeq (0.8\times10^{16}\ \mathrm{cm})\ (\Gamma_n/300)$. This leads to a neutron-decay trail extending to a distance $R_{\rm trail} \sim 10R_\beta \sim 10^{17}$ cm. The trailing proton blastwave would interact with this neutron trail and form a possible bump on the lightcurve.[335]

Third, if GRB's originate in star forming regions, the existence of dust grains would cause a new emission component, in the form of a "dust echo."[338] Conversely, dust may be destroyed by the reverse-shock-induced UV-optical flash,[339] or by the X-ray afterglow.[340] The small-angle scattering of X-rays off the dust grains, on the other hand, may cause a soft X-ray bump on the afterglow lightcurve.[341]

Finally, as supported by accumulating evidence, at least some GRB's are associated with supernova.[28,29,45,46] Therefore, a supernova lightcurve, usually signaled by a reddened bump, is superposed on the optical afterglow lightcurve (as well as, in some bright cases, by a distinctive supernova spectrum).[45,46]

4. Problems

A cartoon of a GRB fireball is shown in Fig. 10. The theoretical model we have been discussing so far only includes those aspects that have been extensively tested, i.e. the afterglow part. This arises in the outest part of the fireball, from which the most of the information is retrieved. This picture is widely accepted in the GRB community, and is used as a standard theoretical framework to confront with the afterglow data. In this section, we zoom in towards the GRB central engine, and discuss several aspects of the model on which consensus has not been quite reached. The closer towards the central engine, the less we know about the physical processes going on. Nonetheless, theorists have utilized the limited information available to construct toy models to tackle the problems. For some of the "problems" discussed below there exists a leading model to interpret the phenomenon (e.g. the internal shock model for GRB prompt emission). However, there is (are) other competing model(s) that is (are) still widely discussed, and there are continuing debates on some of these issues.

A first major problem is the nature of the GRB prompt γ-ray emission itself. The origin of the gamma-rays has been debated since their discovery. Although the discovery of the afterglows settled the GRB distance issue, eliminated many GRB models, and achieved a consensus on attributing the afterglow radiation mainly to synchrotron radiation from a blast wave, there are still several variants within the current generic fireball scheme which are argued about for interpreting the limited GRB prompt emission data.[342] The major uncertainties include (a) the unknown fireball content (e.g. how important are magnetic fields in generating GRB's) and the unknown fireball energy dissipation mechanism (shocks or magnetic dissipation such as reconnection), (b) the unknown location where the GRB prompt emission is produced (i.e. internal or external), and (c) the radiation mechanism (e.g.

synchrotron radiation or other mechanisms such as Comptonization). These three aspects of the problem will be discussed below in Subsecs. 4.1–4.3. We then move on to discuss the global GRB jet structures (Subsec. 4.4), progenitors (Subsec. 4.5) and central engines (Subsec. 4.6). Finally we discuss GRB environments (Subsec. 4.7) and uncertainties in the shock physics (Subsec. 4.8). Some of these questions have also been discussed in Refs. 12 and 11.

4.1. *Fireball content: kinetic energy or magnetically dominated?*

4.1.1. *Internal shock model*

The leading model for the GRB prompt γ-ray emission is the internal shock model,[64] although there are various problems with it which are unresolved. The main motivation for this model is to explain short time (down to ms) time variabilities,[75] more efficiently than e.g. external shocks[97] (cf. Ref. 106). In this model, as discussed in Sec. 3, it is assumed that the fireball behavior is essentially hydrodynamical, i.e. baryon or kinetic energy dominated, and magnetic fields are neglected when considering the dynamics, and are only introduced when discussing synchrotron radiation through the parameterization of ϵ_B. As discussed in Subsec. 3.3, if the central engine ejects energy intermittently in the form of mini-shells, due to non-uniformity of the shell Lorentz factor, these mini-shells would collide at a typical distance of $r_{is} \sim 2\Gamma_0^2 ct_v \sim (6 \times 10^{13}$ cm$)\ (\Gamma_0/100)^2 (t_v/0.1$ s). The collisions produce a pair of shocks propagating into both shells, heating them and accelerating electrons (and protons). With some magnetic fields in the shell (either carried from the central engine or generated *in situ*, e.g. by turbulent dynamos or instabilities), these electrons emit synchrotron radiation in the gamma-ray band after Doppler boosting. This is taken to be responsible for the observed GRB emission. In order to have the gamma-rays from an optically thin medium (to avoid thermalization), only internal shocks whose radii are above the baryonic photosphere ($r_{\rm ph}$, see Subsec. 3.3) contribute to the observed emission. There is another relevant issue currently widely considered. The gamma-rays generated from some closer-in internal shocks may have a large opacity to $\gamma\gamma$ interactions leading to pair production. The pairs generated at small radii form a Thomson scattering screen for the gamma-rays, and the radius where this Thomson scattering optical depth drops to unity defines another "pair photosphere," $r_{\rm pair}$, which is usually above $r_{\rm ph}$, limiting the minimum variability timescale of the lightcurves.[103,343,73] The optical depth in this pair photosphere may be self-regulated to be around a moderate value of a few, so that variabilities shorter than the scale defined by $r_{\rm pair}$ are only smoothed rather than completely washed out.[73]

The internal shock model has been extensively studied by various groups.[99–103] Its most successful aspect is its ability to model the complex temporal profiles of GRB prompt emission lightcurves. Numerical simulations indicate that the observed temporal behavior essentially reflects the temporal behavior of the central engine.[99] A caveat is that the central engine needs to be "bare," or have a channel leading

out to a lower density environment where optically thin shock radiation can be observed. In the currently favored "collapsar" scenario (Subsec. 4.5) for long bursts, the central engine is located deep inside a collapsing star, whose envelope may act as an additional agent to regulate the variability of the relativistic flow.[93] In any case, as long as the relativistic jet emanates intermittently from the central engine, internal shocks are likely to develop (a different situation arises if the jet is strongly Poynting-dominated, as discussed next). The combination of the internal shock emission and the pair photosphere screen can reproduce the $-5/3$ slope and the 1 Hz break in the power density spectra (PDS),[101,102,343] which are deduced from the data.[197,198] An important characteristic of the internal shock model is that outflows with lower Lorentz factors have higher E_p values.[103,342] This is contrary to most other models, and against the simple intuitive expectation that lower Γ (dirty) shells have a smaller Lorentz boost so the emission should be softer. The reason for this behavior in internal shocks is mainly that the low Γ shells collide at closer distances relative to the central engine, where the magnetic fields are stronger. A major reason why this model is widely studied is that there is a clear operational theoretical framework (e.g. shock jump condition, equipartition parameters, synchrotron radiation, etc.) upon which model simulations could be performed (in contrast to the Poynting-dominated model discussed below).

There are several problems or caveats about the internal shock model. First, the emission energy has to be extracted from the *relative* kinetic energy between the colliding shells, so the radiation efficiency is typically small, e.g. around 1%–5%.[344,101,102] In order to achieve a high GRB efficiency, as required by the fact that the prompt gamma-ray energy is of the same order of the afterglow kinetic energy,[36,131] some novel suggestions have been made, including nonlinear dissipation[345] and quasi-elastic collisions.[346] In any case, a large relative Lorentz factor dispersion in the flow is required.[345,346,103] A recent study indicates that by taking into account neutron-decay in the fireball, the efficiency of the internal shock mechanism is further reduced.[347] Second, the BATSE bright burst spectral sample[178] suggests that, at least within the same burst, the peak photon energy E_p distribution is narrow. This is hard to achieve within the internal shock model unless one invokes a strong bimodal distribution of the Lorentz factors,[103,348] which lacks straightforward physical origins. Although none of the above criticisms has displaced the internal shock model from its position as the leading paradigm, these are important issues to consider. For this reason, other models are worth investigating as well, some of which are discussed below.

4.1.2. *Fireball with strong magnetic fields*

There are several motivations for considering strong magnetic fields in the GRB problem, as opposed to the weaker fields needed to produce the observed radiation. (1) Electromagnetic energy is "clean" in the baryon load sense, and can propagate in vacuum. Since a GRB fireball typically requires a very small baryon loading

to achieve the high Lorentz factor needed to solve the compactness problem, a Poynting flux potentially makes it easy to transport a large amount energy without carrying much baryons.[374,349] (2) Current GRB central engine models commonly invoke a rapidly rotating black hole circulated by a debris torus, or a millisecond neutron star. During collapse, magnetic flux conservation naturally gives a field of order $10^{12} - 10^{13}$ G, as is observed in radio pulsars. With rapid rotation near the breakup frequency, magnetic fields are likely to be magnified via an $\alpha - \Omega$ dynamo[140] to achieve $\sim 10^{15}$ G or higher. Such high fields are also argued to be possible for direct collapse of a magnetized white dwarf.[63] A strongly-magnetized rapidly-rotating central engine is a likely scenario. (3) Magnetic fields are a possible agent to tap energy from the two energy reservoirs in the engine, i.e. the gravitational energy of the torus, and the rotational energy of the black hole or the neutron star.[151,152,63,154,156] Another energy extraction mechanism, i.e. the neutrino pair production process ($\nu\bar{\nu} \rightarrow e^+e^-$),[60] is found to just barely power GRB's when a beaming correction is taken into account,[350,351] so that the magnetic power is at least helpful in meeting the GRB energetic needs.[352] (4) Magnetic fields are helpful in collimating jets.[353,354] Other arguments in favor of the magnetic mechanism include its ability to alleviate the inefficiency problem of the internal shock model,[355,356] and the possibility of achieving narrow E_p distributions.[342] There are also three observational facts that support a strongly magnetized central engine (although the flow is not necessarily completely Poynting flux dominated). (1) The strong gamma-ray polarization[42] (cf. Refs. 43 and 44) may indicate a strongly magnetized central engine, either in pure Poynting-flux-dominated form,[357] or in conventional hydrodynamical form but with a globally organized magnetic field configuration.[358,359] (2) Modeling the reverse shock emission of GRB 990123 indicates that the reverse shock region should anchor a stronger field (by a factor of 15 in strength) than in the forward shock region.[120,360] A similar conclusion may also apply to GRB 021211.[120,143] These may indicate that the fireball contains a large portion of "primordial" magnetic fields carried from the central engine, although the flow is *not* completely Poynting-flux dominated (otherwise there should be no strong reverse shock).[361] (3) The non-detection of a bright photospheric thermal component may need a significant fraction of energy being stored in the magnetic form, so that the photosphere temperature could be reduced (i.e. by a factor of $(1+\sigma)^{-1/4}$) to evade detection[362,342] (but see Ref. 363).

The evolution of a fireball with large magnetic content, and how such a fireball may give rise to prompt gamma-ray emission has been reviewed by various authors,[364–368] but so far there is not a standard framework, as is the case for non-magnetic fireballs. The magnetic fireball evolution is now an interplay among three components, an electromagnetic component, an internal energy component, and a kinetic (bulk) energy component. On the other hand, traditional fireballs only invoke the latter two, with the internal component dominating in the beginning, converted into the kinetic form via radiation-pressure-driven acceleration, and partially re-converted back to the internal form via shock dissipation. The total amount

of energy of these two components is essentially conserved except for the radiation losses. To simplify the magnetic case, one can regard the sum of these two components (internal and kinetic) as one single component, through a parameter σ which is the ratio between the energy densities in the electromagnetic component and in the internal+kinetic component.[342] At the beginning of the fireball evolution when the kinetic component is negligible, σ is just the ratio of the cold component (in the form of Poynting flux) and the hot component (in the form of photon-pair fireball generated via neutrino annihilation).[d] During the coasting regime, when radiation-driven acceleration is completed, the ratio σ is simply the Poynting-flux-to-kinetic-energy ratio, as has been widely discussed in pulsar wind nebula theories.[361] In the GRB problem that we are interested in, σ is likely to decrease with radius, since the rapid spin of the central engine greatly eases the conditions for magnetic reconnection and other instabilities to develop, so that the magnetic field energy is dissipated.[e] About a half of the dissipated magnetic energy is converted to internal energy, and the other half energy is used to accelerate the fireball through magnetic pressure gradient.[370,353,354,368] If the dissipation radius is smaller than the photosphere radius, the internal energy is also converted to the kinetic energy eventually. However, if the dissipation radius is above the photosphere radius, the internal energy would be radiated eventually via the non-thermal electrons accelerated during the reconnection event.[355,356,369] This is the main source to interpret the GRB prompt emission in this model.

The inclusion of electromagnetic fields greatly complicates the GRB problem. For not very high σ flows, one simplified treatment is magnetohydrodynamical (MHD). The adoption of such an approximation is justified when the electromagnetic fields are fully coupled to the fluid, i.e. no dissipation is allowed. As discussed above, a GRB magnetic fireball needs to be intrinsically dissipative in order to interpret bursts, thus casting the MHD approach in question. Nonetheless, the MHD method is still used to solve the first-order problem,[353,354] assuming that the radiation loss is only minor, and that σ is not very high. For a high-enough σ (e.g. $\sigma \gtrsim \sigma_{c2} \sim 200$), the MHD condition breaks down globally at a distance $r_{\rm MHD} = (1.8\times10^{18}\ {\rm cm})\ L_{52}^{1/2}[\sigma(1+\sigma)]^{-1/2}\ (t_v/1\ {\rm ms})\Gamma_{0,2}^{-1} \simeq (1.8\times10^{15}\ {\rm cm})\ L_{52}^{1/2}\sigma_3^{-1}(t_v/1\ {\rm ms})\Gamma_{0,2}^{-1}$, which is defined by the condition that the real plasma density is lower than what is required for corotation,[342,371] i.e. the Goldreich–Julian density.[372] Here σ_{c2} is a critical value (at the relevant radius, here it is $r_{\rm MHD}$) above which $r_{\rm MHD} < r_{\rm dec}$, i.e. the MHD condition globally breaks down before the fireball starts to decelerate.[342] Although the condition for this to happen is stringent,[355,342] when it happens there might be a global magnetic field dissipation region within

[d]We note that the coexistence of both components is natural when a cataclysmic collapse event forms a rapidly rotating, strongly-magnetized compact object.

[e]In pulsar wind nebula theories, on the other hand, it has been a long-standing problem to solve the so-called σ-problem, i.e. there is no obvious reason to reduce σ from a pulsar wind (slow pulsars compared with GRB central engine) from a very high value (10^4) to a low enough value (< 1) required in the reverse shocks as inferred from the observations.

the radius range $r_{\rm MHD} < r < r_{\rm dec}$, in which electrons are randomly accelerated by turbulent electromagnetic fields, giving rise to high energy radiation (although the typical energy may not coincide with the observed GRB emission).[373] In reality, such a high σ value may not persist out to such a large radius, especially when reconnection-induced dissipation presumably take place from the very beginning of the flow.[355,356,370] Thus the MHD description may be generally valid to reveal the first order physics.[353,354] However, the formalism is still rather complicated even in the MHD formalism.[353,354] When the fluid inertia is negligible, as is the case in the high-σ limit, an even simpler treatment, i.e. the force-free approximation, could be introduced.[366,367] Using this simple formalism, the evolution of a magnetic bubble is delineated, and is found (surprisingly) to be similar[367] to the self-similar evolution of the hydrodynamical blast waves.[52] An interesting finding is that a point explosion gives rise to structured angular distributions of energy and Lorentz factor,[367] i.e. $\epsilon(\theta) \propto (\sin\theta)^{-2}$ and $\Gamma \propto (\sin\theta)^{-1}$, which nicely matches one version of the structured jet model[135,136] (see more discussions in Subsec. 4.4).

An important characteristic of a high-σ flow is that strong shocks cannot develop.[361] This is because the shock frame pressure is dominated by the magnetic field pressure, and not much internal energy is available for radiation. The condition for strong shocks is $\sigma < 0.1$.[361] As a consequence, the internal shock mechanism has to be replaced by something else if the GRB outflow is strongly magnetized (e.g. $\sigma > 1$). Besides reconnection-induced magnetic dissipation,[355,356] other mechanisms include Comptonization of the photospheric emission via Alfven turbulence[374] and synchro-Compton radiation in a large amplitude electromagnetic wave as the magnetic fireball is decelerated by the ambient ISM.[375] In all these models, however, the fireball is assumed to be converted back to hydrodynamical in the deceleration phase in order to match with the achievements of the hydrodynamical afterglow theories.[364,367] Such a transition, although seemingly *ad hoc* at first sight, is reasonable since such low-σ flows have been observed in pulsar wind nebula, even though the initial pulsar wind is definitely a high-σ flow.

The development of magnetic-field-powered GRB models is still preliminary. The difficulties that hamper its progress lie in the intrinsic complication introduced by electromagnetic fields so that it lacks an operational scenario to work on. The situation has been changing recently.[355,356,370,353,354,367] One potential problem for the magnetic model is that an extremely clean fireball involves much fewer baryon-associated electrons. Giving the same amount of dissipation energy, the typical energy of emission tends to be much harder than the sub-MeV band,[342] unless secondary pairs are involved in the problem, which on the other hand tend to smear out sharp variabilities observed in GRB lightcurves.[342,73]

4.2. *GRB location: internal or external?*

A second uncertainty in understanding GRB prompt emission lies in its location within the fireball. The non-thermal spectra suggest that the location should be

above the photosphere radius $r_{\rm ph}$, below which the emission is opaque. Afterglow radiation limits the prompt emission radius to be smaller than the deceleration radius $r_{\rm dec}$. Generally there are three suggestions in the literature:[342] (1) the external models[59,104−107,375] that suggest GRB prompt emission occurs upon strong deceleration of the fireball at $r_{\rm dec}$; (2) the internal models[64,99,100,356,373] that suggest GRB occurring at a radius $r_{\rm ph} < r < r_{\rm dec}$; and (3) the (innermost) models[374,72,73,343] that suggest GRB occurring right above the photosphere ($r \gtrsim r_{\rm ph}$). The current consensus is that the third emission component at most contributes to the final emission only partially, while a debate between the first two components (external or internal) is still going on.

The main focus of the debate is the GRB variability. For those bursts containing chaotic, spiky pulses, an internal scenario seems more natural to interpret the data. The variability could be due to the intermittent behavior of the central engine,[99] including that induced from black hole accretion instability, the intrinsic instabilities involved in the jet propagation within the stellar envelope,[93] and the intrinsic chaotic behavior of magnetic reconnections and instabilities for the high-σ scenario.[355] However, whether an external model can be ruled out is still uncertain. For a burst running into a homogeneous ISM, if the central engine lifetime ($t_{\rm eng}$) is shorter than the deceleration time $t_{\rm dec}$ (thin shell), the observed duration of the burst should be simply[58] $t_{\rm ang} \sim r_{\rm dec}/\Gamma^2 c$, which is caused by the effect of angular spreading (Subsec. 3.2).[266] In such a case, an external shock model only produces one single smooth pulse. The caveat is that the external medium could be clumpy, and if the scale of the clumps d is smaller than $r_{\rm dec}/\Gamma$, variable lightcurves could be produced. An argument against this scenario is that the process is likely to be inefficient,[97] the efficiency being estimated as $\eta \sim t_v/T \ll 1$ to first-order. However, taking into account the angle-dependent flux effect,[106,376] the observed variability may still be reproduced with a higher efficiency. So solely from this fact, the two scenarios can not be differentiated. The attractive feature of the internal scenario is that the variability arises more naturally, while the external shock model has to introduce an additional assumption, i.e. a (very) clumpy medium. It would then be worth checking the further consequences of such a clumpy medium. Evidence for non-uniform medium has been suggested for GRB's with X-ray line features,[187,34] and for GRB 021004 which shows a wiggling afterglow lightcurve[296] and multi-component absorption features.[159] However, in most afterglows, a clumpy medium is not required, and limits may be posed on the clumpiness. It is worth modeling the prompt GRB emission as well as the afterglow emission within the external shock model in a unified manner, and to make use of any statistical correlations between the GRB variability and afterglow irregularity, so as to validate or falsify the external shock GRB model.

Other clues can also help to differentiate the two scenarios. (1) The correlation between the "waiting time" and the amplitude of the next pulse[201] favors an engine-dominated (internal shock) model. (2) In the external shock model, one expects pulse widths spreading with time,[377] which is not apparent from the data. However,

this objection may be circumvented by taking into account the noise.[376] (3) The absolute value of Γ enters the problem in the external shock model, while only the relative value of Γ enters the problem in the internal shock model. As a result, some correlations are expected in the external shock model but not necessarily in the internal shock model. For example, a brightness-spectral hardness correlation is expected in the external shock model,[104] which is consistent with the data. A brightness-duration anti-correlation is also expected, but this is not apparent from the data. A statistical study about various correlations among duration, flux and E_p in the external shock model[378] suggests that the model predictions are compatible with the data, but the results do not necessarily rule out the internal shock model. For the internal shock model, the burst duration is expected to be independent with flux and spectral behavior.

Eventually, close monitoring of gamma-ray and X-ray emission at the transition time of prompt emission and early afterglow would be able to clearly identify whether or not there is a distinct afterglow component to stick in at a later time. This would eventually settle down the internal-external debate. This is in principle doable in the *Swift* era. Preliminary evidence of two distinct components is available for some bursts.[379]

One point (which is sometimes overlooked) is that the internal and external shock scenarios are *not* mutually exclusive. Both are part of a generic cosmological fireball[64,75,277] model. It is likely that both components coexist in all fireballs, with different components dominating in different bursts. For example, in some parameter regimes, the internal shock radii are beyond the external shock (deceleration) radius, so in this case, no internal shock is expected. There is a small fraction of BATSE bursts that show single-peak FRED-like temporal profiles.[1] These events are most naturally interpreted as of external shock origin.[143]

In the magnetically-dominated fireballs, there are also an internal emission component[356,373] and an external emission component.[375,367] A highly variable lightcurve is possible for both cases.

4.3. *GRB emission mechanism: synchrotron, or others?*

The leading radiation mechanism introduced to interpret GRB prompt emission is synchrotron emission.[59,67,74,64,77] This is the most natural mechanism, which is found successful to interpret afterglows and many other astrophysical phenomena. Many of the observed GRB spectra are found to be consistent with this interpretation.[380–382] However, the synchrotron model encounters difficulties in explaining several observational facts. The most notable one is sometimes called the synchrotron "line-of-death,"[383] i.e. the low energy photon number spectral index should not exceed $-2/3$ (or the F_ν spectral index should not exceed $1/3$). A good portion of GRB's seem to violate this limit[383] in the BATSE database. Proposals to solve this apparent inconsistency include critical discussions of the sensitivity of BATSE to give accurate slopes at such low energies,[385] introducing small-angle

"jitter" radiation,[384] anisotropic electron pitch angle distribution and synchrotron self-absorption,[385,212] as well as photospheric emission.[72] Another argument raised against the synchrotron mechanism is the cooling problem.[386] The synchrotron cooling time scale is very short, so that the prompt emission should be in the "fast-cooling" regime with very low cooling frequency, and an F_ν spectral index of $-1/2$ (or photon number spectral index $-3/2$) should obtain, either for the low or the high energy spectral indices, but this is not apparent in the data. This criticism may be circumvented by taking into account more complicated acceleration and cooling processes.[385,212]

Although the synchrotron model continues to be the main paradigm, some other alternative radiation mechanisms have been proposed to interpret GRB prompt emission. The most natural extension is to involve Compton scattering. If the typical synchrotron emission frequency is well below the sub-MeV band,[f] then the GRB prompt emission could be due to synchrotron self-inverse-Compton (IC) emission.[318] Caveats about this suggestion include, (a) a strong IC component indicates a higher radiation energy density than the magnetic field energy density, so that a positive feedback results in increasingly stronger higher-order IC components before the Klein–Nishina cutoff, and consequently greatly increases the total energy demand of GRB's;[388] (b) a strong IC component requires $\epsilon_B/\epsilon_e \ll 1$,[317,319,186] while in internal shock scenarios, ϵ_B can not be too small given a strongly-magnetized central engine;[342] (c) an IC mechanism tends to further increase the E_p dispersion due to the high power dependence of E_p on Γ,[342] which within the internal shock synchrotron model is already problematic without introducing bimodal distributions of Γ's.[103,348]

Alternatively, models involving Comptonization of thermal or quasi-thermal particles have been proposed. A saturated Compton cooling model seems to be able to fit the observed GRB spectra well.[389,390] It is unclear how the model parameters needed for the fitting could be generated naturally in the fireball model, although attempts to transplant this mechanism to the internal shock scheme has been made.[391] Comptonization off a very dense photon bath as the GRB mechanism was also proposed,[392] with the soft photon field being provided by the trapped photons within the massive envelope funnel of the progenitor.[393] A proton-induced pair-photon synchrotron cascade model, in analogy to a nuclear pile reaction, was proposed recently,[394] which can reproduce a narrow E_p distribution around 0.5 MeV (electron rest mass). For the usual luminosity and bulk Lorentz factors, and when cosmological redshift correction is incorporated, this seems to be consistent with the data.[214]

The report of a measurment of strong linear polarization in gamma-ray prompt emission,[42] though subject to debate and confirmation,[43,44] potentially holds

[f]In fact, in order to raise the synchrotron typical energy to the BATSE band, some parameters need to be pushed to extremes. For example, ϵ_B needs to be close to unity, and ξ_e (the fraction of electrons that are shock accelerated from the shocks) needs to be sometimes less than unity,[387,100] so that the energy per electron is raised.

important clues for understanding the GRB radiation mechanism. Models producing high gamma-ray polarization include Poynting-flux dominated models,[357] models involving hydrodynamical shells with entrained globally structured magnetic fields,[358,359] and models involving off-beam observers for a narrow jet.[358,395] Models involving inverse Compton scattering with offset beaming angles can also give rise to large degrees of polarization.[396–399] Additional information is needed to differentiate these possibilities.

Within the standard synchrotron model, there is a straightforward way to analyze the E_p distribution within various fireball models.[342] In general, one can write

$$E_p \sim \Gamma\gamma_e^2(\hbar eB/m_e c)(1+z)^{-1} \sim (2\times 10^{-8}\ \mathrm{eV})\ (\Gamma B)\gamma_e^2(1+z)^{-1}\,, \qquad (14)$$

where Γ is the bulk Lorentz factor of the fireball, γ_e is the typical comoving Lorentz factor of the electrons, B is the comoving magnetic field, and z is the redshift. In principle, B could have two origins, one carried from the central engine, and another generated locally (e.g. in the shocks), although recent evidence suggests that the engine-origin magnetic field may be the dominant component.[42,120] Both field components, once formed, have the same dependence on distance from the central engine, i.e. the transverse part goes as $B \propto r^{-1}$, if both σ and ϵ_B do not change substantially with radius.[342] The r-dependence of γ_e essentially determines the E_p distributions for the models. For the internal shock model, since γ_e is determined by the relative Lorentz factor between the two colliding shells, it should not have a strong dependence on r, so that $E_p \propto (\Gamma B)(1+z)^{-1} \propto L^{1/2}r^{-1}(1+z)^{-1}$. The dependence on r reflects a strong dependence of Γ ($E_p \propto \Gamma^{-2}$),[342,400] which is the intrinsic reason why the internal shock model cannot reproduce a very narrow E_p distribution within a same burst (unless bimodal distribution of Γ is introduced).[103,348] For a magnetic-dominated acceleration model, on the other hand, γ_e should only depend on the local field B, generally as $\gamma_e \propto B^{-1/2}$ if Lamor acceleration is involved. This nicely cancels out all the apparent r-dependences, so that $E_p \propto \Gamma(1+z)^{-1}$. Only the dispersions of Γ and z contribute to the final dispersion of E_p. This is one of the strengths of the magnetic-dominant internal models.[342] For the external models in both the low- and high-σ regimes, the case is more complicated with the interplay of the ambient density. The E_p distribution is found to be wide if independent distributions of luminosity and other parameters are assumed.[342] A narrower E_p distribution is possible if some intrinsic parameter correlations are taken into account.[378]

Different dependences of E_p on various parameters for different models provide a framework on which the models may be tested with future data (Table 2).[342] The parameters on which E_p depends in various models are in principle measurable with current and future observational facilities. Two major unknown parameters include the redshift and the bulk Lorentz factor. The former is measurable in large numbers based on expected localizations by the next generation GRB mission, *Swift*,[401] scheduled to be launched in 2004. The latter could be also measured or at least constrained in the *Swift* era using the early afterglow data[120] (see more in

Table 2. The dependences of the GRB spectral break energy (E_p) on various parameters in the synchrotron emission model of various fireball variants.[342] Model parameters: L: initial total luminosity (not just the radiated luminosity) of the fireball; E: initial total energy of the fireball; Γ: bulk Lorentz factor at the radius of GRB radiation; t_v: typical variability time scale; n: ISM density; z: redshift.

Model	E_p-dependences
Internal shock model	$E_p \propto L^{1/2}\Gamma^{-2}t_v^{-1}(1+z)^{-1}$
Internal magnetic dissipation model	$E_p \propto \Gamma(1+z)^{-1}$
External shock model	$E_p \propto \Gamma^4 n^{1/2}(1+z)^{-1}$
External magnetic dissipation model	$E_p \propto \Gamma^{8/3}L^{1/2}E^{-1/3}n^{1/3}(1+z)^{-1}$
Pair photosphere model	$E_p \propto \Gamma(1+z)^{-1}$
Baryonic photosphere model (wind coasting)	$E_p \propto L^{-5/12}t_v^{1/6}\Gamma^{8/3}(1+z)^{-1}$
Baryonic photosphere model (shell coasting)	$E_p \propto L^{-1/12}t_v^{-1/6}\Gamma(1+z)^{-1}$
Baryonic photosphere model (shell acceleration)	$E_p \propto L^{1/4}t_v^{-1/2}(1+z)^{-1}$

Subsec. 5.1), or could be measured using the high energy cutoff of GRB spectra[402] when the next generation gamma-ray mission, GLAST,[403] is launched in 2006. So there are good prospects for pinning down the location and mechanism of the GRB prompt emission as well as the content of GRB fireball through statistical analyses of future data.

4.4. *GRB jet: uniform or quasi-universal?*

The geometrical configurations are an essential ingredient in characterizing and understanding astrophysical phenomena (e.g. pulsars, AGN's, etc.). Evidence suggests that GRB's are very likely collimated. Understanding the degree of collimation as well as the possible structure of the collimated flow would be very essential to understand the burst mechanisms and the true event rates.

The main evidence for GRB fireball collimation is provided by the achromatic steepening breaks in some GRB optical afterglow lightcurves. The simplest model, i.e. a conical jet with a uniform energy distribution within the cone and sharp energy depletion at the jet edge[108–113] (Fig. 12), and how this model interprets the lightcurve steepenings, have been introduced in Subsec. 3.7.1. In this model, the time for the break to occur roughly corresponds to a measure of the jet opening angle, if the observer's line-of-sight is in the jet cone and not too close to the jet edge. Data show that such inferred jet opening angles (θ_j) show a large dispersion among different bursts, and so do the corresponding isotropic gamma-ray emission energies ($E_{\gamma,\rm iso}$). However, both dispersions conspire in such a way that $E_{\gamma,\rm iso}\theta_j^2$ is essentially a constant.[36,256] Within the uniform jet model, the total energy released in a GRB event can be estimated as $E_\gamma = (E_{\gamma,\rm iso}/4\pi)2\pi(1-\cos\theta_j)\times 2 = (1-\cos\theta_j)E_{\gamma,\rm iso}$, where $2\pi(1-\cos\theta_j)$ is the solid angle of a conical jet with an opening angle θ_j, and the factor 2 takes into account the consideration that the jet is likely bipolar. At

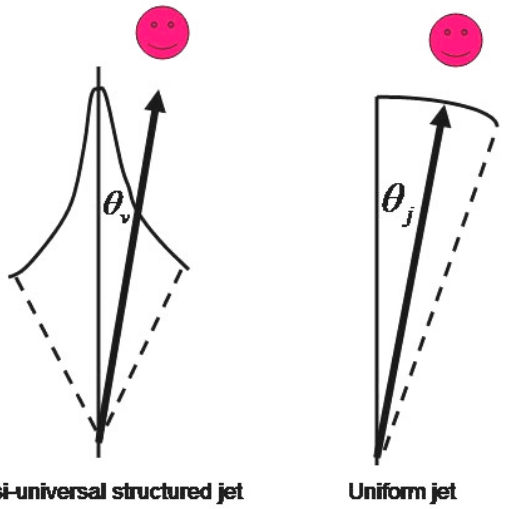

Fig. 12. Comparison between the quasi-universal structured jet model and the uniform jet model. The jet profile is for energy per solid angle (and Lorentz factor).

small angles, $f_b \equiv (1-\cos\theta_j) \simeq \sin^2\theta_j/2 \simeq \theta_j^2/2$. Therefore the above fact suggests a constant energy reservoir among long GRB's.[36,256] One puzzle posed by this is why a standard energy is collimated to rather different degrees among different bursts.

An alternative, and in principle more elegant, model is to postulate a quasi-universal jet structure, i.e. there is a non-uniform distribution of energy per solid angle within the jet, so that all burst outflows could more or less retain a similar geometry, the diversity of observed jet break times and isotropic-equivalent energies being caused by a different line-of-sight relative to the jet axis.[135−137] If such a model can be constructed, GRB's might be considered to have not only a standard energy reservoir, but also a standard geometric configuration. A lesson learned from AGN studies is that much of the apparent diversity in AGN's is simply caused by viewing-angle effects. Current unified AGN models try to find a paradigm interpreting various phenomena as being due to a standard configuration which is viewed at different angles.[404,405] A unified picture for the GRB phenomenology would be similarly appealing.[g] A straightforward speculation is to introduce a jet configuration such that $\epsilon(\theta) \propto \theta^{-2}$, as suggested by the $E_{\gamma,\rm iso}\theta_j^2 \sim$ const empirical law.

[g]Before the standard energy suggestion,[36] it had been speculated[406] that GRB's may have both a standard energy reservoir and a possible standard configuration based on a handful of afterglow data. The geometric configuration suggested in that work includes three distinct components, i.e. two uniform cones (one narrow and one wide) and one quasi-isotropic component. There was no discussion about whether such a configuration is consistent with the afterglow lightcurves, and it seems that such a structure is likely to cause some distinct lightcurve signatures that may violate most of the data. It is, however, likely to be consistent with the recent two-component jets as proposed to exist in GRB 030329.[48,49]

Here $\epsilon(\theta)$ is the energy per solid angle along the direction defined by θ (the angle between the viewing direction and the jet axis), which by definition, is equivalent to $E_{\gamma,\rm iso}/4\pi$. Such a configuration has been discussed earlier to model afterglow lightcurves,[85,407] but the viewing direction was placed close to the jet axis. By placing the viewing direction at arbitrary angles,[135,136] it is found that an achromatic lightcurve steepening is naturally reproduced. However, the time for the steepening to occur now corresponds to the epoch when the line-of-sight Lorentz factor $\Gamma(\theta_v)$ is decelerated to a value below $1/\theta_v$ (rather than below $1/\theta_j$ for the uniform jet model, see Fig. 12 for a comparison between both jet models).

The $\epsilon(\theta) \propto \theta^{-2}$ model (hereafter $k = 2$ power law structure model) is directly motivated by the $E_{\gamma,\rm iso}\theta_j^2 \sim$ const empirical law. From the point of view of reproducing jet steepening with the viewing-angle effect, more general types of jet structure, e.g. power law models with more general indices $\epsilon(\theta) \propto \theta^{-k}$, or even non-power-law structure, such as a Gaussian profile, may do the same job.[136] This is because well before the lightcurve "jet break" time, what is relevant for the dynamical evolution as viewed by the observer is only the average energy per solid angle within the $\sim 1/\Gamma(\theta_v)$ cone, which essentially remains the same as the relativistic beaming cone gets larger (given that $[\Gamma(\theta_v)]^{-1} \ll \theta_v$ is satisfied initially). After the jet axis enters the field of view, eventually all the initial jet structure is expected to be smeared out due to energy redistribution and sideways expansion. So in both asymptotic regimes, a structured jet has the same temporal evolutions as the uniform jet model.[136] Different jet structures only manifest themselves differently around the jet break time. The lightcurve of a $k = 2$ power-law structured jet has been modeled by various authors.[135,408–412] It is found that the sideways expansion effect is not prominent, so that a model with this effect neglected can still roughly reproduce the basic feature of the lightcurve.[408,409] For more general structure functions, the lightcurves are not adequately modeled due to the complicated physics involved. Modeling a Gaussian-type jet evolution[408] revealed a substantial energy structure redistribution during the evolution. The resultant lightcurves are consistent with the jet data.

A structured jet model has several clear predictions to confront with the observations. (1) Because different luminosities in this model are caused by different viewing angles, whose probability is well defined, the structured jet model has a specific prediction about the GRB luminosity function for bursts in the same redshift bin.[136,135] For the power-law model, one has $N(L)dL \propto L^{-1-2/k}dL$, while for the Gaussian model, one has $N(L)dL \propto L^{-1}dL$.[136] When the parameters are allowed to have some dispersion (quasi-universal picture), Monte Carlo simulations show that the above scalings still hold, with turnovers near the low- or high-luminosity ends caused by the variations of the parameters.[413] In contrast, the uniform jet model has no predictive power concerning the luminosity function. Currently, although the sample of bursts with spectroscopic redshifts is too small to allow a direct luminosity function study, the luminosity functions derived using various statistical methods[191,220,221,176,222] are not inconsistent with the model

predictions, especially for the quasi-universal Gaussian jet model with a roughly constant total energy.[136,413] (2) By taking into account the cosmological effects, the total number of bursts detected (regardless of redshifts) as a function of jet angle also has a clear prediction in the structured jet model.[414] For the $k = 2$ power-law model, a distribution peak is predicted around 0.12 rad, which is in rough agreement with the current data. The peak is expected to move to larger angles when the detection sensitivity threshold is increased.[414] Again the uniform jet model has no prediction power on this. In the *Swift* era, both of the above predictions may be tested.

Besides the preliminary support suggested by the above two tests, the structured jet model may also be indirectly supported by other theoretical and experimental results. (1) Progenitor and central engine studies naturally give rise to jet structures. Simulations within the collapsar model naturally predict a jet structure after the jet penetrates through the stellar envelope.[93] Studies of magnetic-threaded central engine models[415] as well as the evolution of a Poynting-flux dominated flow[367] result in jet structures in a natural way. (2) The interpretation of some GRB prompt emission empirical relations, such as the spectral lag-luminosity correlation,[210] are consistent with the structured jet hypothesis,[416,260,176] so that a coherent picture is achievable for both the prompt emission and the afterglow emission within this theoretical framework.

Several criticisms and caveats for the structured jet model have been raised. (1) The uniform jet model is very simple, and one can question the need to introduce jet structures unless it is necessary.[410] On the other hand, a non-uniform jet structure is a natural expectation in any jet model which arises from realistic physical processes. Even if the structured jet model so far does only the same job as the uniform jet model, it is still worth exploring, and should be explored, since it is more physical and has more predictive power. (2) The $k = 2$ power-law model predicts some anomalous signatures inconsistent with the data.[409] However, this may be caused by the assumption of an unphysical singular point in the jet structure, and is removable when more realistic jet structure (e.g. Gaussian) is considered.[408] (3) Recent optical polarization data[245] indicate a near 90° change of the polarization angle, which is inconsistent with the model prediction in the $k = 2$ power-law jet model.[417] However, the data accumulated are so far not conclusive.[418]

An important contraint on the jet structure is posed by the recent discovery that an $E_p \propto (E_{\rm iso})^{1/2}$ correlation earlier proposed for GRB's appears now to extend all the way down to energies characteristic of XRF's.[215,217,218] This relation, together with the $E_{\rm iso} \propto \theta_j^{-2}$ correlation,[36,256] immediately leads to the inference $E_p \propto \theta_j^{-1}$, which poses important constraints on both the simple $k = 2$ power law model and the on-beam uniform jet model.[137] This is because E_p varies by two orders of magnitude from GRB's to XRF's, so that even if XRF's correspond to events viewed at the equator (or isotropic events), GRB's would have to be events viewed within the 1° viewing angle (or corresponding to jets with opening angle smaller than 1°). For

the $k = 2$ model, this tends to greatly over-generate XRF's,[218,137] which are found to contribute $\sim 1/3$ of the total GRB/XRF population in the HETE-2 data.[218] For the uniform jet model, the very narrow GRB jets contradict or do not address some important afterglow jet break data.[137] All current GRB/XRF prompt emission and afterglow data are however consistent with a quasi-universal Gaussian-like structured model,[137] with an angular structure $\epsilon(\theta) = \epsilon_0 \exp(-\theta^2/2\theta_0^2)$. In this model, GRB/XRF jets still retain a characteristic angle θ_0, with a mild structure inside and a rapid exponential decay outside. The XRF's are only those events with viewing angles $\theta_v \sim (3-4)\theta_0$, which greatly decreases the number of XRF's. Statistically, the jet parameters (e.g. the typical angle, total energy in the jet, etc.) are allowed to have some scatter.[413] Monte Carlo simulations indicate that with reasonably small scatter, a quasi-universal jet model not only solves the GRB/XRF population problem, but can also reproduce the $E_{\rm iso} \propto \theta_j^{-2}$ correlation.[413,137] More complicated jet structures are in principle possible, at least for some events.[48,49,419] These models along with the simplest Gaussian model should be more extensively confronted with the various data, including luminosity function, redshift distribution, opening angle distribution, etc.[420,421]

To conclude, evidence in support of long GRB's having a quasi-universal structured jet configuration is mounting, but issues and contradictions remain. This paradigm will be more fully tested with the advent of the extensive data sets expected in the *Swift* era.

4.4.1. *Other jet models*

Besides the uniform conical and the structured jet models, there are some other GRB geometric suggestions discussed in the literature.

One suggestion is that GRB jets might be cylindrical,[422,423] as observed in some AGN's. Current modeling assumes that the line-of-sight is on the jet beam, and no apparent jet break is predicted from the model, so that the model can not interpret the whole GRB phenomenology, but may account for a sub-category of the bursts. A cylindrical jet is however consistent with the MHD description of a Poynting-flux-dominated flow.[353,354]

Another suggestion is that GRB's arise not from fluid fireball jets, but are rather "cannon balls" ejected from the central engine.[424,425] The picture is taken from analogy with some observations in micro-quasars and supernova remnants, although it is unclear how cannon balls could survive as compact entities against instabilities if they are accelerated over a large dynamical range of radii and Lorentz factors. The model has several distinct predictions (e.g. superluminal motion of the source, non-resolvable source image, and so on) that are different from those from the standard fireball jet models, which may be easily tested with the current and future data. The association of GRB's with supernovae has been quoted as a strong support for this model,[426,427] but this piece of evidence is not exclusive, since GRB-SN associations are widely expected within the standard model as well.[84,227−229,45,46]

Yet another suggestion has been that GRB's (and SGR bursts) are due to blazar-like jets that precess.[428−430]

4.4.2. *Orphan afterglows*

An important phenomenon associated with jets is the issue of "orphan afterglows," i.e. the implication of low energy (X-ray, optical, radio) decaying transients which are not detectable in γ-rays. The idea initially comes from the uniform jet model.[108] The starting point is that the initial GRB jet is sufficiently misaligned respect to the line-of-sight so that the observer misses the bright gamma-ray emission. Later as the jet decelerates, it expands sideways and enters the line-of-sight. The observer is expected to see a steeply rising afterglow lightcurve followed by a normal decaying lightcurve.[301,286] The event rates of orphan afterglows in various bands have been estimated,[431−433] but so far there is no firm detection of orphan afterglows of any kind.

The issue of orphan afterglows is complicated by two other considerations. First, there might be cosmological "dirty fireballs"[434] or "failed GRB's"[435] whose Lorentz factors are too small to allow transparent gamma-rays to be detected. These fireballs, when decelerated by ISM, can also give rise to orphan afterglows, whose signatures are not easy to differentiate from those of the jet orphan afterglows.[435] Some possible differences between these two types of transient events have recently been proposed.[436] Second, if GRB jets are structured,[135,136] the orphan afterglows of the first kind would be greatly reduced since there is no sharp jet edge cutoff. The second-type (dirty fireball) may exist due to higher baryon loadings at larger angles from the jet axis. So far there is no detailed discussion about the orphan afterglow rates for the structured jet model.

4.4.3. *Nature of X-ray flashes*

So far there have been essentially four types of interpretations for the XRF phenomenon (see Subsec. 2.1.1) in the literature,[342] i.e. dirty fireballs or failed GRB's,[434,37,435] high redshift GRB's,[37] fireballs dominated by photosphere emission,[73,370,343] and geometry-related events, such as offbeam GRB's in the uniform jet model[437−439] and large viewing angle events in the structured jet model.[137,440] The last model is likely in an anisotropic supernova explosion.[441] The first two interpretations attribute the soft-faint nature of XRF's to the dispersion of two parameters, i.e. Γ and z, respectively.[h] In reality, E_p is determined by the combination of many parameters (Subsec. 4.3), so that XRF's are likely to be caused by the dispersion of at least several of them, as indicated by Monte Carlo simulations.[342] The photospheric interpretation[73,370,343] may require the E_p distribution of the combined sample of GRB's/XRF's to have a separate component in

[h]Note that the dirty fireball interpretation is inconsistent with the internal shock model, which predicts a higher E_p for a lower Γ (Subsec. 4.3). It is however consistent with the external models and with the internal model involving a Poynting-dominated flow.[342]

the low energy regime. Recent afterglow observations for XRF's[442] pose important constraints on some of the above suggestions (e.g. high-z bursts and dirty fireball). In the *Swift* era, a large data set of both GRB's and XRF's, with photometric redshift measurements, should become available, which will eventually pin down the nature of XRF's. At present, the geometric model, especially the one involving structured jets, appears to us the most promising possibility.[137]

4.5. *Long burst progenitor*: *collapsar or supranova?*

Accumulating evidence suggests that long GRB's are associated with deaths of massive stars. The first cosmological progenitor scenario, i.e. the NS–NS merger model,[54,60,61] is now disfavored for long bursts,[91,444] although it is still a leading contender to account for the short, hard bursts (cf. Ref. 443). A variety of GRB progenitor models have been discussed in the literature,[150,352] and the requirement for accretion onto a central object occurring in a neutrino-dominated regime significantly constrains some of the scenarios.[155] Although the formation of GRB's from binary systems[445] or from gravitational energy loss in neutron stars[446] have also been discussed, the leading scenario for causing a long GRB involve the core collapse of a single massive star.[62,91,92,365,447,448,93,449,450] A supernova (SN) explosion is naturally expected to be associated with the GRB (although in its first incarnation[62] this was thought to be a "failed" supernova, i.e. a core collapse which did not to achieve ejection of its stellar envelope). The launching of a GRB jet is widely believed to involve a black hole–torus system resulting from this core collapse, hence the GRB is either produced simultaneously with the SN, if the star collapses to a black hole promptly, or the GRB is delayed with respect to the SN (which would give rise to a "supramassive", rapidly rotating neutron star as a first step, the neutron star later collapsing to a black hole in the second step, after loosing the angular momentum required to sustain the mass). Either a one-step or a two-step collapse are possible, depending on the mass and angular momentum of the progenitor core as well as the details of the collapse. If the delay of the second collapse is not very long, e.g. from minutes to hours as suggested by numerical simulations,[91] the two-step collapse scenario is not very different from the one-step collapse, and both of these cases can be referred to as "collapsars,"[8] e.g. of type I and II. On the other hand, an alternative suggestion is that the delay is long, e.g. from days to weeks or even months, and in this case the burst environment is very different from the collapsar case, due to the role of the well-separated supernova shell and a central pulsar wind.[96,451] This latter case is referred to as a "supranova." Figure 13 is a cartoon picture for the geometric configurations in both models.

The collapsar model has many merits in reproducing the data. First and foremost, it naturally involves a GRB-SN association, and predicts that GRB's are associated with star forming regions. The existence of the stellar envelope helps to collimate a jet with angular structure and can help to regulate the jet flow intermittently.[93] The duration of a burst is set by the fall-back timescale rather than

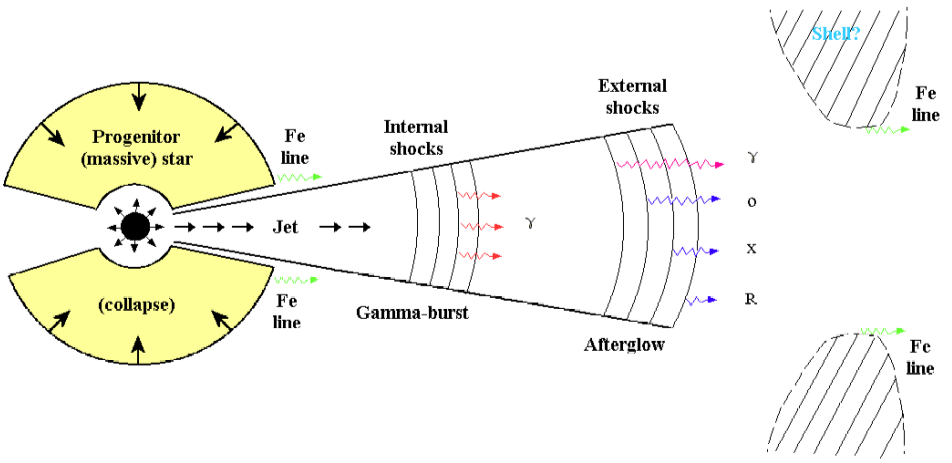

Fig. 13. Comparison of the geometric configurations of the collapsar and supranova models (from Ref. 5).

the accretion timescale,[91] which is natural to interpret the durations of long GRB's. The requirements for the progenitor include a narrow range of specific angular momentum (3×10^{16} cm^2 s$^{-1} < j < 2 \times 10^{17}$ cm^2 s^{-1}) and poor metallicity. Such a constraint may lead to a rough standard GRB energy and an event rate roughly consistent with the true GRB event rate after collimation corrections.[91,8] All of these are consistent with the main-stream observational and theoretical progress in the GRB field. The interaction of such a jet with the stellar envelope should also make unique signatures in both electromagnetic and non-electromagnetic forms. For example, prompt X-ray and gamma-ray signals are expected when the jet breaks through the envelope, leading to precursors to the main burst,[94,452,95] or even leading to short, hard bursts.[93] The breakout of the jet cocoon is also a candidate to interpret the X-ray emission line features hours after the burst trigger as have been seen in several bursts. Internal shocks within the jet before penetrating through the envelope would lead to high energy protons to produce prompt TeV neutrino signals.[172] A collapsar may also harbor a long-lived central engine that continuously pumps energy into the GRB fireball.[125,8] Recently, more realistic MHD simulations for the collapsar model have commenced,[450] which reveals the possible magnetic nature of the fireball. The difficulties of the collapsar model include the following. First, the required high angular momenta of the progenitors are difficult to achieve.[8] Second, although the top candidates for collapsars are those stars that have undergone intense mass loss before collapse (e.g. Wolf–Rayet stars), the expected wind environment for GRB afterglows[90,289] is not commonly identified in current afterglow studies.[36,131]

The supranova model was initially introduced to alleviate the baryon-loading problem, since weeks to months after the SN explosion, the environment is relatively clean.[96] Two other incentives were added later. One was its promise in interpreting

X-ray line emission (and absorption) features. The conjectured supernova shell (or some torus-like or funnel-like remnant) located at $10^{15} - 10^{17}$ cm from the central engine (which was ejected by the progenitor days or months before, moving with a speed of $\sim 0.1c$) provides a large mass of heavy elements for the inferred photo-ionization (by the continuum emission of the GRB or the early afterglow) and recombination processes to produce a strong Fe line,[124,453,127] which are thought to be responsible for many of the discrete X-ray spectral features detected in several bursts.[187,34,35] (More recently, lower-Z elements such as Ca, N, S, etc. have also been reported in bursts.)[35,454] Such an interpretation of X-ray lines has some distinct energetical advantages,[128] but it requires high clumpiness in the shell. This type of model is not the only one able to explain the lines. A competitive class of models attributes the spectral features to continuous photoionization by a long-lived central engine jet,[125] or by a jet cocoon in the collapsar scenario, interacting with the outer layers of the stellar progenitor.[94] These models make the line at smaller distances ($r \sim 10^{12} - 10^{13}$ cm, hence they are referred to as "nearby" models) and in a much higher density medium, require a smaller iron mass. Detailed modeling indicates that they are able to reproduce the data.[129] Other X-ray line models proposed include irradiation of a pair screen caused by the early afterglow back-scattering[455] and the Cerenkov line emission mechanism.[456]

A second advantage of the supranova suggestion is related to the role of the central pulsar after the SN explosion; this can form a pulsar wind bubble which can modify the ambient medium from a pre-stellar wind to a medium with constant (lower) number density, consistent with the observations. High equipartition parameters (ϵ_e, ϵ_B) as inferred from afterglow modeling[131,132] are also naturally interpreted.[458,451]

The supranova model suffers several criticisms. First, it takes fine-tuning to make a GRB months to years after the formation of the neutron star.[459,8] Second, lacking an envelope, the collapse may not produce a long burst[460] and the collimation mechanism is not clear. Third, detailed radiative transfer calculations[129] indicate that a supranova model would reproduce the reported large equivalent widths only for very shallow incidence angles of the ionizing continuum, which is naturally expected in nearby models (involving ionizing the walls of a stellar funnel), but is less natural for an ionizing continuum incident on a distant supernova shell, where normal incidence is expected. Finally, the most severe objection to the supranova model is the fact that the association between GRB 030329 and SN 2003dh allows at most a very short delay (less than two days) between the SN and the GRB events, and is compatible with both events being simultaneous.[46] This rules out the supranova model at least for this burst, and supports the collapsar model. Unfortunately, no high spectral-resolution X-ray observation were obtained in a timely manner, thus missing an excellent chance to test whether a pre-ejected SN shell is a requisite to generate X-ray lines.

Although the GRB 030329/SN 2003dh association greatly changes the balance between the collapsar/supranova debate, the issue of whether there is still a

sub-category of bursts originating from supranovae (e.g. perhaps those which have strong X-ray lines) is still unsettled. Future observations can help to distinguish the two scenarios. First, both scenarios predict different electromagnetic signals. The collapsar scenario predicts X-ray and gamma-ray precursors,[452,95] while the interaction between the fireball and the SN shell in the supranova model tends to produce enhanced, delayed GeV emission signals;[461–463] Second, the supranova model predicts more emission components and stronger flux levels of high energy neutrino emission due to the existence of the SN shell which provides more target photons and protons for $p\gamma$ and pp interaction.[464–467] This leads to stronger neutrino signals both for individual sources and for the diffuse background. Detection/non-detection of such an excess emission may be used to prove/disprove the supranova model. Finally, suggestions to directly search for SNs that occurred weeks to years before GRB's have been proposed,[468] which may pose direct constraints on the models.

4.6. *Central engine: what is behind?*

The most widely discussed GRB central engine invokes a central black hole and a surrounding torus. There are three ultimate energy sources, the gravitational binding energy of the torus, the spin energy of the black hole and the magnetic energy. Two ways of extracting the accretion energy and black hole spin energy have been considered, i.e. the neutrino mechanism,[60,350,351,91,469] and the Blandford–Znajek[151] mechanism. The former mechanism typically powers the conventional "hot" fireball, while the latter invokes strong magnetic fields threading the black hole[152,154,156,470] which typically would power a (at least initially) "cold" Poynting-flux dominated flow. The identification of the content of the fireball (Subsec. 4.1) and the mechanism of GRB prompt emission (Subsec. 4.3) would shed light on the mechanism that powers the central engine. The study of the extracting energy from the GRB central engine may potentially shed light into the black hole physics.[471–473]

Another type of GRB central engine often discussed in the literature is a millisecond magnetar.[63,371,374,474–477,446,365,478] These are rapidly rotating neutron stars with surface magnetic fields of order $\sim 10^{15}$ G and higher. These objects would be able to satisfy the essential conditions for a GRB central engine, e.g., energy, duration, variability, baryon loading, birth rate, etc, especially when considering the possibility that initially there could be some temporary toroidal fields with much stronger strength (10^{17} G).[476,446,478] The ultimate energy sources include the spin energy of the pulsar and the magnetic energy. For a pulsar engine, there could be in principle two energy components, an initial prompt component (either via neutrino mechanism or via destroying the temporary toroidal field) powering the prompt GRB, and another long-term component due to spindown of the millisecond pulsar. The latter component continuously injects energy into the fireball through Poynting flux,[300,303,267] or in the form of electron-positron pairs analogous

to the pulsar wind bubbles,[457] leaving well-defined bump signatures in the afterglow lightcurves.[300,303,267,457] It also naturally provides a long-lived central engine, which may be the agent to continuously photo-ionize the Fe in the stellar envelope, giving rise to the observed X-ray emission features.[125] In the two-step supranova models, a pulsar is invoked in the first step, which powers a pulsar wind to drive a magnetic-enriched bubble.[458]

A more exotic central engine mechanism involves a phase transition from normal neutron matter to strange quark matter.[479] The process is likely in a detonative mode, leaving behind a star completely composed of strange quark matter, called a strange star.[480] The possibility of neutron star–strange star phase transitions powering GRB's has been discussed within the context of one-step[481,482,483,484] or two-step processes.[485,486] A special category of such models lead to intermittent energy injection due to unstable photon decay, which is arguably a viable GRB central engine.[487] A common caveat about the strange star mechanism is that there is no evidence yet about the existence of strange quark matter, while the existence of black holes and pulsars have been widely tested.

4.7. *Environment*: *What is in front?*

On the galactic scale, GRB host galaxies are broadly similar to the normal, star-forming faint field galaxies at comparable redshifts and magnitudes.[13] The distribution of GRB-host offsets from the galactic center[444] as well as heavy element abundance studies[13] are fully consistent with a progenitor population associated with sites of massive star formation.

On the parsec scale, issues about the GRB immediate environment include whether the medium is (quasi-)uniform, wind-like or clumpy, and how dense the medium is on average. Insights about these questions can be obtained through afterglow lightcurve modeling[131−134] and time-dependent absorption feature modeling,[488] but the situation is still controversial. As mentioned earlier, the constant density medium model is consistent with most afterglow data,[131−134] but the $\rho \propto r^{-2}$ wind model also works in some bursts.[289,292,293] It is puzzling how different GRB's could have quite different immediate environments if they come from the same type of progenitor. There are so far no studies on correlations of the inferred GRB environment with other properties of the gamma-ray prompt and afterglow emission. The inferred medium density also varies significantly among different modelers, even for the same burst,[36,131,89,133,235] although there seems to be a trend towards favoring a universal moderate-dense medium with $n \sim 10$ cm^{-3} as being consistent with (most of) the data.[489] Suggestions that GRB's are embedded in dense molecular clouds have been made,[490] which is consistent with the high hydrogen column density (N_H) inferred from some X-ray afterglow studies.[491,488] A high-density afterglow model, on the other hand, although inferred from some afterglow fits,[89,492] may not be compatible with broad-band data[131,133] or with the inferred standard energy budget for all GRB's.[36] The dust grains expected in molecular

clouds may reflect and irradiate the afterglow emission and form a bright dust echo,[338] or a distinct IR signature detectable for nearby bursts.[493] Dust extinction is also invoked as the cause of at least some optically dark bursts.[253] So far direct evidence of dusts is not yet collected, and the possibility of dust destruction[339,340] further complicates the issue. In some bursts (e.g. GRB 021004), multiple spectral absorption systems and density bumps are identified, which refer to a rather non-uniform, bumpy environment from close to the burst to much farther away from the burst.[159,494] It has been suggested that time-dependent absorption features both in optical[495] and in X-rays[488] are powerful tools to study GRB environments, but the present data are too sparse to allow firm conclusions to be drawn.

4.8. *Shock parameters: Universal or unpredictable?*

It is almost certain that afterglows are produced by collisionless relativistic shocks energized by the GRB's. The physics involved in these shocks is however poorly known. A widely discussed scenario is that particles are accelerated via repeatedly crossing a shock front, and achieve a power-law energy distribution through the well-known Fermi mechanism.[496,497] The latest numerical simulations indicate that the resultant power-law index is universal,[147] i.e. $p \simeq (2.2 - 2.3)$. This is in sharp contrast with what is inferred from afterglow fit data. With the simplest jet model, broadband modeling indicates that p is quite unpredictable, ranging from $\sim$1.4 to $\sim$2.8 among different bursts.[131,132] An important caveat for the modeling is that the present model only uses the simplest assumptions, e.g. uniform jets, non-evolving equipartition parameters, etc. The electron power-law index p is usually derived from the temporal decay index after the lightcurve break or steepening attributed to jet properties (i.e. $F_\nu \propto t^{-p}$). Whether the inconsistency between broad-band modeling and shock acceleration simulations is caused by the simplified GRB model or by the simplified shock acceleration model is not known, and developments in both directions are needed. For example, proposals for generating a flat electron index have been outlined,[387] while more complicated afterglow models (e.g. post injection, or refreshed shocks) could change the inferred p value considerably.[498]

Unlike the magnetic fields invoked in the magnetically dominated scenarios for GRB emission, which may be carried out from the central engine, the origin of the magnetic fields in the internal or external shocks needs to be addressed. The inferred fields are typically of strength $\epsilon_B \sim 10^{-3} - 1$. Loosely, these have been attributed to a turbulent dynamo mechanism behind the shocks. More specific suggestions include two-stream instability of relativistic plasma[271] or the presence of an intrinsically magnetized ambient medium.[458]

5. Prospects

We have summarized the achievements (Secs. 2 and 3) and uncertainties (Sec. 4) in our current understanding of the nature of GRB's. Guided by past experience, we believe that our knowledge of GRB's will be further advanced in the coming

years, driven by future observational breakthroughs. In this section, we attempt to foresee some of the possible milestones in the upcoming new epoch of GRB study led by NASA's two future missions, *Swift* and *GLAST*, as well as by experimental developments in new channels of GRB study. We will discuss how our knowledge of GRB's is likely to be extended in the temporal domain (i.e. early afterglows and the prompt emission — afterglow bridge, Subsec. 5.1), in the spectral domain (i.e. GeV–TeV emission, Subsec. 5.3), and into other, non-electromagnetic regimes, including high energy neutrinos (Subsec. 5.5) and gravitational waves (Subsec. 5.6). We will also discuss how future developments could unveil the mystery of short GRB's (Subsec. 5.2) and how GRB studies may become a unique tool to investigate the early universe (Subsec. 5.7). Some of the topics covered in this section are also discussed in Refs. 14 and 15.

5.1. *Early afterglow and reverse shock emission*

The discovery of GRB afterglows[23–25] has greatly extended our knowledge of GRB's, both in the temporal domain (hours to even years after the burst) and in the spectral domain (below the gamma-ray band, from X-rays down to radio). However, due to the long time scale of the mission alerts up to the ground, the slowness in slewing instruments and/or bad weather in ground-based instruments, in most cases afterglow observations have started hours after the burst trigger. At this phase, the afterglow blastwave has been decelerated and entered a self-similar regime, and the behavior is essentially determined by the total energy per solid angle in the fireball and the properties of the ambient medium. Some precious information characterizing the fireball, such as the initial Lorentz factor of the fireball and the magnetic content of the fireball, has at this stage been lost. In order to retrieve such information, very early afterglows need to be studied, which contain information on the emission from the reverse shock region during and shortly after the reverse shock crosses the fireball shell at the very beginning of the shell-medium interaction. As of October 2003, only four bursts (GRB 990123, GRB 021004, GRB 021211 and GRB 030418) were caught within less than 10 minutes after the triggers,[226] thanks greatly to the growing number of robotic optical telescopes spreading all over the world. The situation will change soon following the launch of *Swift* in June 2004, which will automatically record essentially all the triggered GRB early afterglows starting from 100 seconds after the triggers with an on-board X-ray telescope (XRT) and a UV-optical telescope (UVOT). The mission will also issue prompt alerts to the ground to allow rapid follow-up observations from the ground-based telescopes.

There are two important differences between the reverse shock emission and the forward shock emission. First, the reverse shock emission is prompt and short-lived. The electrons are continuously accelerated only until the reverse shock crosses the initial shell. This happens at the deceleration radius $r_{\rm dec}$, which is r_Γ for the thin shell case, or at $r_\times$ for the thick shell case (see Subsec. 3.3). The observer's time at which this occurs is usually defined by the crossing time, $t_\times = \max(t_\Gamma, T)$,

where $t_\Gamma = [(3E/4\pi\Gamma_0^2 n m_p c^2)^{1/3}/2\Gamma_0^2 c](1+z)$, and Γ_0 is the initial Lorentz factor. The fireball Lorentz factor at the shock crossing time is $\Gamma_\times = \min(\Gamma_0, \Gamma_c)$, where $\Gamma_c \simeq 125 E_{52}^{1/8} n^{-1/8} T_2^{-3/8} [(1+z)/2]^{3/8}$. The reverse shock emission is therefore generally divided into two segments. For $t < t_\times$, the synchrotron spectrum is (as in the forward shocks) characterized by a four-segment broken power law, as discussed in Subsec. 3.5. However, for $t > t_\times$, since no new electrons are accelerated, the emission above the cooling frequency ($\nu > \nu_c$) totally disappears. Second, before and at the crossing time, both the pressure and the internal energy density across the contact discontinuity are the same, but the particle density in the shocked shell is much larger than that in the shocked ISM. Given similar electron equipartition and injection factors (ϵ_e and ξ_e), the typical energy per electron is much lower in the reverse shock region than in the forward shock region. So the synchrotron peak frequency in the reverse shock is much lower than that of the forward shock, mainly in the optical/IR regime,[67,78,114–116] and the reverse shock emission component is typically characterized by an optical flash and a radio flare.

Compared with the forward shock emission, there are more physical cases to be considered in the reverse shock lightcurves, depending on issues such as whether one is in the thick shell regime (in which case the reverse shock becomes relativistic before $t_\times$) or in the thin shell regime (in which case the reverse shock only becomes mildly relativistic at $t_\times$); what is the relative position of the observational band with respect to the break frequencies both before and after the crossing time; fast or slow cooling; whether the medium is an ISM or wind-like, etc. For the ISM case, a full discussion about various lightcurve cases (12 altogether) is presented in Ref. 118. A closer investigation taking reasonable parameters reveals that there are only 4 most relevant cases.[120] This depends on whether the shell is thin or thick and another parameter $\mathcal{R}_\nu \equiv \nu_R/\nu_{m,r}(t_\times)$, which defines whether the observational band is above (for $\mathcal{R}_\nu > 1$) or below (for $\mathcal{R}_\nu < 1$) the injection synchrotron frequency of the reverse shock emission at the shock crossing time. A common feature of all four cases is that the final temporal decay slope is steep, i.e. $F_\nu \propto t^{-\alpha}$ with $\alpha \sim 2$. This segment of the lightcurve has been identified in GRB 990123[30] and GRB 021211,[41,40] lending credence to the reverse shock scenario. The reverse shock lightcurve component eventually joins the forward shock component at a later time. A generic expected early optical afterglow lightcurve involves two peaks[120] (Fig. 14), i.e. a reverse shock peak with $(t_{r,p}, F_{\nu,r,p})$ at the beginning of the $\alpha \sim 2$ lightcurve segment (which usually corresponds to $t_\times$), and a forward shock peak with $(t_{f,p}, F_{\nu,f,p})$, which is the transition point from the $\propto t^{1/2}$ lightcurve segment to $\propto t^{-1}$ lightcurve segment corresponding to the typical forward shock synchrotron frequency crossing the band.

The information provided by the reverse shock emission has been widely used to estimate the initial Lorentz factor of the fireball.[114,117,499,500,360,119,285] Usually these methods rely on the poorly known shock parameters (ϵ_e, ϵ_B and p) as constrained by the forward shock modeling, and they implicitly assume that same parameters also apply in the reverse shock. The absolute values of these parameters

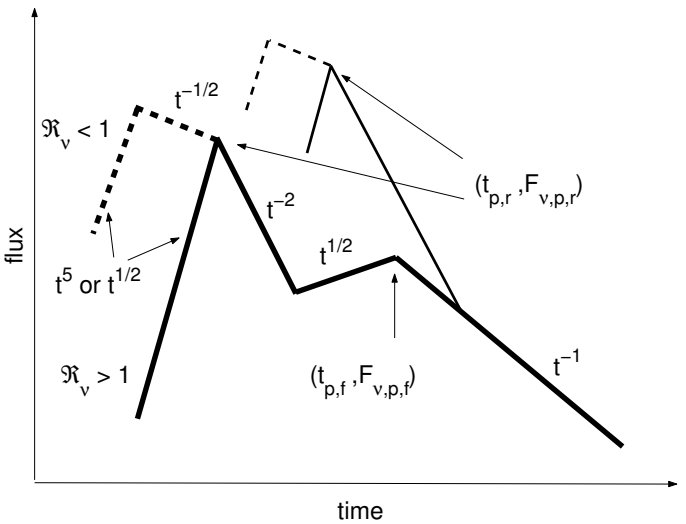

Fig. 14. Typical GRB optical early afterglow lightcurve that includes the contributions from both the forward and the reverse shock emission components (from Ref. 120).

are usually used to derive Γ_0. We have recently proposed a straightforward recipe to derive Γ_0 by using a combined reverse and forward shock analysis.[120] The only information needed is the time and flux level at both the reverse shock and the forward shock emission peaks. Since there are simple correlations of ν_m, ν_c and $F_{\nu,m}$ between both shocks at the crossing time,[119,120] the *ratios* of the two peak fluxes and the two peak times only depend on Γ_0 (or a parameter depending on both Γ_0 and Γ_c for the thick shell case), the unknown parameter $\mathcal{R}_\nu$, as well as the *ratios* of the shock microphysics parameters. Suppose that both lightcurve peaks are detected in an idealized observational campaign; the two unknown parameters Γ_0 and $\mathcal{R}_\nu$ can be then solved for from the peak flux ratio and the peak time ratio, if the microphysics parameters are the same in both shocks (as implicitly assumed in almost all previous studies). Otherwise, the derived Γ_0 (and $\mathcal{R}_\nu$) can be still derived as a function of the *ratios* of the microphysics parameters. No knowledge about the absolute values of ϵ_e and ϵ_B is needed. If the GRB central engine is strongly magnetized, it is natural to expect that the magnetic field in the reverse shock region could be stronger than that in the forward shock region. Generally, one can introduce a free parameter $\mathcal{R}_B \equiv B_r/B_f$ into the problem, which may be solved for along with Γ_0 if enough information is available. Case studies suggest that the central engines of GRB 990123 and possibly of GRB 021211 are strongly magnetized[120] (see also Ref. 143). This provides an independent clue in addition to the gamma-ray polarization[42] (cf. Refs. 43 and 44) to suggest a magnetized central engine and fireball, although not necessarily Poynting-flux dominated.

For a wind-like environment, the early afterglow lightcurves are considerably different.[121,122] Due to the high density of the shell at the shock crossing time,

the reverse shock emission is in the fast cooling regime. After shock crossing, the emission above ν_c disappears. For typical parameters, $\nu_{c,r}(t_\times)$ is well below the optical band, so that the lightcurve after the crossing time is characterized by a steep decay ($\propto t^{-2-\beta}$, β is the spectral index, which is either 1/2 or $p/2$ for the case we are interested in) due to the off-axis angular time delay effect.[122] This rapidly-decaying reverse shock emission joins the forward shock lightcurve later, at a time earlier than the one found in the ISM case.[120] There is also a simple correlation between the fluxes and times for the shock crossing and the crossing of the forward-shock typical frequency across the band,[122] which is related to Γ_0 and $\mathcal{R}_B$ as well, so that a similar recipe to derive Γ_0 and $\mathcal{R}_B$ as in the ISM case[120] can be utilized. When the self-absorption frequency is above injection frequency in the fast cooling regime (which is likely in some parameter regimes), balance between cooling and self-absorption heating implies a bump in the electron distribution spectrum, and hence, in the emission spectrum.[123] This gives rise to interesting spectral and temporal signatures for the early afterglows that can be used to diagnose fireball and wind parameters. Such a signature is also relevant for the case of a constant dense medium.[123]

The significance of Γ_0 and the forward to reverse field ratio $\mathcal{R}_B$ in understanding the GRB fireball content and prompt emission mechanism has been discussed in Subsecs. 4.1 and 4.3. In the *Swift* era, an abundance of early afterglow data is expected, and it is desirable to systematically analyze these data to retrieve essential fireball parameters such as Γ_0 and $\mathcal{R}_B$. A follow-up statistical analysis of these data along with other measurable parameters would eventually lead to constraints on, or even the identification of, the GRB prompt emission site and mechanism.[342]

The early afterglow data may also reveal distinct emission features from processes such as pair loading[333] and neutron decay.[335] More detailed studies about these emission signatures (spectra, lightcurves) are needed in order to differentiate them from other signatures such as post-injection[267] and the passage of the forward shock synchrotron injection frequency across the observational band.[119]

5.2. *Short GRB's and other possible sub-categories*

In a discussion of prospects for future GRB studies, one major unsolved puzzle concerns the nature of the short, hard bursts, which comprise of about 1/4 of the total GRB population detected. Essentially all the information about (and from) afterglows discussed above is for the long bursts. The short bursts still remain as mysterious as the long bursts were before 1997. So far only upper limits on the afterglow emission of a few short bursts are available, although the BATSE archival data contain marginal evidence of weak X-ray afterglows.[501] Directly scaling the long burst afterglow models to the short bursts, the calculations are straightforward with most parameters unchanged except for a smaller total energy (due to the short duration if a roughly similar luminosity is assumed), or a smaller energy per solid angle, if the jets are broader (as might be expected e.g. for NS–NS mergers where a

large envelope is absent), and possibly a lower external medium density (as might be expected for the NS–NS merger scenario if these wander out of the galaxy). The results indicate that the afterglows for short bursts are faint and consistent with the current upper limits, and that the afterglows are most easily detected in the X-ray band.[502]

On the theoretical side, simple estimates reveal that NS–NS mergers are likely to result in central engines with short durations.[61,350,153] Extensive numerical modeling has been carried out to reveal how compact merger events can produce GRB's.[503−505] Significant collimation is required for the neutrino-driven model in order to account for the detected isotropic-equivalent luminosity of short bursts (assuming cosmological distances), while a magnetically-driven model requires fields as high as 10^{17} G to account for the short duration (otherwise, mergers may also produce long bursts).[505] There is a suggestion that short bursts may also originate from collapsars.[93] Other ideas to account for the long and short bimodal distribution include whether or not strong GW losses are incurred in the spin down of a millisecond magnetar model,[63] and whether, in a collapsar scenario, the black hole–torus involves hyper-accretion or suspended accretion because of the interplay between the disk and the black hole spin.[443] All these models, however, have to be developed further to circumvent the apparent lack of the afterglows for the short bursts.

Swift, thanks to its rapid slewing ability, is expected to be able to catch faint afterglows of some short bursts, if they exist, and consequently to lead to the identification of their locations, redshifts as well as other properties such as jet beaming. Merger events have been expected to occur in regions with a large offset from the host galaxy center (due to the asymmetric kicks during the formation of the NSs;[506] see, however, Ref. 507), the position information alone relative to the host is already very essential for the short bursts' identity and/or merger rate. Furthermore, ambient density estimates through broadband afterglow modeling will also help to reveal whether short bursts are located in low-density regions as expected at least for some fraction of short bursts in the merger models.

Will new sub-categories of GRB's be identified in the *Swift* era? One can only speculate on this. It may be unlikely to identify new categories based on one piece of information alone (e.g. burst duration). However, cross correlations among multiple parameters may well lead to the identification of new categories, such as XRF's,[37] long-lag bursts associated with the super-galactic plane,[176] and the sub-energetic rapid-decaying bursts.[256] Alternatively, and equally exciting, it may be that new observational evidence and theoretical modeling might reveal that there is a unified paradigm behind all these apparently diverse sub-classes.

5.3. *High energy photon emission*

A new and so far barely explored window in the electromagnetic spectrum is the high energy extension in the GeV–TeV photon energy range, which holds significant

promise for a better understanding of GRB. After the better studied MeV prompt emission and the broad band low energy (X-ray down to radio) afterglow studies, the final (photon) frontier is the very high and ultra-high energy domain, extending at least up to TeV, and possibly beyond. In the intermediate range of 10–100 MeV observations, while not numerous, have been made with relatively high significance, e.g. with SMM,[179] OSSE,[508] COMPTEL,[509] etc. In the GeV–TeV energy range, only modest to low significance detections of a handful of GRB's have been reported, e.g. with the EGRET spark chamber[21] or TASC calorimeter,[47] or from ground-based air Cerenkov telescopes.[185,510] These results have wetted the appetite for obtaining higher quality data, and have led to a vigorous campaign of instrumental developments in high energy regimes. Several planned new missions including GLAST,[403] AGILE,[511] and several planned or already operating ground-based air and water Cerenkov telescopes (e.g. Milagro, VERITAS, HESS, MAGIC and CANGAROO-II, etc.),[512] are expected to open a new era of unveiling the mystery of GRB high energy emission.

On the theoretical side, high energy photons in this energy range are expected from the leptonic component of the fireball (e.g. electron IC emission from various emission sites) as well as from the hadronic component of the fireball (e.g. proton synchrotron, π^+ synchrotron emission and π^0 decay from $p\gamma$ or pn, pp interactions in various emission sites (see Subsec. 3.7.5 for a discussion in the external shock case). Within the standard fireball shock scenario sketched in Fig. 2, we could in principle have the following components emitting in the GeV range and possibly above.

- Electron self-IC component from the external forward shock.[75,315,186]
- Electron self-IC components from the external reverse shock or the cross-IC components between the reverse and the forward shocks.[74,513]
- Electron self-IC component from the internal shocks.[311,313,514]
- Proton synchrotron emission in the external shock.[321−323,186]
- Photo-meson cascade emission from the external shocks.[322,515,467]
- Proton synchrotron emission and photo-meson cascade emission from the internal shocks.[516]
- Cascade emission resulting from pn inelastic collisions during the early phase of fireball evolution when the neutron component decouples from the proton component.[334,171]
- Baryonic photosphere component (and possibly its Comptonization component) extending into the GeV regime (for a low-σ high-entropy fireball).[72,374,73,342,362,363]

The above is not an exhaustive list, and not all of the above processes may be operative at any time or in all models. For example, if the GRB wind is strongly dominated by a Poynting flux so that the GRB prompt emission is not due to internal shocks, the internal-shock-related GeV components in the above list would be suppressed or absent. Within the collapsar scenario, the precursor emission

tends to be in low energies. But in the supranova scenario, the presence of the presumed supernova shell and hot pulsar wind nebula provide copious photon or baryonic targets for additional IC and hadronic interaction components in the GeV range.[461−463] Such emission components may also exist for the type-II collapsars such that the delay between the SN and GRB is short.

An ideal theoretical approach would start with a study of all the above emission components within a unified theoretical GRB framework, to compare the relative importance of each component as well as the parameter regimes for each component to dominate. An example of the potential benefits of such an approach is provided by the recent observation of a distinct GeV component identified recently from an archival BATSE burst, GRB 941017,[47] which evolves separately from the usual low-energy (sub-MeV) component. So far, only the high energy emission components in the external forward shock have been studied in a consistent manner.[322,186] In the external shock scenario, the proton synchrotron and hadronic cascade emission components are only important when ϵ_B is close to equipartition, while ϵ_e is very small, i.e. of order of m_e/m_p or below. This is especially so when the proton power-law index is taken as 2.2–2.3 rather than the special value 2 adopted in most of the previous studies.[321−323] For the more commonly invoked parameters (e.g. $\epsilon_e \sim 0.1$, $\epsilon_B \sim 0.01$) as inferred from afterglow studies,[131,132] the hadronic components are completely buried beneath the electron IC component, which forms a distinct bump in the GeV regime. The MeV–GeV lightcurve predicted within this parameter regime reveals a second broad bump lasting hours or even a day,[186,315] and the flux level is detectable by GLAST for typical bursts at $z \sim 1$ (Fig. 15).[186] According to this scenario, the previously detected long-lived GeV emission for GRB 940217[21] is simply a nearby burst whose GeV afterglow is bright enough to be caught by EGRET,[75,186] although alternative interpretations have been also proposed.[517,518,519] GLAST would reveal whether such kind of long-term GeV emission is common, which would pose important constraints on GRB shock physical parameters.[186]

Another prominent issue is the high energy photon cutoff in the GRB spectra. During the escape of high energy photons, they are subject to absorption through $\gamma\gamma$ interactions with low energy photons within the source and in space. These interactions produce electron-positron pairs and greatly degrade the photon flux level in the original energy range received at the detector. At the source, such $\gamma\gamma$ pair process would result in a well-defined high energy cutoff[402] that would serve as an important diagnostic of the fireball initial Lorentz factor.[261] Another consequence of this process is that the produced pairs are expected to synchrotron radiate again within the local magnetic fields. Depending on the compactness of the emission region, the secondary emission process may modulate the emergent spectrum under certain conditions and possibly smear out the initial distinct spectral features.[313,73] A self-consistent pair-modulated numerical model for the GRB prompt emission is needed for this. Even if a high energy photon is not absorbed, it may be scattered by the resulting pairs and even by the electrons associated

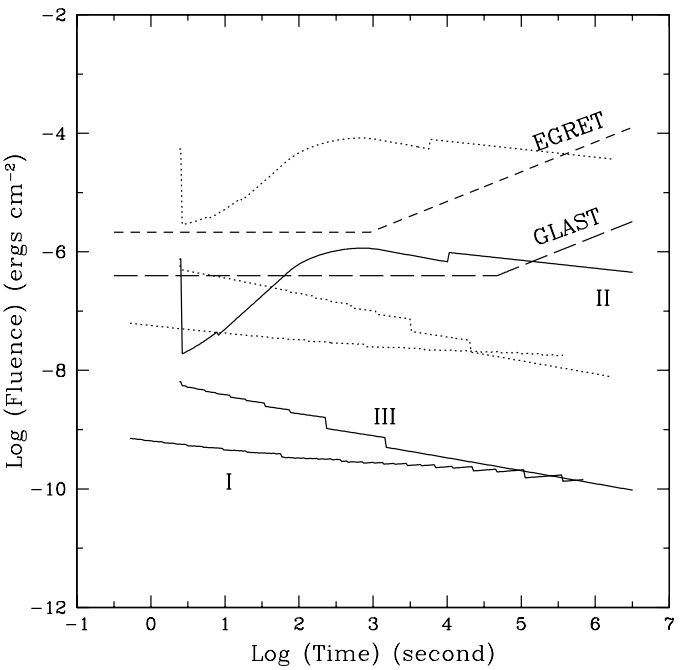

Fig. 15. Model prediction of GRB high energy (GeV) afterglows as compared with the GLAST sensitivity. For the IC-dominated parameter regime, an extended (hour-long) GeV afterglow should be detectable by GLAST for bursts at typical cosmological distances (from Ref. 186).

with the fireball baryons.[261] Furthermore, if the source is not compact enough so that the TeV photons can escape, they may be still absorbed by the cosmic IR background.[520] The mean free path of the TeV photon depends on both the GRB and the IR background models. Typically a TeV source cannot be detected beyond $z \sim 0.1$ (Ref. 520) so that the Milagrito event for GRB 970417,[185] if real, has to come from a nearby source whose compactness is not high (e.g. from external shocks, or with a high Lorentz factor). For typical cosmological sources, TeV emission will be almost completely absorbed by the IR background. The resultant pairs, if within a not-too-strong intergalactic medium (IGM) magnetic field, would IC up-scatter the cosmic microwave background photons to produce a delayed, GeV emission component.[521,522,519] A well-measured spectrum may be used to infer the poorly known strength of the IGM magnetic field (see Ref. 523 for a detailed modeling within the context of blazars).

5.4. *Ultrahigh energy cosmic rays*

Two important non-electromagnetic channels in which GRB may be prominent sources are cosmic rays and neutrinos, neither of which have so far been measured. GRB models involving shocks as sites to accelerate electrons which produce prompt

gamma-rays and long-term afterglows naturally suggests that baryons (most likely protons) should be accelerated by the same shocks as well. These accelerated ions, if not bound in the system and not destroyed during their propagation, would arrive at Earth and be detected as cosmic rays. Those trapped within the system would interact with photons and other baryons to produce high energy neutrinos that might also be detected from Earth.

Discussions about GRB's as cosmic ray accelerators are mainly aimed at explaining the portion of the cosmic ray spectrum above 10^{18} eV, the "ankle" region where the spectrum starts to harden, and in particular the ultrahigh energy cosmic rays (UHECR's) above 10^{20} eV, also referred to as Greisen–Zatsepin–Kuzmin ("GZK") cosmic rays. The isotropic distribution of their arrival directions and small magnetic deflection they suffer at these high energies suggest their extra-galactic origin, and the requirement that they must survive the attenuation by the cosmic microwave background through photomeson interaction constrains their distances to radius of about 50–100 Mpc, the so-called "GZK" volume.[524,525] Two broad classes of models have been suggested to interpret the UHECR's, the "top-down" scenarios that attribute UHECR's to decay products of fossil Grand Unification defects, and the "bottom-up" scenarios that suggest UHECR's are hadrons accelerated in astrophysical objects to these high energies. Among the few serious viable candidates in the bottom-up scenario, GRB's (and/or perhaps AGN's) are considered to be likely sources which may fulfill the constraints on the known data.[168,169,526] Two sub-scenarios of the bottom-up GRB cosmic ray origin have been suggested, one involving acceleration in the internal shocks[168,527,528] and one involving acceleration in the external shock.[169,529] The suggestion is based on two coincidences, i.e. the shock conditions required to accelerate protons to $\sim 10^{20}$ eV are similar to the conditions required for generating the observed prompt gamma-rays, and the observed UHECR energy injection rate into the universe ($\sim 3 \times 10^{44}$ erg Mpc^{-3} yr^{-1}) is similar to the local GRB energy injection rate.[168,169] Both coincidences have been questioned on various grounds,[530,147,531–533] but these have been met by effective counter-arguments, using new data and additional considerations for both the internal shock[528] and the external shock[529] scenarios. At this time, GRB's are, and remain, a promising candidate for the UHECR origin. However, both the internal shock and the external shock scenarios have major caveats, which have nothing to do with their ability to accelerate cosmic rays but rather with the generic shock itself. The internal shock scenario relies on the assumption that the GRB prompt emission is due to synchrotron from electrons accelerated in internal shocks. Although this is the leading scenario, there is no direct proof so far, unlike in the case of the external shock, where this origin of the radiation is quite convincing. For example, a Poynting-flux-dominated GRB model would be incompatible with an internal shock origin of both synchrotron γ-rays and UHECR's. The external shock model, on the other hand, may have to rely on a magnetized external medium[529] (as expected in pulsar wind bubbles[458] in the supranova scenario[96]) in order to reach the desired cosmic ray energy. On the other hand, the radiation

and acceleration responsible for γ-rays is itself unexplored in Poynting-dominated scenarios, and the need for such a scenario is at the moment not proven, while the supranova scenario is incompatible with the observations of GRB030329/SN2003dh. The uncertainties surrounding possible alternative scenarios are significant, and the standard shock scenario remains the most amenable to quantitative modeling and testing, which should be able to provide useful constraints on its ability to explain the data.

A direct proof of the GRB-origin of UHECR's is not easy. The next generation of cosmic ray detectors, such as the *Auger Observatory*, will have a substantially enhanced effective target area, which will greatly improve the cosmic ray count statistics. It also combines elements of the two techniques which currently lead to different results, namely air fluorescence telescopes and water Cerenkov surface tanks. This will help to disentangle the two distinct scenarios (top-down or bottom-up) and to reveal whether a GZK feature indeed exists. Within the bottom-up scenario, the direction information may prove or significantly constrain the AGN model (the close competitor of the GRB model) and eventually shed light on whether GRB's are indeed the sources of UHECR's.

5.5. *High energy neutrinos*

Regardless of whether GRB's can accelerate protons to ultrahigh (GZK) energies, they must be able to accelerate protons to some high energies. The implication is that high energy neutrinos and high energy photons (as discussed in Subsec. 5.3) must accompany the current emission seen at sub-MeV energies. Widely discussed processes for high energy neutrino emission include

- $p\gamma$ process: $p\gamma \to \Delta^+ \to n\pi^+ \to ne^+\nu_e\bar{\nu}_\mu\nu_\mu$;
- pp process: $pp \to \pi^\pm/K^\pm \cdots \to \mu\nu_\mu \cdots \to e\nu_e\bar{\nu}_\mu\nu_\mu \cdots$;
- pn process: $pn \to \pi^\pm/K^\pm \cdots \to \mu\nu_\mu \cdots \to e\nu_e\bar{\nu}_\mu\nu_\mu \cdots$.

The dominant $p\gamma$ process occurs at the Δ-resonance, which has the threshold condition $\epsilon_p\epsilon_\gamma \gtrsim 0.3$ GeV^2 in the center of mass frame. In the case of GRB's, this is usually translated in the observer's frame to $\epsilon_p\epsilon_\gamma \gtrsim 0.3$ GeV$^2\Gamma^2$ when both protons and photons are generated in relativistic shocks (e.g. in the internal or external shock scenarios), or to $\epsilon_p\epsilon_\gamma \gtrsim 0.3$ GeV$^2\Gamma$ when the photons are generated in SN shells while protons are accelerated from relativistic shocks (e.g. in the supranova model), where Γ is the bulk Lorentz factor. The threshold condition for both pp and pn interactions is that the relative drift energy between these baryons exceed the pion rest mass, i.e. $\epsilon' \geq 140$ MeV. Since both p and n have a rest mass close to 1 GeV, the threshold of pp and pn interaction only demands semi-relativistic relative motions.

In a GRB event, there are multiple sites where neutrinos with different energies are generated. Below is an non-exhaustive list which encompasses most of the processes discussed in the literature, in a sequence of ascending neutrino energy:

- MeV neutrinos: If GRB's are originated from stellar collapses, they should be associated with strong thermal MeV neutrinos like supernovae. In most models, the collapse results in a black hole–torus system, and the thermal neutrino annihilation is one of the leading processes for launching the fireball. However, these thermal neutrinos are extremely difficult to detect from cosmological distances, due to the very low cross section for νN interactions at these energies (the cross section increases $\propto E_\nu^2$ at lower energies, and $\propto E_\nu$ at higher energies).
- Multi-GeV neutrinos: GRB fireballs may be neutron-rich.[334,336] During the fireball acceleration phase, neutrons can decouple from protons when the elastic scattering condition breaks down. The relative drift between both species results in inelastic *pn* interactions giving rise to 5–10 GeV neutrinos.[171] A similar process also occurs within sub-photospheric internal shocks,[337] which extends significantly the parameter space for the inelastic neutrino collision condition. It was suggested earlier[65] that *pp* interactions within the internal shocks can also give rise to a 30 GeV neutrino burst, although magnetic fields can inhibit the inter-penetration of charged species streams with different velocities.
- Multi-TeV neutrinos: Within the collapsar scenario, the relativistic jet launched from the base of the flow (presumably the black hole and the torus) has to penetrate through the stellar envelope before breaking out and generating the GRB. The internal shocks below the envelope accelerate protons that interact with thermal photons within the envelope (i.e. $p\gamma$ interaction). Regardless of whether the jet finally penetrates through the envelope or gets choked, it will generate strong multi-TeV neutrino signals.[172] The signature is enhanced or even dominated by *pn*, *pp* interactions and could be used as a diagnostic about the type of progenitor stars.[534]
- PeV neutrinos: $p\gamma$ interactions within the conventional internal shocks which produce prompt gamma-rays typically generate $10^{14} - 10^{16}$ eV neutrinos.[170,535]
- EeV neutrinos: $p\gamma$ interactions within the external reverse shock give rise to even higher energy neutrinos. For a constant density medium the typical energy is $10^{17} - 10^{19}$ eV,[536] while for a wind medium, the typical energy is $\sim 3 \times (10^{15} - 10^{17})$ eV and extending above it.[537,540] In the forward shock region, assuming the blast wave can accelerate protons to ultrahigh energies,[169,529] a neutrino afterglow is expected with the peak energy $\sim 10^{18}$ eV.[538,539]

In the supranova progenitor scenario, the existence of the pre-supernova shell provides extra targets for additional neutrino signals. These signals are broadband with a distinct 1–10 TeV feature with high count rates, which is a useful clue to test the supranova hypothesis.[466,464,465,467] The neutrino model predictions for GRB 030329 indicate that a 2-day delay supranova model predicts a one order of magnitude higher neutrino event rate in the TeV–PeV regime than the conventional burst, a signal which is detectable by the full ICECUBE as an individual source at the observed redshift $z = 0.17$. If it were to be detected by under-ice neutrino detectors at the South Pole or underwater detectors such as

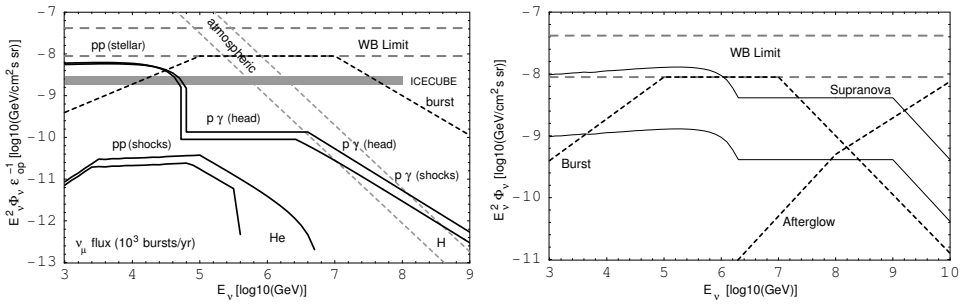

Fig. 16. The diffuse muon neutrino flux for both the collapsar and the supranova models (from Refs. 466 and 534, copyright (2003) by the American Physical Society).

Antares, this would further significantly constrain the validity of the supranova model.[540] The diffuse muon neutrino flux for both the collapsar and supranova models are presented in Fig. 16.[466,534]

All the above are theoretical predictions. So far there is no report about the solid detection of neutrino signals from GRB's. Current and future ice or water Cerenkov neutrino detectors, AMANDA-II, ANTARES, and ICECUBE will make it possible to detect GRB's from this new channel. If detected, neutrinos from GRB's could be also used to study neutrino physics itself as well as to have ramifications for cosmology.[541]

5.6. *Gravitational waves*

Binary compact-object mergers, such as NS–NS, NS–BH, BH–BH mergers, as well as BH–WD, BH-Helium star etc., have long been considered as possible sources of gravitational waves (GW),[542−544,173,350,545,174] as have core-collapse events.[546−548,174,549−551] This has been regardless of whether they could produce GRB's. However, since these events are also the leading candidates for being GRB progenitors, it is natural to expect that a GRB is associated with a GW burst (although the contrary may not hold). For a selection of GRB-related GW literatures, see e.g. Ref. 552. A coincidence between a GW signal and a gamma-ray signal would greatly enhance the statistical significance of the former,[553] making searching for GRB-associated GW bursts of great interest. A binary coalescence process can be divided into three phases, i.e. in-spiral, merger, and ring-down.[554,174] For collapsars, a rapidly rotating core could lead to development of a bar and to fragmentation instabilities that would resemble similar GW signals as in the binary merger scenarios, although a larger uncertainty is involved since the fragment masses and coherence times are unknown. The GW frequencies of the various phases cover the $10 - 10^3$ Hz band which is relevant for the *Laser Interferometer Gravitational-wave Observatory* (LIGO) and other related detectors such as VIRGO, GEO600 and TAMA300.[555] Because of the faint nature of the typical GW strain, only nearby sources (e.g. within $\sim$ 200 Mpc for NS–NS and NS–BH

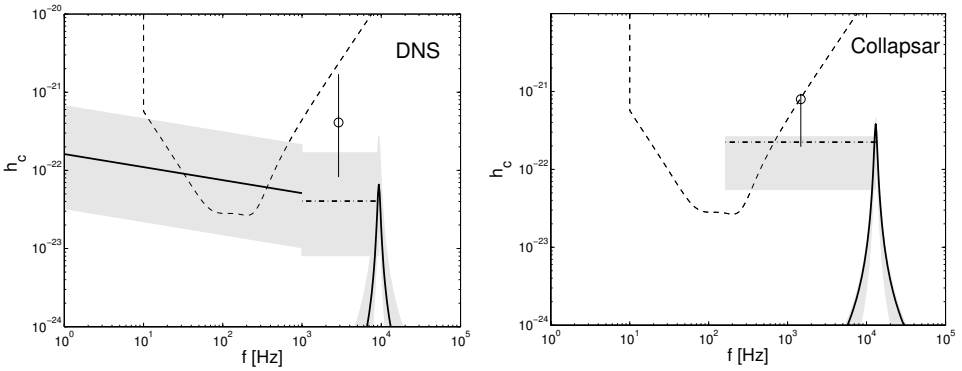

Fig. 17. The gravitational wave strains for the merging double neutron star model and the collapsar model as compared with the sensitivity of the advanced LIGO (from Ref. 174).

mergers, and within ~ 30 Mpc for collapsars)[174] have strong enough signals to be detectable by LIGO (Fig. 17). When event rates are taken into account,[150,556] order of magnitude estimates indicate that after one-year of operation of the advanced LIGO, one in-spiral chirp event from a NS–NS or NS–BH merger, and probably one collapsar event (subject to uncertainties) would be detected.[174,551] Other binary merger scenarios such as BH–WD and BH-Helium star mergers are unlikely to be detectable,[174] and they are also unfavored as sources of GRB's according to other arguments.[155]

The detection of a GRB–GW burst association would have profound implications on GRB studies, and could shed light on several unsettled issues as discussed in Sec. 4. First, it would help to identify the GRB progenitor. GW signals in the advanced LIGO band are different for the merger and the core collapse scenarios, so that a coincident GRB–GW detection would unambiguously differentiate both scenarios. Furthermore, two GW bursts are expected for the supranova model, with the second one coincident with the GRB. Either detection or non-detection of the precedent GW signal would provide additional information on the validity of the supranova model (in addition to the negative evidence provided by GRB030329/SN2003dh). Second, it could help to settle the debate about the location of GRB prompt emission (i.e. external or internal, Sec. 4.2). Because the GW wave is associated with the formation of the central engine (and therefore the launch of the GRB jet), a short time delay ($\lesssim 0.1$ s) for the GRB emission with respect to the GW emission would favor an internal scenario (e.g. the internal shock model), while a longer delay ($\sim 10 - 100$ s) would favor an external model (e.g. the external shock model).[150,557] Third, it would help to pin down the GRB jet configuration. It has been suggested that the GW polarization information could be used to infer geometric information such as the angle between the rotation axis (presumably the center of the jet) and the line-of-sight direction.[558] Such information, when combined with other information from prompt gamma-ray and afterglow, could help to confirm whether and how the GRB jets are structured.[135–137]

5.7. *GRB's and cosmology*

One of the most exciting prospects for future GRB studies is that they could become a unique tool to investigate the high redshift universe. Although this possibility is currently mainly a theoretical expectation, the prospects look bright. The optimism is underpinned by two aspects of the evidence obtained so far.

Clearly, since GRB's are stellar events, some GRB's are expected to exist at (much) higher redshifts than $z = 6.4$, the current high redshift record held by a SDSS quasar.[559] This is supported by several lines of reasoning. (1) Preliminary polarization data on the cosmic microwave background collected by *Wilkinson Microwave Background Probe* (WMAP) indicate a high electron scattering optical depth, hinting that the first stellar objects in the universe should have formed as early as $z \sim 20$.[560] (2) Independent theoretical simulations of the formation of the first stars similarly conclude that these should have formed at redshifts $z \sim (15 - 40)$.[561,562] Because there is convincing evidence that at least long GRB's are associated with the deaths of the massive stars, it is conceivable that high-z GRB's ($z \lesssim 15 - 20$ or even higher) exist. In fact, although subject to considerable model uncertainties, theoretical modeling of GRB redshift distribution based on the standard ΛCDM universe suggests that $\gtrsim 50\%$ of all GRB's on the sky originate at $z \gtrsim 5$.[165] (3) Using empirical luminosity laws such as gamma-ray variability or spectral lag correlations with fluence, rough redshifts can be derived for a large sample of the BATSE bursts, and it is found that a good fraction of these bursts have redshifts in excess of $z \sim 6$, and some even in excess of 10.[202,220,221]

The emission properties of GRB's and afterglows offer several unique advantages for studying the high-z universe (see also Ref. 563). (1) The luminosities of the prompt gamma-ray[163] and the afterglows[164,166] do not fade rapidly with increasing redshifts (unlike for quasars). In fact, the highest-redshift burst identified so far (GRB 000131 at $z = 4.5$) does not show a substantially fainter luminosity in both the prompt emission and the afterglow. This is due to a favorable combination of the cosmological time-dilation and the flux decay resulting in a positive K-correction effect. For afterglows, the high-z bursts are especially favorable for detection in the IR band, thanks to the bright reverse shock emission at early epochs.[166] (2) GRB afterglows have intrinsically high luminosities ($10^{51} - 10^{54}$ erg s^{-1}). Even at moderate redshifts, they already greatly out-shine the host galaxies. This would especially be the case at higher redshifts, as the host galaxies are expected to become even fainter. (3) The intrinsic power-law spectrum of GRB afterglows greatly simplifies the extraction of the IGM absorption features.

The implications of high-z GRB studies for cosmology include the following. (1) WMAP only provides an integral constraint on the reionization history of the universe. Since both the number density and the intrinsic luminosity of quasars are expected to fade rapidly beyond $z \sim 6$, only GRB's and afterglows may be able to act as bright beacons to illuminate the end of cosmic dark age,[564,565,167] and to probe the reionization history of the early universe.[566] (2) Although the

apparent GRB luminosity is by no means a standard candle, the beaming-corrected GRB luminosity appears to be standard.[36,131,259,256] Also empirical luminosity laws,[202,210] if confirmed, could be used as Cepheid-like correlations. These raise the possibility of using GRB's to derive a Hubble-diagram and to constrain the cosmological parameters,[256,567] although the current data are too sparse to draw a firm conclusion. (3) Preliminary evidence on the evolution of GRB properties has been suggested.[221,223] If this is confirmed by future photometric redshift measurements of a large amount of GRB's (with selection effects properly taken care of), the data could be used to infer the possible redshift-evolution of the GRB progenitors. (4) Future extensive afterglow monitoring for many bursts would help to constrain the local environments of GRB's as well as their redshift evolution. In particular, it would tell whether GRB afterglows are decelerated by the IGM (with an increasingly high density at higher redshifts) or by a stratified constant medium bubble cleared by the progenitor star.[166] (5) Insights into cosmological structure formation and star forming history would be gained through studying distributions of the GRB host galaxies.[568]

A great challenge in the *Swift* era is to get direct redshift information for high-z bursts. The most straightforward approach is to search for the Lyα break (at $1.216\ \mu\mathrm{m}[(1+z)/10]$) and the Gunn–Peterson trough[569] in the IR band. However, since the *Swift* UVOT wave-band extends only to $\sim 0.6\ \mu$m in the long wavelength regime, the redshifts of GRB's with $z > 5$ cannot be directly identified with *Swift*. This poses however a great opportunity for ground-based IR cameras to follow up *Swift*-triggered GRB's within a short period of time. Several IR follow-up teams have developed such instruments and plan to carry out *Swift* follow-up observations in a timely manner. Besides IR follow-up, it is also possible that X-ray lines could become important distance indicators at high redshifts.[570] Alternatively, redshifts for GRB's whose radio afterglows are unusually bright may be identified with 21 cm absorption features.[571,572]

Finally, using cosmological GRB high energy data, one may constrain a number of ideas in fundamental physics, e.g. related to quantum gravity effects, Lorentz invariance violations, etc.[573]

6. Summary

Our understanding of GRB's has been greatly advanced since their discovery about 30 years ago. Extensive observational efforts have revealed a rich phenomenology of both prompt emission and afterglows (Sec. 2), and a successful theoretical framework has been set up, which is able to interpret most of the observational data so far (Sec. 3). However, there remain many questions at the present stage of GRB studies, which greatly stimulate further observational and theoretical efforts (Sec. 4). Some of these questions will undoubtedly be addressed in the coming years with the advent of a number of upcoming space- and ground-based experiments, both in the electromagnetic and non-electromagnetic channels (Sec. 5). As suggested by past

experience, new challenges and surprises are bound to emerge, which will stimulate further extensive observational campaigns and theoretical efforts. The GRB field is likely to remain one of the most active fields in contemporary astrophysics in the next decades.

Acknowledgments

This research was partly supported through NASA ATP NAG5-13286, NSF AST0098416 and the Monell Foundation. We thank S. Kobayashi, T. Lu, D. A. Frail, J. P. Norris, E. D. Feigelson, Z. G. Dai, Y. F. Huang, and S. Razzaque for helpful comments.

References

1. G. J. Fishman and C. A. Meegan, *Annu. Rev. Astron. Astrophys.* **33**, 415 (1995).
2. J. van Paradijs, C. Kouveliotou and R. A. M. J. Wijers, *Annu. Rev. Astron. Astrophys.* **38**, 379 (2000).
3. S. Kulkarni *et al.*, in *Gamma-ray Bursts, 5th Huntsville Symposium*, eds. R. M. Kippen, R. S. Mallozzi and G. J. Fishman (Huntsville, Alabama, USA, 1999); *AIP Conf. Ser.* **526**, 277 (2000).
4. T. Piran, *Phys. Rep.* **314**, 575 (1999).
5. P. Mészáros, *Annu. Rev. Astron. Astrophys.* **40**, 137 (2002).
6. P. Mészáros, *Science* **291**, 79 (2001).
7. K. S. Cheng and T. Lu, *Chin. J. Astron. Astrophys.* **1**, 1 (2001).
8. S. E. Woosley, W. Zhang and A. Heger, astro-ph/0206004 (2002).
9. K. Hurley, R. Sari and S. G. Djorgovski, astro-ph/0211620 (2002).
10. C. D. Dermer, astro-ph/0202254 (2002).
11. B. Zhang, P. Mészáros and J. Wang, astro-ph/0212015 (2002).
12. G. Ghisellini, astro-ph/0301256 (2003).
13. S. G. Djorgovski *et al.*, *SPIE* **4834**, 238 (2003).
14. E. Waxman, astro-ph/0303517 (2003).
15. P. Mészáros, S. Kobayashi, S. Razzaque and B. Zhang, astro-ph/0305066 (2003).
16. R. Klebesadel, I. Strong and R. Olson, *Astrophys. J.* **182**, L85 (1973).
17. E. P. Mazets, S. V. Golenetskii and V. N. Ilinskii, *JETP Lett.* **19**, 77 (1974).
18. C. Meegan *et al.*, *Nature* **355**, 143 (1992).
19. C. Kouveliotou *et al.*, *Astrophys. J.* **413**, L101 (1993).
20. D. Band *et al.*, *Astrophys. J.* **413**, 281 (1993).
21. K. Hurley *et al.*, *Nature* **372**, 652 (1994).
22. J. P. Norris *et al.*, *Astrophys. J.* **424**, 540 (1994).
23. E. Costa *et al.*, *Nature* **387**, 783 (1997).
24. J. van Paradijs *et al.*, *Nature* **386**, 686 (1997).
25. D. Frail *et al.*, *Nature* **389**, 261 (1997).
26. M. R. Metzger *et al.*, *Nature* **387**, 878 (1997).
27. S. R. Kulkarni *et al.*, *Nature* **393**, 35 (1998).
28. T. J. Galama *et al.*, *Nature* **395**, 670 (1998).
29. S. R. Kulkarni *et al.*, *Nature* **395**, 663 (1998).
30. C. Akerlof *et al.*, *Nature* **398**, 400 (1999).
31. S. R. Kulkarni *et al.*, *Astrophys. J.* **522**, L97 (1999).

32. S. R. Kulkarni *et al.*, *Nature* **398**, 389 (1999).
33. F. A. Harrison *et al.*, *Astrophys. J.* **523**, L121 (1999).
34. L. Piro *et al.*, *Science* **290**, 955 (2000).
35. J. N. Reeves *et al.*, *Nature* **416**, 512 (2002).
36. D. Frail *et al.*, *Astrophys. J.* **562**, L155 (2001).
37. J. Heise *et al.*, astro-ph/0111246 (2001).
38. M. Kippen *et al.*, astro-ph/0203114 (2002).
39. D. Fox *et al.*, *Nature* **422**, 284 (2003).
40. D. Fox *et al.*, *Astrophys. J.* **586**, L5 (2003).
41. W. Li *et al.*, *Astrophys. J.* **586**, L9 (2003).
42. W. Coburn and S. E. Boggs, *Nature* **423**, 415 (2003).
43. R. B. Rutledge and D. B. Fox, *Mon. Not. R. Astron. Soc.*, in press, astro-ph/0310385 (2003).
44. S. E. Boggs and W. Coburn, *Mon. Not. R. Astron. Soc.*, submitted, astro-ph/0310515 (2003).
45. K. Z. Stanek *et al.*, *Astrophys. J.* **591**, L17 (2003).
46. J. Hjorth *et al.*, *Nature* **423**, 847 (2003).
47. M. M. Gonzalez *et al.*, *Nature* **424**, 751 (2003).
48. E. Berger *et al.*, *Nature* **426**, 154 (2003).
49. K. Sheth *et al.*, *Astrophys. J.* **595**, L33 (2003).
50. A. K. Harding, *Phys. Rep.* **206**, 327 (1991).
51. M. Ruderman, *Ann. N.Y. Acad. Sci.* **262**, 164 (1975).
52. R. Blandford and C. McKee, *Phys. Fluids* **19**, 1130 (1976).
53. G. Cavallo and M. J. Rees, *Mon. Not. R. Astron. Soc.* **183**, 359 (1978).
54. B. Paczyǹski, *Astrophys. J.* **308**, L43 (1986).
55. J. Goodman, *Astrophys. J.* **308**, L47 (1986).
56. B. Paczyǹski, *Astrophys. J.* **363**, 218 (1990).
57. A. Shemi and T. Piran, *Astrophys. J.* **365**, L55 (1990).
58. M. J. Rees and P. Mészáros, *Mon. Not. R. Astron. Soc.* **258**, P41 (1992).
59. P. Mészáros and M. J. Rees, *Astrophys. J.* **405**, 278 (1993).
60. D. Eichler, M. Livio, T. Piran and D. Schramm, *Nature* **340**, 126 (1989).
61. R. Narayan, B. Paczyński and T. Piran, *Astrophys. J.* **395**, L8 (1992).
62. S. E. Woosley, *Astrophys. J.* **405**, 273 (1993).
63. V. V. Usov, *Nature* **357**, 472 (1992).
64. M. J. Rees and P. Mészáros, *Astrophys. J.* **430**, L93 (1994).
65. B. Paczyński and G. Xu, *Astrophys. J.* **427**, 708 (1994).
66. P. Mészáros, P. Laguna and M. J. Rees, *Astrophys. J.* **415**, 181 (1993).
67. P. Mészáros and M. J. Rees, *Astrophys. J.* **418**, L59 (1993).
68. T. Piran, A. Shemi and R. Narayan, *Mon. Not. R. Astron. Soc.* **263**, 861 (1993).
69. J. I. Katz, *Astrophys. J.* **422**, 248 (1994).
70. R. Sari and T. Piran, *Astrophys. J.* **455**, L143 (1995).
71. S. Kobayashi, T. Piran and R. Sari, *Astrophys. J.* **513**, 669 (1999).
72. P. Mészáros and M. J. Rees, *Astrophys. J.* **530**, 292 (2000).
73. P. Mészáros, E. Ramirez-Ruiz, M. J. Rees and B. Zhang, *Astrophys. J.* **578**, 812 (2002).
74. P. Mészáros, M. J. Rees and H. Papathanassiou, *Astrophys. J.* **432**, 181 (1994).
75. P. Mészáros and M. J. Rees, *Mon. Not. R. Astron. Soc.* **269**, L41 (1994).
76. B. Paczyński and J. E. Rhoads, *Astrophys. J.* **418**, L5 (1993).
77. J. Katz, *Astrophys. J.* **432**, L107 (1994).
78. P. Mészáros and M. J. Rees, *Astrophys. J.* **476**, 232 (1997).

79. R. A. M. I. Wijers, M. J. Rees and P. Mészáros, *Mon. Not. R. Astron. Soc.* **288**, L51 (1997).
80. M. Vietri, *Astrophys. J.* **478**, L9 (1997).
81. E. Waxman, *Astrophys. J.* **485**, L5 (1997).
82. E. Waxman, *Astrophys. J.* **489**, L33 (1997).
83. M. Tavani, *Astrophys. J.* **483**, L97 (1997).
84. B. Paczyński, *Astrophys. J.* **494**, L45 (1998).
85. P. Mészáros, M. J. Rees and R. A. M. I. Wijers, *Astrophys. J.* **499**, 301 (1998).
86. R. Sari, T. Piran and R. Narayan, *Astrophys. J.* **497**, L17 (1998).
87. M. J. Rees and P. Mészáros, *Astrophys. J.* **496**, L1 (1998).
88. Z. G. Dai and T. Lu, *Mon. Not. R. Astron. Soc.* **298**, 87 (1998).
89. Z. G. Dai and T. Lu, *Astrophys. J.* **519**, L155 (1999).
90. R. Chevalier and Z.-Y. Li, *Astrophys. J.* **520**, L29 (1999).
91. A. I. MacFadyen and S. E. Woosley, *Astrophys. J.* **524**, 262 (1999).
92. M. Aloy *et al.*, *Astrophys. J.* **531**, L119 (2000).
93. W. Zhang, S. E. Woosley and A. I. MacFadyen, *Astrophys. J.* **586**, 356 (2003).
94. P. Mészáros and M. J. Rees, *Astrophys. J.* **556**, L37 (2001).
95. E. Waxman and P. Mészáros, *Astrophys. J.* **584**, 390 (2003).
96. M. Vietri and L. Stella, *Astrophys. J.* **507**, L45 (1998).
97. R. Sari and T. Piran, *Astrophys. J.* **485**, 207 (1997).
98. R. Sari and T. Piran, *Mon. Not. R. Astron. Soc.* **287**, 110 (1997).
99. S. Kobayashi, T. Piran and R. Sari, *Astrophys. J.* **490**, 92 (1997).
100. F. Daigne and R. Mochkovitch, *Mon. Not. R. Astron. Soc.* **296**, 275 (1998).
101. A. Panaitescu, M. Spada and P. Mészáros, *Astrophys. J.* **522**, L105 (1999).
102. M. Spada, A. Panaitescu and P. Mészáros, *Astrophys. J.* **537**, 824 (2000).
103. D. Guetta, M. Spada and E. Waxman, *Astrophys. J.* **557**, 399 (2001).
104. A. Panaitescu and P. Mészáros, *Astrophys. J.* **492**, 683 (1998).
105. J. Chiang and C. D. Dermer, *Astrophys. J.* **512**, 899 (1999).
106. C. D. Dermer and K. E. Mitman, *Astrophys. J.* **513**, L5 (1999).
107. C. D. Dermer, M. Böttcher and J. Chiang, *Astrophys. J.* **515**, L49 (1999).
108. J. E. Rhoads, *Astrophys. J.* **487**, L1 (1997).
109. J. E. Rhoads, *Astrophys. J.* **525**, 737 (1999).
110. R. Sari, T. Piran and J. P. Halpern, *Astrophys. J.* **519**, L17 (1999).
111. A. Panaitescu and P. Mészáros, *Astrophys. J.* **526**, 707 (1999).
112. R. Moderski, M. Sikora and T. Bulik, *Astrophys. J.* **529**, 151 (2000).
113. Y. F. Huang, L. J. Gou, Z. G. Dai and T. Lu, *Astrophys. J.* **543**, 90 (2000).
114. R. Sari and T. Piran, *Astrophys. J.* **517**, L109 (1999).
115. R. Sari and T. Piran, *Astrophys. J.* **520**, 641 (1999).
116. P. Mészáros and M. J. Rees, *Mon. Not. R. Astron. Soc.* **306**, L39 (1999).
117. S. Kobayashi and R. Sari, *Astrophys. J.* **542**, 819 (2000).
118. S. Kobayashi, *Astrophys. J.* **545**, 807 (2000).
119. S. Kobayashi and B. Zhang, *Astrophys. J.* **582**, L75 (2003).
120. B. Zhang, S. Kobayashi and P. Mészáros, *Astrophys. J.* **595**, 950 (2003).
121. X. F. Wu, Z. G. Dai, Y. F. Huang and T. Lu, *Mon. Not. R. Astron. Soc.* **342**, 1131 (2003).
122. S. Kobayashi and B. Zhang, *Astrophys. J.* **597**, 455 (2003).
123. S. Kobayashi, P. Mészáros and B. Zhang, *Astrophys. J.* **601**, L13 (2004).
124. D. Lazzati, S. Campana and G. Ghisellini, *Mon. Not. R. Astron. Soc.* **304**, L31 (1999).
125. M. J. Rees and P. Mészáros, *Astrophys. J.* **545**, L73 (2000).

126. C. Weth, P. Mészáros, T. Kallman and M. J. Rees, *Astrophys. J.* **534**, 581 (2000).
127. M. Vietri *et al.*, *Astrophys. J.* **550**, L43 (2001).
128. D. Lazzati, E. Ramirez-Ruiz and M. J. Rees, *Astrophys. J.* **572**, L57 (2002).
129. T. Kallman, P. Mészáros and M. J. Rees, *Astrophys. J.* **593**, 946 (2003).
130. R. A. M. J. Wijers and T. J. Galama, *Astrophys. J.* **523**, 177 (1999).
131. A. Panaitescu and P. Kumar, *Astrophys. J.* **560**, L49 (2001).
132. A. Panaitescu and P. Kumar, *Astrophys. J.* **571**, 779 (2002).
133. F. A. Harrison *et al.*, *Astrophys. J.* **559**, 123 (2001).
134. S. Yost, F. A. Harrison, R. Sari and D. A. Frail, *Astrophys. J.* **597**, 459 (2003).
135. E. Rossi, D. Lazzati and M. J. Rees, *Mon. Not. R. Astron. Soc.* **332**, 945 (2002).
136. B. Zhang and P. Mészáros, *Astrophys. J.* **571**, 876 (2002).
137. B. Zhang, X. Dai, N. M. Lloyd-Ronning and P. Mészáros, *Astrophys. J.* **601**, L119 (2004).
138. E. P. Mazets *et al.*, *Nature* **290**, 378 (1981).
139. T. Murakami *et al.*, *Nature* **335**, 234 (1988).
140. R. C. Duncan and C. Thompson, *Astrophys. J.* **392**, L9 (1992).
141. C. Thompson and R. C. Duncan, *Mon. Not. R. Astron. Soc.* **275**, 255 (1995).
142. C. Thompson and R. C. Duncan, *Astrophys. J.* **473**, 322 (1996).
143. P. Kumar and A. Panaitescu, *Mon. Not. R. Astron. Soc.* **346**, 905 (2003).
144. A. Heger *et al.*, *Astrophys. J.* **591**, 288 (2003).
145. S. E. Woosley, W. Zhang and A. Heger, astro-ph/0211063 (2002).
146. K. Iwamoto *et al.*, *Nature* **395**, 672 (1998).
147. A. Achterberg *et al.*, *Mon. Not. R. Astron. Soc.* **328**, 393 (2001).
148. D. C. Ellison and G. P. Double, *Astropart. Phys.* **18**, 213 (2002).
149. M. Lemoine and G. Pelletier, *Astrophys. J.* **589**, L73 (2003).
150. C. L. Fryer, S. E. Woosley and D. H. Hartmann, *Astrophys. J.* **526**, 152 (1999).
151. R. D. Blandford and R. L. Znajek, *Mon. Not. R. Astron. Soc.* **179**, 433 (1977).
152. P. Mészáros and M. J. Rees, *Astrophys. J.* **482**, L29 (1997).
153. R. Popham, S. E. Woosley and C. Fryer, *Astrophys. J.* **518**, 356 (1999).
154. H. K. Lee, R. A. M. J. Wijers and G. E. Brown, *Phys. Rep.* **325**, 83 (2000).
155. R. Narayan, T. Piran and P. Kumar, *Astrophys. J.* **557**, 949 (2001).
156. M. H. P. M. van Putten, *Phys. Rep.* **345**, 1 (2001).
157. X. Wang and A. Loeb, *Astrophys. J.* **535**, 788 (2000).
158. Z. G. Dai and T. Lu, *Astrophys. J.* **565**, L87 (2002).
159. B. E. Schaefer *et al.*, *Astrophys. J.* **588**, 387 (2003).
160. J. S. Bloom, S. G. Djorgovski, S. R. Kulkarni and D. A. Frail, *Astrophys. J.* **507**, L25 (1998).
161. A. S. Fruchter *et al.*, *Astrophys. J.* **519**, L13 (1999).
162. J. S. Bloom, S. R. Kulkarni and S. G. Djorgovski, *Astron. J.* **123**, 1111 (2002).
163. D. Q. Lamb and D. E. Reichart, *Astrophys. J.* **536**, 1 (2000).
164. B. Ciardi and A. Loeb, *Astrophys. J.* **540**, 687 (2000).
165. V. Bromm and A. Loeb, *Astrophys. J.* **575**, 111 (2002).
166. L. J. Gou, P. Mészáros, T. Abel and B. Zhang, *Astrophys. J.* **604**, 508 (2004).
167. J. Miralda-Escudé, *Science* **300**, 1904 (2003).
168. E. Waxman, *Phys. Rev. Lett.* **75**, 386 (1995).
169. M. Vietri, *Astrophys. J.* **453**, 883 (1995).
170. E. Waxman and J. Bahcall, *Phys. Rev. Lett.* **78**, 2292 (1997).
171. J. N. Bahcall and P. Mészáros, *Phys. Rev. Lett.* **85**, 1362 (2000).
172. P. Mészáros and E. Waxman, *Phys. Rev. Lett.* **87**, 171102 (2001).
173. C. Kochanek and T. Piran, *Astrophys. J.* **417**, L17 (1993).

174. S. Kobayashi and P. Mészáros, *Astrophys. J.* **589**, 861 (2003).
175. R. J. Nemiroff *et al.*, *Astrophys. J.* **423**, 432 (1994).
176. J. P. Norris, *Astrophys. J.* **579**, 386 (2002).
177. P. Bhat *et al.*, *Astrophys. J.* **426**, 604 (1994).
178. R. D. Preece *et al.*, *Astrophys. J. Suppl.* **126**, 19 (2000).
179. M. Harris and G. Share, *Astrophys. J.* **494**, 724 (1998).
180. R. S. Mallozzi *et al.*, *Astrophys. J.* **454**, 597 (1995).
181. B. L. Schaefer, *Astrophys. J.* **583**, L71 (2003).
182. J. P. Norris *et al.*, *Astrophys. J.* **301**, 213 (1986).
183. L. Ford *et al.*, *Astrophys. J.* **439**, 307 (1986).
184. B. L. Dingus *et al.*, in *High Energy Gamma Ray Astronomy*, ed. F. A. Aharonian and H. J. Völk (AIP, New York, 2001), p. 383.
185. R. Atkins *et al.*, *Astrophys. J.* **533**, L119 (2000).
186. B. Zhang and P. Mészáros, *Astrophys. J.* **559**, 110 (2001).
187. L. Amati *et al.*, *Science* **290**, 953 (2000).
188. D. A. Frail *et al.*, GCN notice 2280 (2003).
189. W. S. Paciesas *et al.*, *Astrophys. J. Suppl.* **122**, 465 (1999).
190. Mészáros, A. *et al.*, *Astrophys. J.* **539**, 98 (2000).
191. M. Schmidt, *Astrophys. J.* **552**, 36 (2001).
192. M. H. P. M. van Putten and T. Regimbau, *Astrophys. J.* **593**, L15 (2003).
193. S. Mukherjee *et al.*, *Astrophys. J.* **508**, 314 (1998).
194. I. Horváth, *Astrophys. J.* **508**, 757 (1998).
195. I. Horváth, *Astron. Astrophys.* **392**, 791 (1998).
196. C. Barraud *et al.*, *Astron. Astrophys.* **400**, 1021 (2003).
197. A. M. Beloborodov, B. E. Stern and R. Svensson, *Astrophys. J.* **508**, L25 (1998).
198. A. M. Beloborodov, B. E. Stern and R. Svensson, *Astrophys. J.* **535**, 158 (2000).
199. H. Li and E. E. Fenimore, *Astrophys. J.* **469**, L115 (1996).
200. E. Nakar and T. Piran, *Mon. Not. R. Astron. Soc.* **331**, 40 (2002).
201. E. Ramirez-Ruiz and A. Merloni, *Mon. Not. R. Astron. Soc.* **320**, L25 (2001).
202. E. E. Fenimore and E. Ramirez-Ruiz, astro-ph/0004176 (2000).
203. D. E. Reichart *et al.*, *Astrophys. J.* **552**, 57 (2001).
204. S. McBreen *et al.*, *Astron. Astrophys.* **393**, L29 (2002).
205. S. McBreen *et al.*, *Astron. Astrophys.* **393**, L15 (2002).
206. S. V. Golenetskii *et al.*, *Nature* **306**, 451 (1983).
207. E. Liang and V. Kargatis, *Nature* **381**, 49 (1996).
208. J. P. Norris *et al.*, *Astrophys. J.* **459**, 393 (1996).
209. L. X. Cheng *et al.*, *Astron. Astrophys.* **300**, 746 (1995).
210. J. P. Norris, G. F. Marani and J. T. Bonnell, *Astrophys. J.* **534**, 248 (2000).
211. A. Crider *et al.*, *Astrophys. J.* **479**, L39 (1997).
212. N. M. Lloyd-Ronning and V. Petrosian, *Astrophys. J.* **543**, 722 (2002).
213. N. M. Lloyd, V. Petrosian and R. S. Mallozzi, *Astrophys. J.* **534**, 227 (2000).
214. N. M. Lloyd-Ronning and E. Ramirez-Ruiz, *Astrophys. J.* **576**, 101 (2002).
215. L. Amati *et al.*, *Astron. Astrophys.* **390**, 81 (2002).
216. J. L. Atteia, *Astron. Astrophys.* **407**, L1 (2003).
217. T. Sakamoto *et al.*, *Astrophys. J.* **602**, 875 (2004).
218. D. Q. Lamb, T. Q. Donaghy and C. Graziani, to appear in *Proc. 2nd VERITAS Symposium on TeV Astrophysics* (Chicago, Illinois, 2003), astro-ph/0309456.
219. Z. Bagoly *et al.*, *Astron. Astrophys.* **398**, 919 (2003).
220. B. E. Schaefer, M. Deng and D. L. Band, *Astrophys. J.* **563**, L123 (2001).

221. N. M. Lloyd-Ronning, C. L. Fryer and E. Ramirez-Ruiz, *Astrophys. J.* **574**, 554 (2002).
222. B. E. Stern, Ya. Tikhomirova and R. Svensson, *Astrophys. J.* **573**, 75 (2002).
223. D. M. Wei and W. H. Gao, *Mon. Not. R. Astron. Soc.* **345**, 743 (2003).
224. D. A. Frail, astro-ph/0309557 (2003).
225. G. Ricker, talk presented in GRB 2003 meeting, Santa Fe, USA (2003).
226. J. Greiner, http://www.mpe.mpg.de/∼jcg/grb.html
227. J. S. Bloom *et al.*, *Nature* **401**, 453 (1999).
228. D. Reichart, *Astrophys. J.* **521**, L111 (1999).
229. T. J. Galama *et al.*, *Astrophys. J.* **536**, 185 (2000).
230. D. Lazzati *et al.*, *Astron. Astrophys.* **378**, 996 (2001).
231. A. J. Castro-Tirado *et al.*, *Astron. Astrophys.* **370**, 398 (2001).
232. K. C. Sahu *et al.*, *Astrophys. J.* **540**, 74 (2000).
233. J. S. Bloom *et al.*, *Astrophys. J.* **572**, L45 (2002).
234. P. Price *et al.*, *Astrophys. J.* **589**, 838 (2003).
235. L. Piro, in *Gamma-Ray Bursts in the Afterglow Era*, eds. E. Costa *et al.* (Springer, Berlin, Heidelberg, 2001).
236. L. Piro *et al.*, *Astrophys. J.* **514**, L73 (1999).
237. A. Yoshida *et al.*, *Astrophys. J.* **557**, L27 (2001).
238. L. A. Antonelli *et al.*, *Astrophys. J.* **545**, L39 (2000).
239. N. R. Butler *et al.*, *Astrophys. J.* **597**, 1010 (2003).
240. V. Connaughton, *Astrophys. J.* **567**, 1028 (2002).
241. http://gcn.gsfc.nasa.gov/gcn/gcn3_archive.html
242. G. Björnsson, astro-ph/0302177 (2003).
243. S. Covino *et al.*, astro-ph/0301608 (2003).
244. J. Greiner *et al.*, *Nature* **426**, 157 (2003).
245. E. Rol *et al.*, *Astron. Astrophys.* **405**, L23 (2003).
246. D. A. Frail *et al.*, *Astron. J.* **125**, 2299 (2003).
247. D. A. Frail, *et al.*, *Astrophys. J.* **600**, 828 (2004).
248. S. R. Kulkarni *et al.*, *Astrophys. J.* **522**, L97 (1999).
249. D. A. Frail *et al.*, *Astrophys. J.* **538**, L129 (2000).
250. D. A. Frail *et al.*, *Astrophys. J.* **537**, 191 (2000).
251. J. Goodman, *New Astron.* **2**, 449 (1997).
252. E. Waxman, S. R. Kulkarni and D. A. Frail, *Astrophys. J.* **497**, 288 (1998).
253. S. G. Djorgovski *et al.*, *Astrophys. J.* **562**, 654 (2001).
254. E. Berger *et al.*, *Astrophys. J.* **581**, 981 (2002).
255. J.-L. Atteia *et al.*, GCN Notice 2432 (2003).
256. J. S. Bloom, D. A. Frail and S. R. Kulkarni, *Astrophys. J.* **594**, 674 (2003).
257. D. L. Freedman and E. Waxman, *Astrophys. J.* **547**, 922 (2001).
258. T. Piran, P. Kumar, A. Panaitescu and L. Piro, *Astrophys. J.* **560**, L167 (2001).
259. E. Berger, S. R. Kulkarni and D. A. Frail, *Astrophys. J.* **590**, 379 (2003).
260. J. D. Salmonson and T. J. Galama, *Astrophys. J.* **569**, 682 (2002).
261. Y. Lithwick and R. Sari, *Astrophys. J.* **555**, 540 (2001).
262. J. H. Krolik and E. A. Pier, *Astrophys. J.* **373**, 270 (1991).
263. E. E. Fenimore, R. I. Epstein and C. Ho, *Astron. Astrophys. Suppl.* **97**, 59 (1993).
264. E. Woods and A. Loeb, *Astrophys. J.* **453**, 583 (1995).
265. M. G. Baring and A. K. Harding, *Astrophys. J.* **491**, 663 (1997).
266. E. E. Fenimore, C. D. Madras and S. Nayakshin, *Astrophys. J.* **473**, 998 (1996).
267. B. Zhang and P. Mészáros, *Astrophys. J.* **566**, 712 (2002).

268. G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics* (Wiley-Interscience, New York, 1979).
269. J. Granot, T. Piran and R. Sari, *Astrophys. J.* **527**, 236 (1999).
270. A. Gruzinov and E. Waxman, *Astrophys. J.* **511**, 852 (1999).
271. M. V. Medvedev and A. Loeb, *Astrophys. J.* **526**, 697 (1999).
272. A. Gruzinov, *Astrophys. J.* **525**, L29 (1999).
273. G. Ghisellini and D. Lazzati, *Mon. Not. R. Astron. Soc.* **309**, L7 (1999).
274. R. Sari, *Astrophys. J.* **524**, L43 (1999).
275. A. Panaitescu and P. Mészáros, *Astrophys. J.* **501**, 772 (1998).
276. E. Cohen, T. Piran and R. Sari, *Astrophys. J.* **509**, 717 (1998).
277. T. Piran and R. Sari, in *Gamma-Ray Bursts*: *4th Hunstville Symp.*, eds. C. A. Meegan, R. D. Preece and T. M. Koshut (AIP Conf. Proc. 428) (AIP, New York, 1998), p. 662.
278. M. Böttcher and C. D. Dermer, *Astrophys. J.* **532**, 281 (2000).
279. R. Sari, *Astrophys. J.* **489**, L37 (1997).
280. E. Waxman, *Astrophys. J.* **491**, L19 (1997).
281. A. Panaitescu and P. Mészáros, *Astrophys. J.* **493**, L31 (1998).
282. R. Sari, *Astrophys. J.* **494**, L49 (1998).
283. J. Granot, T. Piran and R. Sari, *Astrophys. J.* **513**, 679 (1999).
284. J. Granot and R. Sari, *Astrophys. J.* **568**, 820 (2002).
285. D. M. Wei, *Astron. Astrophys.* **402**, L9 (2003).
286. J. Granot, A. Panaitescu, P. Kumar and S. E. Woosley, *Astrophys. J.* **570**, L61 (2002).
287. J. Granot *et al.*, astro-ph/0103038 (2001).
288. J. K. Cannizzo, N. Gehrels and E. T. Vishniac, *Astrophys. J.* **601**, 380 (2004).
289. R. A. Chevalier and Z.-Y. Li, *Astrophys. J.* **536**, 195 (2000).
290. P. Kumar and A. Panaitescu, *Astrophys. J.* **541**, L9 (2000).
291. L. J. Gou, Z. G. Dai, Y. F. Huang and T. Lu, *Astron. Astrophys.* **368**, 464 (2001).
292. Z.-Y. Li and R. A. Chevalier, *Astrophys. J.* **589**, L69 (2003).
293. Z.-Y. Li and R. A. Chevalier, astro-ph/0110002 (2001).
294. E. Ramirez-Ruiz, L. M. Dray, P. Madau and C. A. Tout, *Mon. Not. R. Astron. Soc.*, **327**, 829 (2001).
295. Z. G. Dai and X. F. Wu, *Astrophys. J.* **591**, L21 (2003).
296. D. Lazzati *et al.*, *Astron. Astrophys.* **396**, L5 (2002).
297. E. Nakar, T. Piran and J. Granot, *New Astron.* **8**, 495 (2003).
298. J. Heyl and R. Perna, *Astrophys. J.* **586**, L13 (2003).
299. P. Kumar and A. Panaitescu, *Astrophys. J.* **541**, L51 (2000).
300. Z. G. Dai and T. Lu, *Phys. Rev. Lett.* **81**, 4301 (1998).
301. A. Panaitescu, P. Mészáros and M. J. Rees, *Astrophys. J.* **503**, 314 (1998).
302. R. Sari and P. Mészáros, *Astrophys. J.* **535**, L33 (2000).
303. B. Zhang and P. Mészáros, *Astrophys. J.* **552**, L35 (2001).
304. P. Kumar and T. Piran, *Astrophys. J.* **532**, 286 (2000).
305. J. Granot, E. Nakar and T. Piran, *Nature* **426**, 138 (2003).
306. P. Kumar and T. Piran, *Astrophys. J.* **535**, 152 (2000).
307. Y. F. Huang, Z. G. Dai and T. Lu, *Mon. Not. R. Astron. Soc.* **309**, 513 (1999).
308. D. Frail, E. Waxman and S. R. Kulkarni, *Astrophys. J.* **537**, 191 (2000).
309. M. Livio and E. Waxman, *Astrophys. J.* **538**, 187 (2000).
310. Y. F. Huang and K. S. Cheng, *Mon. Not. R. Astron. Soc.* **341**, 263 (2003).
311. H. Papathanassiou and P. Mészáros, *Astrophys. J.* **471**, L91 (1996).

312. R. Sari, R. Narayan and T. Piran, *Astrophys. J.* **473**, 204 (1996).
313. R. P. Pilla and A. Loeb, *Astrophys. J.* **494**, L167 (1998).
314. D. M. Wei and T. Lu, *Astrophys. J.* **505**, 252 (1998).
315. C. D. Dermer, M. Böttcher and J. Chiang, *Astrophys. J.* **537**, 255 (2000).
316. C. D. Dermer, J. Chiang and K. E. Mitman, *Astrophys. J.* **537**, 785 (2000).
317. A. Panaitescu and P. Kumar, *Astrophys. J.* **543**, 66 (2000).
318. A. Panaitescu and P. Mészáros, *Astrophys. J.* **544**, L17 (2000).
319. R. Sari and A. A. Esin, *Astrophys. J.* **548**, 787 (2001).
320. X. Y. Wang, Z. G. Dai and T. Lu, *Astrophys. J.* **546**, L33 (2001).
321. M. Vietri, *Phys. Rev. Lett.* **78**, 4328 (1997).
322. M. Böttcher and C. D. Dermer, *Astrophys. J.* **499**, L131 (1998).
323. T. Totani, *Astrophys. J.* **502**, L13 (1998).
324. N. Masetti *et al.*, *Astron. Astrophys.* **374**, 382 (2001).
325. K. Z. Stanek *et al.*, *Astrophys. J.* **563**, 592 (2001).
326. Z. G. Dai and K. S. Cheng, *Astrophys. J.* **558**, L109 (2001).
327. A. Panaitescu and P. Kumar, *Mon. Not. R. Astron. Soc.*, in press, astro-ph/0308273 (2003).
328. E. Rossi and M. J. Rees, *Mon. Not. R. Astron. Soc.* **339**, 881 (2003).
329. P. Madau and C. Thompson, *Astrophys. J.* **534**, 239 (2000).
330. C. Thompson and P. Madau, *Astrophys. J.* **538**, 105 (2000).
331. C. D. Dermer and M. Böttcher, *Astrophys. J.* **534**, L155 (2000).
332. P. Mészáros, E. Ramirez-Ruiz and M. J. Rees, *Astrophys. J.* **554**, 660 (2001).
333. A. M. Beloborodov, *Astrophys. J.* **565**, 808 (2002).
334. E. V. Derishev, V. V. Kocharovsky and Vl. V. Kocharovsky, *Astrophys. J.* **521**, 640 (1999).
335. A. M. Beloborodov, *Astrophys. J.* **585**, L19 (2003).
336. A. M. Beloborodov, *Astrophys. J.* **588**, 931 (2003).
337. P. Mészáros and M. J. Rees, *Astrophys. J.* **541**, L5 (2000).
338. A. A. Esin and R. Blandford, *Astrophys. J.* **534**, L151 (2000).
339. E. Waxman and B. T. Draine, *Astrophys. J.* **537**, 796 (2000).
340. A. Fruchter, J. H. Krolik and J. E. Rhoads, *Astrophys. J.* **563**, 597 (2001).
341. P. Mészáros and A. Gruzinov, *Astrophys. J.* **543**, L35 (2000).
342. B. Zhang and P. Mészáros, *Astrophys. J.* **581**, 1236 (2002).
343. S. Kobayashi, F. Ryde and A. MacFadyen, *Astrophys. J.* **577**, 302 (2002).
344. P. Kumar, *Astrophys. J.* **532**, L113 (1999).
345. A. M. Beloborodov, *Astrophys. J.* **539**, L25 (2000).
346. S. Kobayashi and R. Sari, *Astrophys. J.* **551**, 934 (2001).
347. E. Rossi, A. M. Beloborodov and M. J. Rees, astro-ph/0401355.
348. K. Asano and S. Kobayashi, *Pub. Astron. Soc. Japan* **55**, 579 (2003).
349. A. Levinson and D. Eichler, *Astrophys. J.* **594**, L19 (2003).
350. M. Ruffert *et al.*, *Astron. Astrophys.* **319**, 122 (1997).
351. M. Ruffert and H.-Th. Janka, *Astron. Astrophys.* **344**, 573 (1999).
352. P. Mészáros, M. J. Rees and R. A. M. J. Wijers, *New Astron.* **4**, 303 (1999).
353. N. Vlahakis and A. Königl, *Astrophys. J.* **596**, 1080 (2003).
354. N. Vlahakis and A. Königl, *Astrophys. J.* **596**, 1104 (2003).
355. H. C. Spruit, F. Daigne and G. Drenkhahn, *Astron. Astrophys.* **369**, 694 (2001).
356. G. Drenkhahn and H. C. Spruit, *Astron. Astrophys.* **391**, 1141 (2002).
357. M. Lyutikov, V. I. Pariev and R. Blandford, *Astrophys. J.* **597**, 998 (2003).
358. E. Waxman, *Nature* **423**, 388 (2003).
359. J. Granot, *Astrophys. J.* **596**, L17 (2003).

360. Y. Z. Fan, Z. G. Dai, Y. F. Huang and T. Lu, *Chin. J. Astron. Astrophys.* **2**, 449 (2002).
361. C. F. Kennel and F. V. Coroniti, *Astrophys. J.* **283**, 694, (1984).
362. F. Daigne and R. Mochkovitch, *Mon. Not. R. Astron. Soc.* **336**, 1271 (2002).
363. G. Ghirlanda, A. Celotti and G. Ghisellini, *Astron. Astrophys.* **406**, 879 (2003).
364. V. V. Usov, in *Gamma-Ray Bursts*: *The First Three Minutes*, eds. J. Poutanen and R. Svensson (ASP Conf. Ser. V109, 1999), p. 153.
365. J. C. Wheeler, I. Yi, P. Hoflitch and L. Wang, *Astrophys. J.* **537**, 810 (2000).
366. R. D. Blandford, astro-ph/0202265 (2002).
367. M. Lyutikov and R. Blandford, astro-ph/0210671 (2002).
368. H. C. Spruit and G. Drenkhahn, astro-ph/0302468 (2003).
369. M. Sikora, M.C. Begelman, P. Coppi and D. Proga, *Astrophys. J.*, submitted, astro-ph/0309504 (2003).
370. G. Drenkhahn, *Astron. Astrophys.* **387**, 714 (2002).
371. V. V. Usov, *Mon. Not. R. Astron. Soc.* **267**, 1035 (1994).
372. P. Goldreich and W. H. Julian, *Astrophys. J.* **157**, 869 (1969).
373. M. Lyutikov and E. G. Blackman, *Mon. Not. R. Astron. Soc.* **321**, 177 (2001).
374. C. Thompson, *Mon. Not. R. Astron. Soc.* **270**, 480 (1994).
375. M. V. Smolsky and V. V. Usov, *Astrophys. J.* **531**, 764 (2000).
376. C. D. Dermer and K. E. Mitman, astro-ph/0301340 (2003).
377. E. E. Fenimore, E. Ramirez-Ruiz and B. Wu, *Astrophys. J.* **518**, L73 (1999).
378. M. Böttcher and C. D. Dermer, *Astrophys. J.* **529**, 635 (2000).
379. T. W. Giblin *et al.*, *Astrophys. J.* **570**, 573 (2002).
380. M. Tavani, *Phys. Rev. Lett.* **76**, 3478 (1996).
381. M. Tavani, *Astrophys. J.* **466**, 768 (1996).
382. E. Cohen *et al.*, *Astrophys. J.* **488**, 330 (1997).
383. R. D. Preece *et al.*, *Astrophys. J.* **506**, L23 (1998).
384. M. V. Medvedev, *Astrophys. J.* **540**, 704 (2000).
385. N. M. Lloyd and V. Petrosian, *Astrophys. J.* **543**, 722 (2000).
386. G. Ghisellini, A. Celotti and D. Lazzati, *Mon. Not. R. Astron. Soc.* **313**, L1 (2000).
387. A. M. Bykov and P. Mészáros, *Astrophys. J.* **461**, L37 (1996).
388. E. V. Derishev, V. V. Kocharovsky and Vl. V. Kocharovski, *Astron. Astrophys.* **372**, 1071 (2001).
389. E. Liang *et al.*, *Astrophys. J.* **479**, L35 (1997).
390. E. P. Liang, *Astrophys. J.* **491**, L15 (1997).
391. G. Ghisellini and A. Celotti, *Astrophys. J.* **511**, L93 (1999).
392. D. Lazzati, G. Ghisellini, A. Celotti and M. J. Rees, *Astrophys. J.* **529**, L17 (2000).
393. G. Ghisellini, D. Lazzati, A. Celotti and M. J. Rees, *Mon. Not. R. Astron. Soc.* **316**, L45 (2000).
394. D. Kazanas, M. Geogranopoulos and A. Mastichiadis, *Astrophys. J.* **578**, L15 (2002).
395. E. Nakar, T. Piran and E. Waxman, *J. Cos. Astropart. Phys.* **10**, 5 (2003).
396. N. J. Shaviv and A. Dar, *Astrophys. J.* **447**, 863 (1995).
397. D. Eichler and A. Levinson, *Astrophys. J.* **596**, L147 (2003).
398. D. Lazzati, E. Rossi, G. Ghisellini and M. J. Rees, *Mon. Not. R. Astron. Soc.* **347**, L1 (2004).
399. A. Dar and A. de Rújula, *Mon. Not. R. Astron. Soc.*, submitted, astro-ph/0308248 (2003).
400. E. Ramirez-Ruiz and N. M. Lloyd-Ronning, *New Astron.* **7**, 197 (2002).
401. http://swift.gsfc.nasa.gov/
402. M. G. Baring, astro-ph/9911061 (1999).

403. http://glast.gsfc.nasa.gov/
404. R. Antonucci, *Annu. Rev. Astron. Astrophys.* **31**, 473 (1993).
405. C. M. Urry and P. Padovani, *Pub. Astron. Soc. Pacific* **107**, 803 (1995).
406. V. M. Lipunov, K. A. Postnov and M. E. Prokhorov, *Astron. Rep.* **45**, 236 (2001).
407. Z. G. Dai and L. J. Gou, *Astrophys. J.* **552**, 72 (2001).
408. P. Kumar and J. Granot, *Astrophys. J.* **591**, 1075 (2003).
409. J. Granot and P. Kumar, *Astrophys. J.* **591**, 1086 (2003).
410. A. Panaitescu and P. Kumar, *Astrophys. J.* **592**, 390 (2003).
411. D. M. Wei and Z. P. Jin, *Astron. Astrophys.* **400**, 415 (2003).
412. J. D. Salmonson, *Astrophys. J.* **592**, 1002 (2003).
413. N. M. Lloyd-Ronning, X. Dai and B. Zhang, *Astrophys. J.* **601**, 371 (2004).
414. R. Perna, R. Sari and D. Frail, *Astrophys. J.* **594**, 379 (2003).
415. G. Barbiellini, A. Celotti and F. Longo, *Mon. Not. R. Astron. Soc.* **339**, L17 (2003).
416. J. D. Salmonson, *Astrophys. J.* **546**, L29 (2001).
417. E. Rossi, D. Lazzati, J. D. Salmonson and G. Ghisellini, astro-ph/0211020 (2002).
418. D. Lazzati *et al.*, *Astron. Astrophys.* **410**, 823 (2003).
419. Y. F. Huang, X. F. Wu, Z. G. Dai, H. T. Ma and T. Lu, *Astrophys. J.* **605**, 300 (2004).
420. X. Dai and B. Zhang, to be submitted (2004).
421. E. W. Liang, X. F. Wu and Z. G. Dai, *Mon. Not. R. Astron. Soc.*, submitted (2003).
422. K. S. Cheng, Y. F. Huang and T. Lu, *Mon. Not. R. Astron. Soc.* **325**, 599 (2001).
423. H. T. Ma, Y. F. Huang, Z. G. Dai and T. Lu, *Chin. J. Astron. Astrophys.* **3**, 225 (2003).
424. S. Dado, A. Dar and A. De Rujula, *Astron. Astrophys.* **388**, 1079 (2002).
425. S. Dado, A. Dar and A. De Rujula, *Astron. Astrophys.* **401**, 243 (2003).
426. S. Dado, A. Dar and A. De Rujula, *Astrophys. J.* **572**, L143 (2002).
427. S. Dado, A. Dar and A. De Rujula, *Astrophys. J.* **594**, L89 (2003).
428. D. Fargion and A. Salis, *Astrophys. Space Sci.* **231**, 191 (1995).
429. D. Fargion *Astron. Astrophys. Suppl.* **138**, 507 (1999).
430. D. Fargion, astro-ph/0307348 (2003).
431. T. Totani and A. Panaitescu, *Astrophys. J.* **576**, 120 (2002).
432. A. Levinson *et al.*, *Astrophys. J.* **576**, 923 (2002).
433. E. Nakar, T. Piran and J. Granot, *Astrophys. J.* **579**, 699 (2002).
434. C. D. Dermer, J. Chiang and M. Böttcher, *Astrophys. J.* **513**, 656 (1999).
435. Y. F. Huang, Z. G. Dai and T. Lu, *Mon. Not. R. Astron. Soc.* **332**, 735 (2002).
436. J. E. Rhoads, *Astrophys. J.* **591**, 1097 (2003).
437. E. Woods and A. Loeb, *Astrophys. J.* **523**, 187 (1999).
438. T. Nakamura, *Astrophys. J.* **522**, L101 (1999).
439. R. Yamazaki, K. Ioka and T. Nakamura, *Astrophys. J.* **571**, L31 (2002).
440. Z. P. Jin and D. M. Wei, *Chin. J. Astron. Astrophys.*, submitted, astro-ph/0308061 (2003).
441. S. E. Woosley, R. G. Eastman and B. Schmidt, *Astrophys. J.* **516**, 788 (1999).
442. A. M. Soderberg *et al.*, *Astrophys. J.*, in press, astro-ph/0311050 (2003).
443. M. H. P. M. van Putten and E. C. Ostriker, *Astrophys. J.* **552**, L31 (2001).
444. J. S. Bloom, S. R. Kulkarni and S. G. Djorgovski, *Astron. J.* **123**, 1111 (2002).
445. G. E. Brown *et al.*, *New Astron.* **5**, 191 (2000).
446. H. C. Spruit, *Astron. Astrophys.* **341**, L1 (1999).
447. A. I. MacFadyen, S. E. Woosley and A. Heger, *Astrophys. J.* **550**, 410 (2001).
448. J. C. Wheeler, D. L. Meier and J. R. Wilson, *Astrophys. J.* **568**, 807 (2002).
449. C. D. Matzner, *Mon. Not. R. Astron. Soc.* **345**, 575 (2003).

450. D. Proga, A. I. MacFadyen, P. J. Armitage and M. C. Begelman, *Astrophys. J.* **599**, L5 (2003).
451. A. Königl, astro-ph/0302110 (2003).
452. E. Ramirez-Ruiz, A. Celotti and M. J. Rees, *Mon. Not. R. Astron. Soc.* **337**, 1349 (2002).
453. M. Böttcher, *Astrophys. J.* **539**, 102 (2000).
454. D. Watson *et al.*, *Astrophys. J.* **595**, L29 (2003).
455. P. Kumar and R. Narayan, *Astrophys. J.* **584**, 895 (2003).
456. W. Wang, Y. Zhao and J. H. You, *Astrophys. J.* **576**, L37 (2002).
457. Z. G. Dai, *Astrophys. J.*, in press, astro-ph/0308468 (2003).
458. A. Königl and J. Granot, *Astrophys. J.* **574**, 134 (2002).
459. S. L. Shapiro, *Astrophys. J.* **544**, 397 (2000).
460. M. Böttcher and C. L. Fryer, *Astrophys. J.* **547**, 338 (2001).
461. X. Y. Wang, Z. G. Dai and T. Lu, *Mon. Not. R. Astron. Soc.* **336**, 803 (2002).
462. D. Guetta and J. Granot, *Mon. Not. R. Astron. Soc.* **340**, 115 (2003).
463. S. Inoue, D. Guetta and F. Pacini, *Astrophys. J.* **583**, 379 (2003).
464. J. Granot and D. Guetta, *Phys. Rev. Lett.* **90**, 191102 (2003).
465. D. Guetta and J. Granot, *Phys. Rev. Lett.* **90**, 201103 (2003).
466. S. Razzaque, P. Mészáros and E. Waxman, *Phys. Rev. Lett.* **90**, 241103 (2003).
467. C. D. Dermer and A. Atoyan, *Phys. Rev. Lett.* **91**, 071102 (2003).
468. J. S. Heyl, *Astrophys. J.* **592**, 401 (2003).
469. C. L. Fryer and P. Mészáros, *Astrophys. J.* **588**, L25 (2003).
470. L.-X. Li, *Astrophys. J.* **544**, 375 (2000).
471. R. Ruffini *et al.*, *Astrophys. J.* **555**, L107 (2001).
472. R. Ruffini *et al.*, *Astrophys. J.* **555**, L113 (2001).
473. R. Ruffini *et al.*, *Astrophys. J.* **555**, L117 (2001).
474. I. Yi and E. G. Blackman, *Astrophys. J.* **494**, L163 (1998).
475. E. G. Blackman and I. Yi, *Astrophys. J.* **498**, L31 (1998).
476. W. Kluźniak and M. Ruderman, *Astrophys. J.* **505**, L113 (1998).
477. T. Nakamura, *Prog. Theor. Phys.* **100**, 921 (1998).
478. M. Ruderman, L. Tao and W. Kluźniak, *Astrophys. J.* **542**, 243 (2000).
479. E. Witten, *Phys. Rev. D.* **30**, 272 (1984).
480. C. Alcock, E. Farhi and A. Olinto, *Astrophys. J.* **310**, 261 (1986).
481. K. S. Cheng and Z. G. Dai, *Phys. Rev. Lett.* **77**, 1210 (1996).
482. I. Bombaci and B. Datta, *Astrophys. J.* **530**, L69 (2000).
483. R. Ouyed, J. Dey and M. Dey, *Astron. Astrophys.* **390**, L39 (2002).
484. Z. Silagadze, *Acta Phys. Polonica* **B35**, 881 (2004).
485. X. Y. Wang *et al.*, *Astron. Astrophys.* **357**, 543 (2000).
486. Z. Berezhiani *et al.*, *Astrophys. J.* **586**, 1250 (2003).
487. R. Ouyed and F. Sannino, *Astron. Astrophys.* **387**, 725 (2002).
488. R. Perna and D. Lazzati, *Astrophys. J.* **580**, 261 (2002).
489. D. Frail *et al.*, in preparation (2003).
490. D. E. Reichart and P. A. Price, *Astrophys. J.* **565**, 174 (2002).
491. T. J. Galama and R. A. M. J. Wijers, *Astrophys. J.* **549**, L129 (2002).
492. L. Piro *et al.*, *Astrophys. J.* **558**, 442 (2001).
493. B. P. Venemans and A. W. Blain, *Mon. Not. R. Astron. Soc.* **325**, 1477 (2001).
494. N. Mirabal *et al.*, *Astrophys. J.* **595**, 935 (2003).
495. R. Perna and A. Loeb, *Astrophys. J.* **501**, 467 (1998).
496. E. Fermi, *Phys. Rev.* **75**, 1169 (1949).
497. R. Blandford and D. Eichler, *Phys. Rep.* **154**, 1 (1987).

498. G. Björnsson *et al.*, *Astrophys. J.* **579**, L59 (2002).
499. X. Y. Wang, Z. G. Dai and T. Lu, *Mon. Not. R. Astron. Astrophys.* **319**, 1159 (2000).
500. A. M. Soderberg and E. Ramirez-Ruiz, *Mon. Not. R. Astron. Astrophys.* **330**, L24 (2002).
501. D. Lazzati, E. Ramirez-Ruiz and G. Ghisellini, *Astron. Astrophys.* **379**, L39 (2001).
502. A. Panaitescu, P. Kumar and R. Narayan, *Astrophys. J.* **561**, L171 (2001).
503. S. Rosswog and M. B. Davies, *Mon. Not. R. Astron. Soc.* **334**, 481 (2002).
504. W. H. Lee and E. Ramirez-Ruiz, *Astrophys. J.* **577**, 893 (2002).
505. S. Rosswog, E. Ramirez-Ruiz and M. B. Davies, *Mon. Not. R. Astron. Soc.* **345**, 1077 (2003).
506. J. S. Bloom, S. Sigurdsson and O. R. Pols, *Mon. Not. R. Astron. Soc.* **305**, 763 (1999).
507. K. Belczynski, T. Bulik and V. Kalogera, *Astrophys. J.* **571**, L147 (2002).
508. S. M. Matz *et al.*, *Astrophys. Space Sci.* **231**, 123 (1995).
509. R. M. Kippen *et al.*, *Adv. Space Res.* **22**, 1097 (1998).
510. J. Poirier *et al.*, *Phys. Rev.* **D67**, 042001 (2003).
511. M. Tavani *et al.*, *SPIE — The International Society for Optical Engineering* **4851**, 1151 (2003).
512. T. Weekes, in *High Energy Gamma-Ray Astronomy*, eds. F. Aharonian and H. Völk, H. (AIP Conf. Proc. 558) (Heidelberg, 2000), p. 15.
513. X. Y. Wang, Z. G. Dai and T. Lu, *Astrophys. J.* **556**, 1010 (2001).
514. D. Guetta and J. Granot, *Astrophys. J.* **585**, 885 (2003).
515. P. C. Fragile *et al.*, *Astroparticle Phys.* **20**, 591 (2004).
516. J. P. Rachen and P. Mészáros, *Phys. Rev.* **D58**, 123005 (1998).
517. J. I. Katz, *Astrophys. J.* **432**, L27 (1994).
518. M. Böttcher and C. D. Dermer, *Astroparticle Phys.* **11**, 113 (1999).
519. Z. G. Dai and T. Lu, *Astrophys. J.* **580**, 1013 (2002).
520. O. C. de Jager and F. W. Stecker, *Astrophys. J.* **566**, 738 (2002).
521. R. Plaga, *Nature* **374**, 430 (1995).
522. L. X. Cheng and K. S. Cheng, *Astrophys. J.* **459**, L79 (1996).
523. Z. G. Dai, B. Zhang, L. J. Gou, P. Mészáros and E. Waxman, *Astrophys. J.* **580**, L7 (2002).
524. K. Greisen, *Phys. Rev. Lett.* **16**, 748 (1966).
525. G. T. Zatsepin and V. A. Kuz'min, *Zh. Esks. Teor. Fiz. Pis'ma Red.* **4**, 144 (1966).
526. M. Milgrom and V. Usov, *Astrophys. J.* **449**, L37 (1995).
527. E. Waxman, *Astrophys. J.* **452**, L1 (1995).
528. E. Waxman, *Astrophys. J.*, in press, astro-ph/0210638 (2002).
529. M. Vietri, D. de Marco and D. Guetta, *Astrophys. J.* **592**, 378 (2003).
530. Y. A. Gallant and A. Achterberg, *Mon. Not. R. Astron. Soc.* **305**, L6 (1999).
531. M. Ostrowski and J. Bednarz, *Astron. Astrophys.* **394**, 1141 (2002).
532. F. W. Stecker, *Astropart. Phys.* **14**, 207 (2000).
533. S. T. Scully and F. W. Stecker, *Astropart. Phys.* **16**, 271 (2000).
534. S. Razzaque, P. Mészáros E. Waxman, *Phys. Rev.* **D68**, 083001 (2003).
535. D. Guetta, M. Spada and E. Waxman, *Astrophys. J.* **559**, 101 (2001).
536. E. Waxman and J. N. Bahcall, *Astrophys. J.* **541**, 707 (2000).
537. Z. G. Dai and T. Lu, *Astrophys. J.* **551**, 249 (2000).
538. C. D. Dermer, *Astrophys. J.* **574**, 65 (2002).
539. Z. Li, Z. G. Dai and T. Lu, *Astron. Astrophys.* **396**, 303 (2002).
540. S. Razzaque, P. Mészáros E. Waxman, *Phys. Rev.* **D69**, 023001 (2004).
541. T. J. Weiler *et al.*, astro-ph/9411432 (1994).

542. K. S. Thorne, in *300 Years of Gravitation*, ed. S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987), p. 330.
543. E. S. Phinney, *Astrophys. J.* **380**, L17 (1991).
544. C. T. Cutler *et al.*, *Phys. Rev. Lett.* **70**, 2984 (1993)
545. H. T. Janka *et al.*, *Astrophys. J.* **527**, L39 (1999).
546. M. Rampp, E. Müller and M. Ruffert, *Astron. Astrophys.* **332**, 969 (1998).
547. C. L. Fryer, D. E. Holz and S. A. Hughes, *Astrophys. J.* **565**, 430 (2002).
548. M. B. Davies *et al.*, *Astrophys. J.* **579**, L63 (2002).
549. M. H. P. M. van Putten, *Astrophys. J.* **562**, L51 (2001).
550. M. H. P. M. van Putten, *Astrophys. J.* **575**, L71 (2002).
551. M. H. P. M. van Putten, *Astrophys. J.* **583**, 374 (2003).
552. http://astrogravs.gsfc.nasa.gov/docs/lit/grb_auth.html
553. L. S. Finn, S. D. Mohanty and J. D. Romano, *Phys. Rev.* **D60**, 121101 (1999).
554. E. E. Flanagan and S. A. Hughes, *Phys. Rev.* **D57**, 4535 (1998).
555. L. S. Finn, *AIP Conference Proceedings* **575**, 92 (2001).
556. K. Belczynski, T. Bulik and B. Rudak, *Astrophys. J.* **571**, 394 (2002).
557. L. S. Finn, B. Krishnan and P. J. Sutton, *Astrophys. J.*, in press, astro-ph/0304228 (2003).
558. S. Kobayashi and P. Mészáros, *Astrophys. J.* **585**, L89 (2003)
559. X. Fan, *et al.*, *Astron. J.* **125**, 1649 (2003).
560. D. N. Spergel *et al.*, *Astrophys. J. Suppl.* **148**, 175 (2003).
561. T. Abel, G. L. Bryan and M. L. Norman, *Science* **295**, 93 (2002).
562. V. Bromm, P. S. Coppi and R. B. Larson, *Astrophys. J.* **564**, 23 (2002).
563. A. Loeb, astro-ph/0307231 (2003).
564. R. Barkana and A. Loeb, *Phys. Rep.* **349**, 125 (2001).
565. A. Loeb and R. Barkana, *Annu. Rev. Astron. Astrophys.* **39**, 19 (2001).
566. A. K. Inoue, R. Yamazaki and T. Nakamura, *Astrophys. J.* **601**, 644 (2004).
567. B. E. Schaefer, *Astrophys. J.* **583**, L67 (2003).
568. S. Mao and H. J. Mo, *Astron. Astrophys.* **339**, L1 (1998).
569. J. E. Gunn and B. A. Peterson, *Astrophys. J.* **142**, 1633 (1965).
570. P. Mészáros and M. J. Rees, *Astrophys. J.* **591**, L91 (2003).
571. C. L. Carilli, N. Y. Gnedin and F. Owen, *Astrophys. J.* **577**, 22 (2002).
572. S. R. Furlanetto and A. Loeb, *Astrophys. J.* **579**, 1 (2002).
573. G. Amelino-Camelia, to appear in a special issue, Fundamental physics on the International Space Station and in space, in *General Relativity and Gravitation*, astro-ph/0309174 (2003).
574. S. McBreen *et al.*, *Astron. Astrophys.* **380**, L31 (2001).
575. F. Quilligan *et al.*, *Astron. Astrophys.* **385**, 377 (2002).