Gamma-ray bursts and their links with supernovae and cosmology

Peter Mészáros$^1$ and Neil Gehrels$^2$

$^1$ Center for Particle and Gravitational Astrophysics, Dept. of Astronomy & Astrophysics and Dept. of Physics, Pennsylvania State University, University Park, PA 16802, USA; nnp@astro.psu.edu

$^2$ Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; neil.gehrels@nasa.gov

Received 2012 June 29; accepted 2012 July 4

Abstract Gamma-ray bursts are the most luminous explosions in the Universe, whose origin and mechanism are the focus of intense interest. They appear connected to supernova remnants from massive stars or the merger of their remnants, and their brightness makes them temporarily detectable out to the largest distances yet explored in the universe. After pioneering breakthroughs from space and ground experiments, their study is entering a new phase with observations from the recently launched Fermi satellite, as well as the prospect of detections or limits from large neutrino and gravitational wave detectors. The interplay between such observations and theoretical models of gamma-ray bursts is reviewed, as well as their connections to supernovae and cosmology.

Key words: gamma-ray sources — gamma-ray bursts — cosmic rays — neutrinos — supernovae — cosmology — intergalactic medium

1 INTRODUCTION

Roughly once a day, somewhere within our Hubble horizon a Gamma-ray burst (GRB) occurs, which during the next few seconds or tens of seconds completely overwhelms the gamma-ray flux from the rest of the Universe, including the Sun. In fact, the GRB’s prompt electromagnetic energy output during tens of seconds is comparable to that of the Sun over $\sim \text{few} \times 10^{10}$ years, or to that of our entire Milky Way over a few years; and their X-ray and optical afterglow over the first day after the outburst can outshine the brightest quasars, as well as supernovae (SNe), making them potentially important probes of the distant Universe. Since the discovery of their X-ray afterglows by the Beppo-SAX satellite in 1997 and the subsequent detection of their optical counterparts, we have measured these objects out to the farthest cosmological distances. Thanks to triggers and measurements from the Swift (Gehrels et al. 2004) and Fermi (Michelson et al. 2010; Meegan et al. 2009) satellites, we now have detailed multi-wavelength data for many hundreds of bursts, and redshifts for over 200 of them, and this data set will continue to grow with the continuation of Swift and Fermi, and the possible upcoming SVOM mission (Paul et al. 2011).

GRBs are thought to arise either when a massive star ($\gtrsim 25 M_\odot$) undergoes core collapse, or possibly when a double neutron star or a neutron star and a black hole (BH) binary merges (Woosley & Bloom 2006). The first scenario applies to the so-called long GRBs (LGRBs), whose $\gamma$-ray light-curves last $t_\gamma \gtrsim 2$ s, while the second scenario is the likeliest one so far for short GRBs (SGRBs), whose $\gamma$-ray light-curves last $t_\gamma \lesssim 2$ s (Kouveliotou et al. 1993) (for the latter, this short duration refers to photons at $\epsilon \gtrsim 100$ keV; some “short” bursts, at softer energies, have tails lasting as long...
as 100 s (Gehrels et al. 2009; Vedrenne & Atteia 2009)). In either scenario, it appears inevitable that a compact core object of a few to several solar masses forms, whose radius is on the order of the Schwarzschild radius for this mass, \( r_g \sim 10^8 (M/3M_\odot) \) cm, over a timescale comparable to a few dynamic (free-fall) times, which is likely to be a BH. Accretion of residual infalling gas leads, if the core is fast rotating (guaranteed for a binary), to an accretion disk whose inner radius is \( r_0 \sim 3r_g \sim 10^7 \) cm, and the typical variability timescale of accretion is \( t_0 \sim (2GM/r_0^3)^{-1/2} \sim 10^{-3} \) s. The bulk of the gravitational energy, of order a solar rest mass or \( 10^{54} \) erg is, as in SNe, rapidly radiated as thermal neutrinos (\( E_{\nu,\text{th}} \sim 10 \) MeV), and some amount is radiated as gravitational waves. A smaller fraction, of order \( E_j \sim 10^{51} \) to \( 10^{52} \) erg, is converted into a fireball of equivalent blackbody temperature \( T_0 \sim \) few MeV. This energy eventually emerges as the burst, in the form of a jet. However, for purposes of the dynamics, one can generally use (see below) the isotropic equivalent energy \( E_0 = E_j (4\pi/\Omega_j) \sim 10^{53}E_53 \) erg, which with a nominal total burst duration of \( t_b \sim 10 \) s implies a nominal isotropic equivalent luminosity \( L_\gamma \sim E_0/t_b = 10^{52}L_52 \) erg s\(^{-1}\) (if most of that energy is emitted as \( \gamma \)-rays). The number density of photons at \( r_0 \) is roughly given by \( \nu_0 = 4\pi r_0^2cn_\nu\varepsilon \) where \( \varepsilon \sim kT \sim \) MeV, and the “compactness parameter” (roughly the optical depth of a photon with energy \( \gtrsim m_ec^2 \) against \( \gamma \)\( \gamma \) \( \rightarrow \) e\(^+\)e\(^-\) pair production) is

\[
\tau' \sim \tau_{\gamma\gamma} \sim n_\gamma \sigma_T r_0 \sim \frac{\alpha \sigma_T L_\gamma}{4\pi r_0^2 c \varepsilon} \sim 10^{15},
\]

where \( \sigma_T \) is the Thomson cross section and \( \alpha \) is the fraction of the luminosity above \( m_ec^2 \). This creates a fireball of gamma-rays, electron-positron pairs and hot baryons, where most of the entropy and pressure is in the photons and leptons. The optical depth is huge, and the radiation pressure far exceeds gravity, so the fireball expands and becomes relativistic. A simple lower limit on the expansion bulk Lorentz factor follows from the observations of photons up to \( \gtrsim \) GeV energies, in some bursts. Such photons, trying to escape the source, would collide against softer photons and produce pairs, \( \gamma + \gamma \rightarrow e^+ + e^- \), degrading the spectrum to \( \lesssim 0.5 \) MeV. However the pair production cross section depends on the optical depth to target photons \( \gtrsim m_ec^2 \) against \( \gamma \)\( \gamma \) \( \rightarrow \) e\(^+\)e\(^-\) pair production) is

\[
\varepsilon_1 \varepsilon_2 (1 - \cos \theta) \lesssim 2m_e^2 c^4.
\]

In a relativistically moving jet, causality implies that only photons within angles \( \theta \lesssim 1/\Gamma \) can interact, so with \( \varepsilon_1 \sim 30 \) GeV, \( \varepsilon_2 \sim 1 \) MeV, and \( \cos \theta \sim 1 - \theta^2/2 \), we see that this implies

\[
\Gamma \gtrsim \sqrt{(\varepsilon_1/m_ec^2)(\varepsilon_2/m_ec^2)/2} \gtrsim 10^2.
\]

A more general constraint on \( \Gamma \) is obtained by considering the typical photon spectral distribution in GRBs, which is a broken power law “Band” function \( n(\varepsilon) \propto \varepsilon^{-\beta} \) ph cm\(^{-3}\) MeV\(^{-1}\), where \( \beta \gtrsim 1 \) if \( \beta \approx 2 \) for \( \varepsilon \) below or above a break frequency \( \varepsilon_\nu \) (for \( \beta \approx 2 \), Band et al. 1993; Fishman & Meegan 1995). The \( \gamma \gamma \) optical depth at each energy \( \gtrsim m_ec^2 \) depends on the optical depth to target photons at \( \lesssim m_ec^2 \) satisfying the threshold condition (2), in the jet’s comoving frame. This optical depth can be shown to be \( \propto \Gamma^{-6} \), so for increasingly high \( \Gamma \) the source becomes optically thin to increasingly higher energy photons. The result is that typical GRBs, even if the highest energy photons observed are only 100 MeV, require bulk Lorentz factors in excess of \( \Gamma \sim 250 \) (Baring & Harding 1997).

If the entire burst energy is released impulsively, injecting an energy \( E_0 \) over a timescale \( t_0 \) inside a radius \( r_0 \), with the numbers comparable to those above, the initial entropy per baryon is \( \eta \sim E_0/M_0c^2 \), where \( M_0 \) is the baryon load of the fireball. If the pressure is mainly due to radiation and pairs, the inertia is due to baryons (“baryonic dynamics” regime) and the bulk Lorentz factor initially accelerates as \( \Gamma(r) \sim (r/r_0) \), e.g. Mészáros (2006). After the baryons have become non-relativistic in their own frame, the expansion changes to a coasting behavior at a saturation radius \( r_{\text{sat}} \sim r_0\eta \), and the fireball continues to expand freely with \( \Gamma \sim \eta \sim \text{constant} \). The observationally estimated values are \( \Gamma \sim \eta \sim 10^2 \) to \( 10^3 \), so the baryon load is typically \( 10^{-5} \) to \( 10^{-6}M_\odot \). The behavior is similar if the energy and mass input is spread out over accretion times (i.e. outflow feeding or ejection times) of \( t_b \sim 10 \) to \( 100 \) s, as inferred for a “long” GRB. On the other hand, if
the fireball pressure, or rather stress tensor, is dominated by magnetic fields, the dynamic behavior is different; depending on the symmetries of the fields, the acceleration behavior can range from \( \Gamma \propto r^{1/3} \), e.g., Drenkhahn (2002); Metzger et al. (2011); Mészáros & Rees (2011) to \( \Gamma \propto r^{\eta} \) where \( 1/3 \leq \eta \leq 2/3 \), at least when the outflow is one-dimensional (Komissarov et al. 2009; Narayan et al. 2010). This regime is referred to as magnetically dominated, or Poynting dominated dynamics.

In practice, the outflow is inferred to be jet-like, rather than isotropic, with an average solid angle \( \langle \Omega_j \rangle / 4\pi \sim 1/500 \) or \( \langle \theta_j \rangle \sim 1/30 \) (Frail et al. 2001; Gehrels et al. 2009). In the case of core collapse (“long”) GRBs, this can be due to the outer parts of the star providing a massive barrier, which is best pierced along the centrifugally lightened rotation axis, along which the fireball escapes. The stellar envelope provides a sideways pressure which channels the jet. However, as long as the jet opening angle \( \theta_j \) exceeds \( 1/\Gamma \sim 10^{-2} \) (i.e. \( \gtrsim 0.5^\circ \)), which is generally the case, the expansion occurs as if it were isotropic: causality prevents the gas from having any knowledge of what happens outside an angle \( 1/\Gamma \). For compact binary mergers, the data on jet opening angles are much sparser, but the average value may not be too different (Fong et al. 2012).

In what follows, Section 1 describes the observations, Section 2 discusses the standard picture of the photon spectrum, Section 3 deals with the prompt MeV emission, Sections 3.1 discusses photospheric models and 3.2 discusses magnetic models, Section 4 describes the GeV phenomenology. Section 6 deals with hadronic models, Section 7 with gravitational waves and Section 8 with neutrinos from a GRB, Section 9 discusses progenitors and the supernova connection, Section 10 is about the high redshift GRBs, Section 11 describes the cosmology connection, and Section 12 discusses future prospects.

## 2 Observations

The *Swift* mission, launched in 2004 November, finds bursts and observes the prompt phase with the Burst Alert Telescope (BAT). The afterglow is then observed with the X-Ray Telescope (XRT) and the UV Optical Telescope (UVOT). Measurements of the redshift and studies of host galaxies are typically done with ground-based telescopes which receive immediate alerts from the spacecraft when GRBs are detected. *Swift* has recorded, by far, the largest number of well-localized bursts, afterglow observations and redshift determinations. As of 2012 April 1, BAT has detected 669 GRBs (annual average rate of \( \sim 90 \) per year). Approximately 80% of the BAT-detected GRBs have rapid repointings (the remaining 20% have spacecraft constraints that prevent rapid slewing). Of those, virtually all long bursts observed promptly have detected X-ray afterglow. Short bursts are more likely to have negligible X-ray afterglow, fading rapidly below the XRT sensitivity limit. The fraction of rapid-pointing GRBs that have UVOT detection is \( \sim 35\% \). Combined with ground-based optical observations, \( \sim 60\% \) of *Swift* GRBs have optical afterglow detection. There are so far (mid 2012) about 200 *Swift* GRBs with redshifts, compared with 41 in the pre-Swift era. The redshift distribution of *Swift* GRBs is shown in Figure 1.

![Fig. 1 Redshift distribution and cosmic look-back time of GRBs. The Swift GRBs are in blue, the pre-Swift GRBs in yellow and the co-moving volume of the Universe is the red curve. The GRBs roughly follow the co-moving volume (Gehrels et al. 2009).](image-url)
The *Fermi* mission, launched in 2008 June, has two instruments, the Gamma-ray Burst Monitor (GBM) and the Large Area Telescope (LAT). The GBM has scintillation detectors and covers the energy range from 8 keV to 40 MeV. It measures spectra of GRBs and determines their position to $\sim 5^\circ$ accuracy. The LAT is a pair conversion telescope covering the energy range from 20 MeV to 300 GeV. It measures spectra of sources and positions them to an accuracy of $< 1^\circ$. The GBM detects GRBs at a rate of $\sim 250$ per year, of which on average 20% are short bursts. The LAT detects bursts at a rate of $\sim 8$ per year.

### 3 THE PHOTON SPECTRUM: STANDARD PICTURE

The expansion converts internal energy into bulk kinetic energy, so that the gas cools in its own rest frame and soon becomes an inefficient radiator. In the absence of dissipation, at the photospheric radius $r_{\text{ph}}$ where the flow becomes optically thin to scattering (which for baryon dominated dynamics usually occurs above the saturation radius $r_{\text{sat}}$), the escaping radiation would carry only a small fraction of the burst’s kinetic energy, and might be expected to have a quasi-blackbody spectrum (Paczynski 1986; Shemi & Piran 1990). This motivated the fireball shock model, where the bulk kinetic energy is reconverted by shocks into random particle energy, and hence into non-thermal radiation, at radii beyond the scattering photosphere, where the flow is optically thin; this could be at either the external shocks (Rees & Meszaros 1992; Meszaros & Rees 1993b), where the jet interacts with external (e.g. interstellar) matter, or at internal shocks (Rees & Meszaros 1994) occurring within the jet, at radii intermediate between the photosphere and the external shock radius. These radii can be expressed as

\[
\begin{align*}
    r_{\text{ph}} &\approx (L/4\pi m_p c^3 \eta^3) \sim 4 \times 10^{12} L_{\gamma,52}\eta_{2.5}^{-3} \text{ cm}, \\
    r_{\text{dis}} &\approx \Gamma^2 c t_v \sim 3 \times 10^{13} \eta_{2.5}^2 t_v^{-2} \text{ cm}, \\
    r_{\text{dec}} &\approx (3E_0/4\pi \eta_{\text{ext}} m_p c^2 \eta^2)^{1/3} \sim 2 \times 10^{17} E_{53}^{-1/2} \eta_0^{-1/2} \eta_{2.5}^{-2/3} \text{ cm}. \quad (4)
\end{align*}
\]

Here the photospheric radius assumes baryonic dynamics (Rees & Meszaros 1994, for magnetic dynamics see Mészáros & Rees 2011). The dissipation (internal shock) radius $r_{\text{dis}}$ follows from the relativistic relation between observer time $t$, the radius $r$ and Lorentz factor, $r \approx \Gamma^2 c t$. Considering two shells of matter ejected at time intervals comparable to the variability timescale of ejection $t_v$, with Lorentz factors differing by $\Delta \Gamma \sim \Gamma$, and the deceleration radius $r_{\text{dec}}$ where external shocks start, following from the energy conservation assumption, $E_0 \approx (4\pi/3)r_{\text{dec}}^3 \eta_{\text{ext}} m_p c^2 \Gamma^2$, when the swept up matter has been shock-heated to an energy comparable to the explosion energy (with $\Gamma \sim \eta$, see e.g. Mészáros 2006 for details).

In the (collisionless) shocks, the particles are reheated to thermal energies comparable to the pre-shock kinetic energies per particle. The internal shocks are semi-relativistic (since $\Delta \Gamma \sim \Gamma$), the relative Lorentz factor $\Gamma_{\text{rel}} = (1/2)[(\Gamma_1/\Gamma_2) + (\Gamma_2/\Gamma_1)]$ is semi-relativistic), and this results in a shock luminosity $L_{\text{sh}} = \epsilon_{\text{sh}} L_0$ where $L_0$ is initial kinetic luminosity $L_0 = E_0 \eta / t_v M c^2$ and the shock dissipation efficiency $\epsilon_{\text{sh}} \lesssim 0.1$. Particles repeatedly bounce across the shock, scattered by magnetic irregularities whose energy density can be assumed to build up to some degree of equipartition with the proton’s thermal energy, so the comoving magnetic field is $B^2/8\pi \approx \epsilon_B 4\Gamma_{\text{rel}}^2 m_p c^2$ (primed quantities are in the comoving frame), where $\epsilon_B \leq 1$. The repeated crossing Fermi accelerates the particles according to a relativistic power law (Rees & Meszaros 1992, 1994); the minimum electron (comoving) random Lorentz factor is $\gamma_{e,\text{m}} \sim \epsilon_e (m_p/m_e) \Gamma_{\text{rel}} \gg 1$ (for internal shocks $\Gamma_{\text{rel}} \ll 1$), while for external ones $\Gamma_{\text{rel}} \sim \eta$ and one expects a power law $N(\gamma_e) \propto \gamma_e^{-p}$ above that, with $p \approx 2 - 2.3$. One also expects something similar for the protons in the flow.

i) **Internal shock prompt radiation:** The electrons in the internal shock will emit synchrotron and inverse Compton (IC) radiation, leading to a non-thermal broken power law photon spectrum, roughly similar to the observed “Band” spectra (Meszaros & Rees 1993b; Rees & Meszaros 1994). For reasonable values of $L$, $\eta$, $t_v$, $\epsilon_B$, $\epsilon_e$ the synchrotron peak energy (observer frame) corresponding
to the minimum $\gamma_{e, m}$ is comparable to the observed Band spectral break energies,

$$\varepsilon_{\text{sy, m}} \sim \varepsilon_{\text{br}} \sim 1 \varepsilon_B^{1/2} \varepsilon_e^{3/2} \frac{L_{52}^{1/2}}{\eta_{2, 5}^5 v_{c, -2}} \text{MeV}. \ (5)$$

Equation (5) assumes the randomized kinetic luminosity of internal shocks $L_{\text{sh}}$ to be related to their $\gamma$-ray luminosity through $L_{\gamma} = \varepsilon_{\text{m}} L_{\text{sh}}$, which is true in the fast cooling regime where cooling time is shorter than the dynamic time $t'_{\text{sync}} \ll t'_{\text{dyn}}$ (Sari et al. 1998), which, for internal shocks, is true (Waxman 1997). In this fast cooling regime, for a typical Fermi electron index of $p \simeq 2$ the photon spectral index above $\varepsilon_{\text{br}}$ is expected to be $n(\varepsilon) \propto \varepsilon^{-(p/2) - 1} \propto \varepsilon^{-2} \text{ph cm}^{-3} \text{s}^{-1}$, as is typical for the high energy branch of the canonical Band spectrum. The synchrotron model predicts that below $\varepsilon_{\text{sy, m}}$ a low energy branch $n(\varepsilon) \propto \varepsilon^{-2/3}$, which through superposition of maxima for various parameters could fit the observed Band average low energy branch $n(\varepsilon) \propto \varepsilon^{-1}$ (Meszaros & Rees 1993b) (but flatter spectra are a problem, see below).

ii) External shock radiation and afterglow: At $r_{\text{dec}}$ the relativistic ejecta has used up about half its initial energy in sweeping up an amount $M_{\text{sw}} \sim M_{\text{ejecta}}/\eta$ of external material, driving a forward shock into the external gas and a reverse shock into the ejecta. This occurs at an observer time

$$t_{\text{dec, 0}} \sim \frac{r_{\text{dec}}}{2\eta v^2} \simeq 10 \left( \frac{E_{53}}{n_0} \right)^{1/3} \eta_{2, 5}^{-8/3} \text{ s.} \ (6)$$

The forward shock is initially highly relativistic, $\Gamma_{\text{sh}} \sim \eta$, so $\Gamma_{\text{rel, fs}} \sim \eta$ and the synchrotron spectrum is in the hard X-rays or gamma-rays. The reverse shock builds up slowly and for usual conditions becomes semi-relativistic, $\Gamma_{\text{rel, rs}} \sim 1$ at the deceleration time, when it crosses the ejecta. For this reason its $\gamma_{e, M}$ is smaller than that of the forward shock electrons, and the reverse shock spectrum peaks in the optical or UV (Meszaros & Rees 1993a, 1997a). Beyond $r_{\text{dec}}$ the expansion continues but it is increasingly slowed down due to an increasing amount of swept up matter. In the adiabatic approximation, the bulk Lorentz factor changes from being $\sim \eta$ $\sim$ constant to a power law, with declining behavior given by $E_0 \propto r^3 \Gamma^{-2}$, or

$$\Gamma \sim \eta (r/r_{\text{dec}})^{-3/2}. \ (7)$$

In both the forward and the reverse shock, one again expects Fermi acceleration of electrons according to a power law distribution leading to synchrotron and inverse Compton radiation, but the synchrotron break energy becomes softer in time as the Doppler boost decreases in accordance with Equation (7). This leads to an afterglow (Meszaros & Rees 1997a) progressing from X-rays through optical to radio lasting from minutes to days to months, with fluxes decaying as power laws in time. This prediction was indeed confirmed by observations with the Beppo-SAX satellite in the X-ray (Costa et al. 1997) and with ground telescopes in the optical (van Paradijs et al. 1997) afterglow of GRB 970228, soon followed by the first confirmation of a cosmological redshift (Metzger et al. 1997) and a radio detection (Frail et al. 1997) for GRB 970508. The amount of data on, and understanding of, afterglows has since increased enormously, see e.g. Gehrels et al. (2009); Vedrenne & Atteia (2009).

The external reverse shock gas, most luminous at $t \sim t_{\text{dec}}$, is in pressure equilibrium with forward shock gas, and having a higher particle density and smaller energy per electron than the forward shock, its synchrotron spectrum peaks in the O/UV. This was predicted to lead to an observable prompt optical emission (Meszaros & Rees 1993a, 1997a), later detected with robotic ground telescopes such as ROTSE (Akerlof et al. 1999) triggered by spacecraft detections; the number of such detections is now several dozen (Gehrels et al. 2009).

The external forward shock is expected to also give rise to an IC component, in particular a synchrotron-self-Compton (SSC) from upscattering its own synchrotron photons (Meszaros & Rees 1993a, 1994), which would appear in the GeV range. Such GeV emission has already been detected by EGRET (Hurley et al. 1994), and more recently by the Fermi LAT, e.g. Abdo et al. (2009b). This is discussed in Section 5.
4 PROMPT MEV EMISSION: ISSUES AND DEVELOPMENTS

Issues arise with the radiation efficiency of internal shocks, which is small in the bolometric sense (5%–10%), unless the different shells have widely differing Lorentz factors (Spada et al. 2000; Beloborodov 2000; Kobayashi & Sari 2001). The MeV efficiency is also substantially affected by IC losses (Papathanassiou & Meszaros 1996; Pilla & Loeb 1998), in the BATSE range being typically $\sim 1\%$–$5\%$, both when the MeV break is due to synchrotron (Kumar 1999; Spada et al. 2000; Guetta et al. 2001) and when it is due to inverse Compton (Panaiteescu & Meszaros 2000).

The synchrotron interpretation of the GRB radiation is the most attractive; however, a number of effects can modify the simple synchrotron spectrum. One is that the cooling could be rapid, i.e. when the comoving synchrotron cooling time $t_{\text{sy}}' = 9m_e^3c^5/(4e^4B^2\gamma_c) \sim 7 \times 10^8/B^2\gamma_c$ s is less than the comoving dynamic time $t_{\text{d}yn}' = r/2c\Gamma$, the electrons cool down to $\gamma_c = (6\pi m_ec/\sigma_TB^2t_{\text{d}yn}')$ and the spectrum above $\nu_c \sim \Gamma(3/8\pi)(eB'/m_ec)c^2\gamma_c^2$ is $F_{\nu} \propto \nu^{-1/2}$ (Sari et al. 1998; Ghisellini et al. 2000).

The radiative efficiency issue has motivated investigating various alternatives, e.g. relativistic turbulence in the emission region (Narayan & Kumar 2009; Kumar & Narayan 2009). This assumes that relativistic eddies with Lorentz factors $\gamma_r \sim 10$ exist in the comoving frame of the bulk $\Gamma \gtrsim 300$ flow, and survive to undergo at least $\gamma_r$ changes over a dynamic time, leading both to high variability and better efficiency. Various constraints may however pose difficulties (Lazar et al. 2009), while numerical simulations (Zhang & MacFadyen 2009) indicate that relativistic turbulence would lead to shocks and thermalization, reducing it to the non-relativistic situation.

The synchrotron spectral interpretation faces a problem from the observed distribution of Band low energy spectral indices $\beta_l$ (where $N_c \propto \epsilon^\beta_l$ below the spectral peak), which has a mean value $\beta_l \sim -1$, but for a fraction of bursts this slope reaches positive values $\beta_l > -2/3$ which are incompatible with a low energy synchrotron asymptote $\beta_l = -2/3$ (Preece et al. 2000). Possible explanations include synchrotron self-absorption in the X-ray (Granot et al. 2000) or in the optical range up-scattered to X-rays (Panaiteescu & Meszaros 2000), low-pitch angle scattering or jitter radiation (Medvedev 2000, 2006), or time-dependent acceleration (Lloyd-Ronning & Petrosian 2002), where low-pitch angle diffusion might also explain high energy indices steeper than predicted by isotropic scattering.

Pair formation can become important (Rees & Meszaros 1994; Papathanassiou & Meszaros 1996; Pilla & Loeb 1998) in internal shocks or dissipation regions occurring at small radii, since a high comoving luminosity implies a large comoving compactness parameter $\ell' \gg 1$. Pair-breakdown may cause a continuous rather than an abrupt heating and lead to a self-regulating moderate optical thickness pair plasma at sub-relativistic temperature, suggesting a Comptonized spectrum (Ghisellini et al. 2000). Copious pair formation in internal shocks may in fact extend the photosphere beyond the baryonic photosphere value (4). Generic photosphere plus internal shock models (Mészáros & Rees 2000b; Mészáros et al. 2002; Ryde 2005), including the emission of a thermal photosphere as well as a non-thermal component from internal shocks outside of it, are subject to pair breakdown, which can produce both steep low energy spectra, preferred breaks and a power law at high energies. A moderate to high scattering depth can lead to a Compton equilibrium which gives spectral peaks in the right energy range (Pe’er & Waxman 2004). Pair enrichment of the outflow (due to back-scattered $\gamma\gamma$ interactions) can in general affect both the radiative efficiency and the spectrum (Madau & Thompson 2000; Thompson & Madau 2000; Mészáros et al. 2001; Beloborodov 2002; Thompson 2006).

4.1 Photospheric Models

In the synchrotron interpretation, the observed peak frequency is dependent on the bulk Lorentz factor, which may be random, and since the observed peaks appear to be concentrated near 0.2–1 MeV (Gehrels et al. 2009), the question can be posed whether this is indeed due to synchrotron, or to some other effect. An alternative is to attribute a preferred peak to a blackbody at the comoving pair recombination temperature in the fireball’s photosphere (Eichler & Levinson 2000). In this
case, a steep low energy spectral slope is due to the Rayleigh-Jeans part of the photosphere, and the high energy power law spectra and GeV emission require a separate explanation (Mészáros & Rees 2000b). A related explanation has been invoked (Thompson 1994), considering scattering of photospheric photons off MHD turbulence in the coasting portion of the outflow, which up-scatters the adiabatically cooled photons to the higher observed break energy and forms a power law.

For a photosphere occurring at \( r < r_{\text{sat}} \), which in a baryon-dominated model requires high values of \( \eta \), the radiative luminosity in the observer's frame is undiminished, since \( E_{\text{rad}}' \propto r^{-1} \) but \( \Gamma \propto r \) so \( E_{\text{rad}} \propto \text{constant} \), or \( L_{\text{ph}} \propto r^2 \Gamma^2 T^\gamma \propto \text{constant} \), since \( T' \propto r^{-1} \). However, for more moderate values of \( \eta \) the photosphere occurs at \( r > r_{\text{sat}} \), and whereas the kinetic energy of the baryons is \( E_{\text{kin}} \propto E_0 \propto r^{-1} \) constant the radiation energy drops as \( E_{\text{rad}}' \propto (r/r_{\text{sat}})^{-2/3}, \) or \( L_{\text{ph}} \propto L_0(r_{\text{ph}}/r_{\text{sat}})^{-2/3} \) (Meszaros & Rees 1993b; Mészáros & Rees 2000b). This weakening of the photospheric luminosity leads again to a lowered efficiency, as well as a lower peak energy than observed. However, if the photosphere is dissipative (due to shocks or other dissipation occurring at or below the photosphere) then a high efficiency is regained, and the thermal peak photon energies are in the range of observed Band peaks (Rees & Mészáros 2005). An important aspect is that Compton equilibrium of internal shock electrons or pairs with photospheric photons lead to a high radiative efficiency, as well as to spectra with a break at the right preferred energy and steep low energy slopes (Rees & Mészáros 2005; Pe'er et al. 2005, 2006). It also leads to possible physical explanations for the Amati (Amati et al. 2002) or Ghirlanda (Ghirlanda et al. 2004) relations between spectral peak energy and burst fluence (Rees & Mészáros 2005; Thompson et al. 2007).

4.2 Magnetic Models

An alternative set of models for the prompt emission assume that this is due to magnetic reconnection or dissipation processes, or else to the external shock. Magnetic models fall into two categories, one where baryons are absent or dynamically negligible, at least initially (Usov 1994; Drenkhahn 2002; Drenkhahn & Spruit 2002; Lyutikov & Blandford 2003), and another where the baryon load is significant, although dynamically sub-dominant relative to the magnetic stresses (Thompson 1994; Meszaros et al. 1994; Thompson 2006). These scenarios would in all cases still lead to an external shock, whose radius would again be given by \( r_{\text{dec}} \) in Equation (4), with a standard forward blast wave, but possibly a weaker or absent reverse shock (Meszaros et al. 1994; Meszaros & Rees 1997a), due to the very high Alfvén (sound) speed in the ejecta. For the same reason, internal shocks may be prevented from forming in magnetized outflows. However, this depends on the magnetization parameter \( \sigma \); if not too large, reverse shocks (Zhang & Kobayashi 2005; Giannios et al. 2008; Narayan et al. 2011) or internal shocks might still form as described by Fan et al. (2004), although with different strengths and radiation characteristics. In fact, “internal” dissipation regions may form due to magnetic reconnection, at radii comparable but differing from \( r_{\text{dis}} \) of Equation (4), where electric fields due to reconnection (instead of a Fermi mechanism) lead to particle acceleration, and a high radiative efficiency is conceivable.

A hybrid dissipation model, entitled ICMART (Zhang & Yan 2011), involves a hybrid magnetically dominated outflow leading to semi-relativistic turbulent reconnection. Here a moderately magnetized \( \sigma = (B_r^2 / 4 \pi \rho' c^2) \lesssim 100 \) MHD outflow undergoes internal shocks as \( \sigma \to 1 \), leading to turbulence and reconnection which accelerates electrons at radii \( r \gtrsim 10^{15} \) cm. These involve fewer protons than usual baryonic models, hence they have less conspicuous photospheres and significant variability, and the efficiency and spectrum are argued to have advantages over those in the usual synchrotron internal shock models.

The baryon-free Poynting jet models resemble pulsar wind models, except for being jet-shaped, as in AGN baryon-poor models. The energy requirements of a GRB (isotropic-equivalent luminosities \( L_{\gamma} \gtrsim 10^{52} \) erg s\(^{-1}\)) require magnetic fields at the base in excess of \( B \sim 10^{15} \) G, which can be produced by shear and instabilities in an accreting torus around the BH. The energy source can be either the accretion energy, or via the magnetic coupling between the disk and BH, with extraction of angular momentum from the latter occurring via the Blandford-Znajek mechanism (Blandford
The stresses in this type of model are initially magnetic, also involving pairs and photons, and just as in purely hydro baryon-loaded models they lead to an initial Lorentz factor growth $\Gamma \propto r$ up to a pair annihilation photosphere (Meszaros & Rees 1997b). This provides a first radiation component, typically peaking in the hard X-ray to MeV, with upscattering adding a high energy power law. Internal shocks are not expected beyond this photosphere, but an external shock provides another IC component, which reaches into the GeV-TeV range.

The baryon-loaded magnetically dominated jets have different acceleration dynamics than the baryon-poor magnetic jets or the baryon dominated hydrodynamic jets: whereas both of the latter cases initially accelerate as $\Gamma \propto r$ and eventually achieve a coasting Lorentz factor $\Gamma_f \sim L/\dot{M}c^2$, the baryon-loaded magnetically dominated jets have a variety of possible acceleration behaviors, generally less steep than the above. In the simplest treatment of a homogeneous jet with transverse magnetic field which undergoes reconnection, the acceleration is $\Gamma \propto r^{2/3}$ (Drenkhahn & Spruit 2002; Mészáros & Rees 2011), but in inhomogeneous jets where the magnetic field and the rest mass varies across the jet, the average acceleration ranges from $\Gamma \propto r^{1/3}$ to various other power laws intermediate between this and $\Gamma \propto r$ (McKinney & Uzdensky 2012; Metzger et al. 2011). Few calculations have been made (Giannios & Spruit 2007) of the expected (leptonic) spectral signatures in the simpler magnetized outflow photospheres, typically in a one-zone steady state approximation, showing that a Band-type spectrum can be reproduced.

5 GEV-TEV PHENOMENOLOGY AND MODELS

The first Fermi GRB observations, starting in late 2008, soon yielded a number of surprises. One of the first bright objects showing radically new features was GRB 080916C (Abdo et al. 2009b), in which the GeV emission started only with a second pulse, which was delayed by $\sim 4$ s relative to the first pulse, visible only in MeV (Fig. 2).

The spectra of GRB 080916C consisted of simple Band-type broken power laws, with the first pulse having a soft high energy index disappearing at GeV, but the second and subsequent pulses having higher high energy indices reaching well into the GeV range. There was no evidence for a second spectral component (such as that expected from inverse Compton or hadronic effects). The peak energy of the Band function evolved from soft to hard and back to soft, but in this as well as in other Fermi LAT bursts, the GeV emission persisted in afterglows typically lasting $\gtrsim 1000$ s. On the other hand, in a few bursts, such as GRB 090902B (Abdo et al. 2009a), a second spectral component did indeed appear, at 5$\sigma$ significance, and also a lower energy power law extension whose significance is lower but suggestive. Another burst with a high energy second component was GRB 090926A (Ackermann et al. 2011), with this one showing a clear cut-off or turnover to the high energy power law (Fig. 3).

A significant advance from Fermi LAT was the discovery of the first GeV short burst, GRB 090510 (Ackermann et al. 2010), whose general behavior (including a GeV delay) was qualitatively similar to that of long bursts. Several more short bursts have been discovered subsequently with the Fermi LAT.

The Fermi LAT extended emission, if one ignores various details, has a relatively simple interpretation in terms of conventional forward shock leptoic synchrotron models (i.e. relying on accelerated electrons or $e^+e^-$ pairs) (Ghisellini et al. 2010; Kumar & Barniol Duran 2009). Such models provide a natural delay between an assumed prompt MeV emission (assumed implicitly to come from, e.g. internal shocks or other “inner” mechanisms) and the GeV emission from the external shock, which starts after a few seconds of time delay. However, by more carefully taking into account the constraints provided by the Swift MeV and X-ray observations, and carefully considering the accompanying IC scattering and Klein-Nishina effects, it is clear that at least during the prompt emission, there must be a subtle interplay between the shorter lasting mechanism providing the MeV radiation and the mechanism or emission region responsible for the bulk of the longer lasting GeV radiation (Corsi et al. 2010; De Pasquale et al. 2010; He et al. 2011). One general shortcoming of these early studies was a postponement of addressing the interaction of the GeV emission...
Gamma-ray Bursts and Their Links with Supernovae and Cosmology

Fig. 2 Light curves of GRB 080916C showing the GBM (top two curves) and the LAT (bottom three curves) energy ranges (Abdo et al. 2009b).

A resolution of this problem is possible if the prompt MeV Band spectrum is due to an efficient dissipative photosphere (baryonic, in this case) with an internal shock upscattering the MeV photons at a lower efficiency, giving the delayed GeV spectrum (Toma et al. 2011b). Alternatively, for a magnetically dominated outflow, where internal shocks may not occur, an efficient dissipative photospheric Band spectrum can be up-scattered by the external shock and produce the observed delayed GeV spectrum (Veres & Mészáros 2012). Depending on the parameters, the combined spectrum can look like a two-component or a single Band spectrum (Fig. 4).

On the other hand, a delayed GeV spectrum can also be expected in hadronic models, which assume the co-acceleration, along with the electrons, of protons which undergo electromagnetic cascades and synchrotron losses along with their secondaries (Razzaque 2010; Asano et al. 2009a; Murase et al. 2012; Asano & Mészáros 2012), see Section 6.
The LAT data show that a fraction of GRBs are emitting (in their own rest frame) photons in the energy range of at least $30 - 90$ GeV. A partial list of Fermi LAT detections (Omodei et al. 2011) of maximum observed photon energies and redshifts ($E_{\gamma, \text{obs}}, z$) is $(13.2, 4.35), (7.5, 3.57), (5.3, 0.74), (31.3, 0.90), (33.4, 1.82), (19.6, 2.10), (2.8, 0.897), (4.3, 1.37)$. This list shows that (i) even $z > 4$ bursts can produce $E_{\gamma} > 10$ GeV photons at the observer, and (ii) some $z \sim 1$ bursts can produce $E_{\gamma} > 30$ GeV photons at the observer. This is highly encouraging for the planned large Cherenkov Telescope Array (CTA), as described in recent reviews (Bouvier et al. 2011; Funk & Hinton 2012). The CTA detection rate is estimated (Bouvier et al. 2011) to be $0.7 - 1.6$ per year, based on the rate of Swift triggers (although GBM triggers on Fermi are more frequent, their positional accuracy is poorer). This rate is affected by uncertainties in the fraction of bursts which emit in the GeV range, relative to those emitting below 100 MeV (Guetta et al. 2011; Beniamini et al. 2011). For example, as of 2011 February, in 2.5 years, Fermi LAT detected four bursts at energies $> 10$ GeV (or 20 at $> 0.1$ GeV) out of some 700 bursts detected by Fermi GBM at $E < 100$ MeV. This very small fraction of the total ($\lesssim 1\%$) of course is in part due to the size constraints under which space detectors must operate.

In the standard internal shock model of prompt emission, the intra-source $\gamma\gamma$ absorption typically prevents photons in excess of a few GeV from emerging (Papathanassiou & Meszaros 1996; Pilla & Loeb 1998), unless the bulk Lorentz factor is above $\sim 700$ (Razzaque et al. 2004). For photospheric models of the prompt emission, e.g. (Beloborodov 2010), photons in excess of 10 GeV can escape the source from radii $r_{\gamma\gamma} \sim 10^{15}$ cm, and such radii are also inferred phenomenologically from one-zone analyses of the Fermi data of GRBs. However, most of the GeV emission occurs during the afterglow, which is good for ground-based TeV Cherenkov telescopes, whose reaction time can be slower. Indeed, the GeV emission can last up to $\sim 1000$ s, far longer than the $\sim 2 - 50$ s of the MeV emission. In the standard external shock scenario, the compactness parameter is smaller than in the internal shock, and inverse Compton scattering is expected to lead to multi-GeV and TeV
Fig. 4 A magnetically dominated leptonic model where the MeV Band spectrum is due to photospheric emission, there are no internal shocks, and the external reverse and forward shock upscatter the MeV photospheric spectrum into the GeV range. Parameters are typical for Fermi LAT GRBs, but in some cases lead (top) to a two-component spectrum, while in others (bottom panel) it can be fitted as a single Band spectrum extending to the GeV range (Veres & Mészáros 2012).

photons (Meszaros et al. 1994; Meszaros & Rees 1994), with the details depending on the electron distribution slope and the radiative regime (e.g. slow or fast cooling). This scenario is thought to be responsible for the afterglows of GRBs (Meszaros & Rees 1997a), and is also thought to be responsible for the extended GeV emission observed by LAT so far (Ghisellini et al. 2010; Kumar & Barniol Duran 2009; Wang et al. 2010; Zhang et al. 2011). Of course, propagation in the intergalactic medium from high redshifts leads to additional $\gamma\gamma \rightarrow e^\pm$ interaction with the extragalactic background light or EBL (Coppi & Aharonian 1997; Finke et al. 2010; Primack et al. 2011), the threshold for which depends on the photon energy and the source redshift.

Thus, if TeV emission is produced, it is mainly expected to be detectable from $z \lesssim 0.5$, while the 10–30 GeV emission should be (and is) detectable from higher redshifts. Thus, the GeV detectability is dictated by the source physics, the source rate and the immediate source environment. The source rate, based on MeV observations, is well constrained (Gehrels et al. 2009), while effects from the
near-source environment can be reasonably parametrized (e.g. Gilmore & Ramirez-Ruiz 2010). The source physics, however, has large uncertainties. For example, in an external shock model the simple synchrotron self-Compton (SSC) model can be additionally complicated by the scattering of photons arising at other locations well inside the external shock, e.g. from the photosphere (Toma et al. 2011b), or from an inner region energized by continued central engine activity (Wang et al. 2006). Similar uncertainties about the soft photon source and location would affect hadronic cascade models. The observed GBM high energy spectral slopes are in many cases steep enough not to expect much GeV emission from their extrapolation (Omodei et al. 2011), while in other cases the LAT spectrum shows a cutoff or turnover, e.g. in GRB 090926B (Ackermann et al. 2011). Nonetheless, all things considered, the estimate (Bouvier et al. 2011) of $0.7 ^{-1.6} CTA$ detections per year appears to be a conservative lower limit.

6 HADRONIC MODELS

If GRB jets are baryonic, or magnetically dominated but with a non-negligible baryon load, the charged baryons should be co-accelerated with the electrons in any shocks or reconnection zones, and hadronic processes would lead to both secondary high energy photons and neutrinos. Monte Carlo codes have been developed to model hadronic effects in relativistic flows, including $p, \gamma$ cascades, Bethe-Heitler interactions, etc. For example, Asano et al. (2009a,b) used such a code to calculate the photon spectra from secondary leptons resulting from hadronic interactions following proton acceleration in the same shocks that accelerate primary electrons in GRBs. The code uses an escape probability formulation to compute the emerging spectra in a steady state, and provides a detailed quantification of the signatures of hadronic interactions, which can be compared to those arising from purely leptonic acceleration. Spectral fits of the Fermi LAT observations of the short GRB 090510 were modeled by Asano et al. (2009a) as electron synchrotron for the MeV component and photohadronic cascade radiation for the GeV’s distinct power law component.

Since acceleration as well as cascade development can take some time, in principle even one-zone models might result in GeV-MeV photon delays. For example, Razzaque (2010) assumes for GRB 090510 the prompt MeV to be electron synchrotron and the GeV to be proton synchrotron, whose cooling time cranks down the photon energy into the GeV range on a delay timescale of a few seconds, with the electron plus proton synchrotron merging into a single Band function with the approximate spectral slope of the GeV photons. A more recent one-zone hadronic calculation (Asano & Mészáros 2012) shows that even when proton synchrotron is not important, hadronic cascade development leads to a second GeV component, with time delays comparable to the observed ones (Fig. 5, right panel). Similar delays can, however, be also obtained in purely leptonic two-zone photosphere plus external shock models (Asano & Mészáros 2011) (Fig. 5, left panel).

Hadronic interactions can also have implications for a low energy photon power law below the Band function, perhaps resulting in a GRB optical prompt flash, as discussed by Asano et al. (2010). For the usual Band MeV spectrum produced by conventional leptonic mechanisms, the acceleration of hadrons leads to secondaries whose radiation produces both a high energy “extra” GeV component and a prompt bright optical emission from secondary synchrotron. This might explain, e.g. the observed “naked eye” 5th magnitude flash of GRB 080319B, e.g. Racusin et al. (2008).

Hadronic binary collisions in baryon-loaded jets can also be important, both for efficient energy dissipation and for shaping the photon spectrum. This is because the baryons will consist of both protons ($p$) and neutrons ($n$), especially if heavy elements are photo-dissociated. The protons are coupled to the radiation during the acceleration phase but the neutrons are carried along only thanks to nuclear $(p, n)$ elastic collisions, whose characteristic timescale at some point becomes longer than the expansion time. At this point the $p$ and $n$ radial relative drift velocity $v$ approaches $c$, leading to the collisions becoming inelastic, $p + n \rightarrow \pi^+, \pi^0$, in turn leading to positrons, gammarays and neutrinos (Bahcall & Mészáros 2000). Such inelastic $(p, n)$ collisions can also arise in jets where the bulk Lorentz factor is transversely inhomogeneous (Mészáros & Rees 2000a), e.g. going from large to small as the angle increases, as expected intuitively from a jet experiencing friction against the surrounding stellar envelope. In such cases, the neutrons from the slower, outer
jet regions can diffuse into the faster inner regions, leading to inelastic \((p, n)\) and \((n, n)\) collisions resulting again in pions. An interesting consequence of either radial or tangential \((n, p)\) drifts is that the decoupling generally occurs below the scattering photosphere, and the resulting positrons and gamma-rays deposit a significant fraction of the relative kinetic energy into the flow, reheating it (Beloborodov 2010). Internal dissipation below the photosphere has been advocated, e.g. Rees & Mészáros (2005), to explain the MeV peaks as quasi-thermal photospheric peaks (Ryde et al. 2011; Pe’er 2011), while having a large radiative efficiency. Such internal dissipation is naturally provided by \((p, n)\) decoupling, and numerical simulations (Beloborodov 2010) indicate that a Band spectrum and a high efficiency is indeed obtained, which remains the case even when the flow is magnetized up to \(\varepsilon_B = 2\) (Vurm et al. 2011), while keeping the dynamics dominated by the baryons. These numerical results were obtained for nominal cases based on a specific radial \((n, p)\) velocity difference, although the phenomenon is generic.

The photon spectral signatures of a magnetically dominated, baryon loaded leptonic plus hadronic GRB model involving nuclear collisions have been calculated by Mészáros & Rees (2011). This uses a realistic transverse structure of a fast core-slow sheath. The analytical results indicate that the transverse neutron collisions become most effective, resulting in GeV photons at radii from which the observer-frame time delay relative to the photospheric MeV photons is appropriate to explain the observed Fermi time lags. The purely leptonic (SSC, EIC) time delays and spectral components of such a baryon-loaded magnetic model, in the absence of drifts and transverse gradients, have been calculated by Bošnjak & Kumar (2012), leading to delays in the range observed by Fermi.

A hadronic model which attempts to self consistently produce the GeV radiation, the MeV Band spectrum and the low energy (optical) power law is discussed in Murase et al. (2012). The protons accelerated in the shocks or magnetic reconnection regions result in hadronic cascades which produce, as in Asano et al. (2010), the GeV and optical power laws, while the cooled leptonic secondaries are re-accelerated via a Fermi second order mechanism in the turbulent MHD waves produced by the same shocks or reconnection regions, leading self-consistently to an MeV Band spectrum.

7 GRAVITATIONAL WAVES FROM GRBS

GRBs may also be sources of gravitational waves (GWs). The most likely such sources are SGRBs (Centrella 2011), if these indeed arise from merging compact objects (Gehrels et al. 2009). The rates
in advanced LIGO and VIRGO may be at least several per year (Leonor et al. 2009). LGRBs are more speculative as sources, since in the favored core collapse scenario the collapse may be more chaotic (Fryer et al. 2002). They may, nonetheless, be weakly detectable as GW sources, especially if the core collapse breaks up into substantial blobs (Kobayashi & Mészáros 2003), or if they go through a magnetar phase leading to a bar (Corsi & Mészáros 2009). More recent, detailed numerical calculations of collapsar (long) GRBs lead to GW prospects which range from pessimistic (Ott et al. 2011) to modest (Kiuchi et al. 2011).

8 HIGH ENERGY NEUTRINOS FROM GRBS

High energy ($10^9\,\text{eV} \lesssim E_{\nu} \lesssim 10^{18}\,\text{eV}$) neutrinos (HENUs) may be expected from baryon-loaded GRBs if sufficient protons are co-accelerated in the shocks (Waxman 2011). The most widely considered paradigm involves proton acceleration and $p\gamma$ interactions in internal shocks, resulting in prompt $\sim 100\,\text{TeV}$ HENUs (Waxman & Bahcall 1997; Murase & Nagataki 2006). Other interaction regions considered are external shocks, with $p\gamma$ interactions on reverse shock UV photons leading to EeV HENUs (Waxman & Bahcall 2000); and pre-emerging or choked jets in collapsars resulting in HENU precursors (Mészáros & Waxman 2001). Also, for baryonic dominated GRBs, a neutrino component may arise from photospheric $p\gamma$ and $pp$ interactions (Murase 2008; Wang & Dai 2009).

An EeV neutrino flux is also expected from external shocks in very massive Pop. III magnetically dominated GRBs (Gao et al. 2011). Current IceCube observations (Ahlers et al. 2011; Abbasi et al. 2011, 2012) are putting significant constraints on the internal shock neutrino emission model, with data from the full array still needing to be analyzed. One caveat is that, since the above analysis, several groups (Hümmer et al. 2012; Li 2012; He et al. 2012) have recalculated the GRB internal shock neutrino production in greater detail, including multi-pion and Kaon production in the $p\gamma$ interactions, and allowing for various astrophysical uncertainties including different values of the Lorentz factor and the accelerated proton to electron ratio $L_p/L_e = 1/f_e$. The conclusion from these revised calculations is that the current IceCube (IC40+IC59) measurements need to be extended for another four to nine years to obtain a strong constraint.

9 PROGENITORS AND THE SUPERNOVA CONNECTION

Including the collimation correction, the GRB electromagnetic emission is energetically quite compatible with an origin in, say, either compact mergers of neutron star-neutron star (NS-NS) or black hole-neutron star (BH-NS) binaries (Paczynski 1986; Eichler et al. 1989; Narayan et al. 1992; Meszaros & Rees 1992), or with a core collapse (hypernova or collapsar) model of a massive stellar progenitor (Woosley 1993; Paczynski 1998a; MacFadyen & Woosley 1999), which would be related to but much rarer than core-collapse SNe (Woosley & Bloom 2006; Gehrels et al. 2009). While in both scenarios the outcome could be, at least temporarily, a massive fast-rotating ultra-high magnetic field neutron star (a magnetar) (Wheeler et al. 2000; Metzger et al. 2011), the mass of the resulting central object substantially exceeds the Chandrasekhar mass and is expected to lead, sooner or later, to the formation of a central BH. The latter will be fed through a (seconds to minutes) accretion episode from the surrounding disrupted core of stellar matter, which provides the energy source for the ejection of relativistic matter responsible for the radiation. This inference appears to be confirmed by numerical simulations, for NS-NS or NS-BH mergers (Ruffert & Janka 1999; Rosswog 2005; Rezzolla et al. 2011) as well as for collapsar models (MacFadyen & Woosley 1999; MacFadyen et al. 2001; Zhang et al. 2004).

The above numerical simulations also indicate that (1) the compact object merger accretion disks are less massive and the accretion episode (when the disk is not highly magnetized) lasts less than a few seconds, compatible with the observations of the canonical “short” GRBs (SGRBs); and (2) in the collapsar models the accretion lasts for tens of seconds or more, compatible with the durations of canonical “long” GRBs (LGRBs).

The observations of LGRBs indicate that they are generally located in active star-forming environments, usually in blue, small or not too massive, gas-rich galaxies (Paczynski 1998a; Woosley &
Bloom 2006; Gehrels et al. 2009), which is where one expects massive stars to be present. Progenitor stars more massive than \( 25 - 28 M_\odot \), following core collapse, are expected to result in a BH central remnant, either directly or through an intermediate neutron star (NS) phase (Fryer et al. 1999; Fryer 2006). Such BH core collapse events, if the core is rotating sufficiently fast, can lead to a fall-back fed accretion disk powering a relativistic jet, which is able to escape a star of \( \sim 10^{11} \) cm (MacFadyen et al. 2001; Zhang et al. 2004; Woosley 2011). This radius corresponds to those of Wolf-Rayet (WR) stars, which are thought to arise from more massive \( M > 25 M_\odot \) progenitors, whittled down by wind mass loss prior to core collapse. The high rotation rate, which favors the wind mass loss and also the formation of a longer lasting accretion disk, is thought to be enhanced when the star arises in a metal-poor environment (Woosley & Bloom 2006), which in fact seems to characterize many LGRB host galaxies (Stanek et al. 2006; Gehrels et al. 2009).

The massive core collapse model of LGRBs is confirmed by the fact that LGRBs are, in some cases, demonstrably associated with type Ib/c SNe, whose explosion is, to within errors, contemporaneous with the GRB (Galama et al. 1998; Hjorth et al. 2003; Della Valle 2011; Hjorth & Bloom 2011). The SNe Ib/c are generally thought to have WR progenitors, whether they are associated with GRBs or not; only a small fraction of order a few percent of SNe Ib/c appear to be associated with LGRBs (Soderberg et al. 2010; Hjorth & Bloom 2011). However, the SNe Ib/c associated with GRBs, as well as a good fraction of those not associated, are classified as hypernovae (HNe), meaning that they have unusually broad spectral lines indicating semi-relativistic envelope ejecta (\( v/c \sim 1 \)), and inferred isotropic energies \( E_{HN,iso} \sim 10^{52-5} \) erg, as opposed to the average SNe with \( v/c \sim 0.1 \) and \( E_{SN,iso} \sim 10^{51} \) erg (Paczynski 1998b; Waxman & Loeb 1999; Nomoto et al. 2010; Thielemann et al. 2011; Lazzati et al. 2012).

For short gamma-ray bursts (SGRB) the most widely favored candidates are mergers of neutron star binaries (DNS) or NS-BH binaries, which lose orbital angular momentum by gravitational wave radiation and undergo a merger. This second progenitor scenario has only now begun to be tested thanks to the Swift detection of short burst afterglows (Gehrels et al. 2005; Berger et al. 2005, 2006; Nakar 2007; Berger et al. 2007). The SGRBs are found both in evolved (elliptical) galaxies (Gehrels et al. 2005) and in galaxies with star formation (Gehrels et al. 2009; Berger 2011), in proportions compatible with that expected for an old population such as NSs. While NSs are expected, and found, in young star-forming galaxies, massive young stars are not expected, and not found in old population ellipticals. Indeed, no SN has ever been found exploding at the same time and location as an SGRB (Gehrels et al. 2009). Of course, NSs are the product of SNe, which could have occurred much earlier than the burst, from progenitors whose initial mass was \( 8 M_\odot \lesssim M_\star \lesssim 25 M_\odot \). A kick is imparted to the NS at the SN event, so the NS can wander off significant distances from its birth site (many \( \text{kpc} \), even outside the host galaxy). If the NS was born in a binary and/or later became a binary, the time between the initial explosion and the eventual merger can range up to \( 10^3 - 10^5 \) years, and is very unlikely to be less than \( 10^9 \) years (Coward et al. 2012; Kiel et al. 2010). Only a handful of SGRBs have yielded reliable lightcurve breaks suitable for determining a jet opening angle. The latest measurements and comparison to previous data (Fong et al. 2012) indicate an average \( \theta_j \sim 5^\circ \) (comparable to the LGRB average value, although there is one outlier at \( 25^\circ \)). This is interesting, since for DNS mergers there is no stellar envelope (as for LGRBs) to provide collimation, at most there would be a wind; however, such narrow jets would be compatible with twisted or hoop-stress collimated MHD jets from DNS mergers (Rezzolla et al. 2011; Shibata et al. 2011).

10 HIGH REDSHIFT GRBS

Long GRBs are astonishingly bright, both in gamma-rays and at longer wavelengths. In the optical, typical brightnesses are \( \sim 18 \)th magnitude a few hours after the trigger (and some have been detected in the 5th-10th magnitude range seconds after the trigger), while a Milky-Way-type galaxy has \( \sim 32 \)nd magnitude at a redshift \( z = 8 \). In fact, GRBs vie with galaxies for the record of the highest confirmed redshift measurements, e.g. GRB 080913 at \( z = 6.7 \) (Greiner et al. 2009), GRB 090423 at \( z = 8.2 \) (Tanvir et al. 2009; Salvaterra et al. 2009) (through spectroscopy), while GRB 090429B has a photometric redshift of \( z \approx 9.4 \) (Cucchiara et al. 2011). It is possible that even much more distant
objects than these have already been detected in the gamma-ray and X-ray detectors of Swift and Fermi, although for such $z > 9$ objects a specific (optical/IR or other) redshift signature is extremely difficult and noisy, so redshift diagnostics are increasingly harder to obtain in this range. The above discoveries do, however, indicate that the prospect of eventually reaching into the realm of Pop. III objects is becoming increasingly realistic.

Population III GRBs at $10 \lesssim z \lesssim 20$ may result, as they do at lower redshifts, from massive, $M > 25 - 30 M_\odot$ metal-poor stars whose core collapses to a black hole (Bromm & Loeb 2006). However, the mass of Pop. III stellar progenitors could be as high as $\sim 1000 M_\odot$, leading to $100 - 500 M_\odot$ black holes (Komissarov & Barkov 2010), although the Pop. III masses are a subject of debate (and could be much lower (Stacy et al. 2009; Norman 2010). For extremely massive black holes, the jets are likely to be Poynting-dominated, e.g. powered by the Blandford-Znajek mechanism. The expansion dynamics and the radiation arising from such very massive Poynting jet GRBs was discussed by Mészáros & Rees (2010). At typical redshifts $z \sim 20$ this implies a “prompt” emission extending to $\lesssim 1$ day which should be detectable by Swift or Fermi, being most prominent initially around $50 \text{ keV}$ due to the jet pair photosphere, followed after a similar time interval by an external shock synchrotron component at a few $\text{keV}$ and an inverse Compton component at $\gtrsim 70 \text{ GeV}$ (Toma et al. 2011a). Both the ‘prompt’ emission and the longer-lasting afterglows (Toma et al. 2011a) of such Pop. III GRBs should be detectable with the BAT or XRT on Swift or the GBM on Fermi. On Swift, image triggers may be the best way to detect them, and some constraints on their rate are provided by radio surveys. They are expected to have GeV extensions as well, but redshift determinations need to rely on $L$-band or $K$-band spectroscopy.

11 THE COSMOLOGY CONNECTION

Since GRBs are seen out to the largest redshifts yet measured, and for periods of hours to days they can outshine any other objects at those distances, their potential usefulness as tools for cosmology has been intriguing for some time.

The simplest way, to use them as distance markers, is unfortunately not straightforward. This is because they are not good “standard candles,” which could be used e.g. in a Hubble-type diagram to plot flux against redshift to compare against cosmological models to deduce a closure parameter or an acceleration rate. Even the collimation-corrected average gamma-ray energy $E_{\gamma,j} \sim 10^{51} \text{ erg}$ has too much variance to be of direct use as a standard candle. The error is still almost twice as large as that obtained from SN Ia, at least at redshifts $z \lesssim 1.5$, which is the most important region for dark energy studies. The hope is, still, that one or more of the various empirical correlations between observed spectral quantities could lead to a calibration, as the Phillips relation does for SN Ia, which could turn GRBs into an effective distance ruler.

One such empirical correlation is between the photon spectral peak energy $E_{\text{pk}}$ and the apparent isotropic energy $E_{\text{iso}}$ (Amati et al. 2002; Amati 2006), namely $E_{\text{pk}} \propto E_{\text{iso}}^{\alpha}$, with $\alpha \sim 1/2$ (Amati relation). Other correlations are between $E_{\text{pk}}$ and the peak luminosity $L_{\text{pk}}$ (Yonetoku et al. 2004) (Yonetoku relation); or for bursts where the jet opening angle is known, between $E_{\text{pk}}$ and the collimation corrected gamma-ray energy of the jet $E_{\gamma,j}$ (Ghirlanda et al. 2004; Ghirlanda 2007) (Ghirlanda relation); or between $E_{\text{pk}}$ and $L_{\gamma,j}$ (Dai et al. 2004); or between $E_{\text{pk}}$, $E_{\text{iso}}$, and the light curve break time $t_{br}$ (Liang & Zhang 2006) (Liang-Zhang relation); or between the X-ray luminosity at break time $L_{X,br}$ and the break time $t_{br}$ (Cardone et al. 2010). Of course, these correlations are of interest in themselves as possible constraints on the radiation mechanism or emission region, and various interpretations have been made, e.g. Zhang & Mészáros (2002); Rees & Mészáros (2005); Thompson et al. (2007); Ghirlanda et al. (2012a). However, for cosmology only the tightness of the empirical correlation is what matters. Observational biases could, of course, pose problems (Nakar & Piran 2005; Ghirlanda et al. 2012b), and circularity issues may be a concern; the latter, however can in principle be minimized by restricting oneself only to directly observed quantities (Graziani 2011). Some recent papers using GRBs in Hubble diagrams are, e.g. Liang et al. (2010) and Demianski & Pedipalumbo (2011), the latter including 109 GRBs with known redshifts calibrated with 567 SN Ia. While strongly suggestive, and increasingly interesting, the sample is still relatively small compared
to SN Ia, and the dispersion remains larger than for SN Ia, so the usefulness of GRBs as statistical distance indicators remains to be seen.

GRBs, however, are likely to be unique as beacons for probing the high redshift Universe. They are detectable with current gamma-ray, X-ray and infra-red detectors out to distances corresponding to the earliest star formation epoch (Lamb & Reichart 2000; Ciardi & Loeb 2000), and they may provide possible redshift signatures (Mészáros & Rees 2003; Gou et al. 2004) extending into the $10 < z < 20$ range. Their strong, featureless power law continuum spectrum shining through the intergalactic medium and intervening young galaxies or protogalaxies provide a sensitive probe of the ionization state, velocity distribution and chemical composition at those redshifts (Loeb & Barkana 2001; Savaglio 2010; Hartmann 2010). An example is shown in Figure 6, indicating the change of the metallicity (given by the oxygen to hydrogen abundance ratio) as a function of redshift. These abundances are determined from spectral absorption lines in the continuum radiation of GRBs and quasars. The GRB lines come mainly from the host galaxy gas in the star forming region where the explosion occurred, while the quasar lines arise in random intervening galaxies along the line of sight. One sees that GRBs provide information out to higher redshifts than quasars, and indicate systematically higher metallicities. This is expected, since the GRBs are sampling gas in the star forming regions, which have been enriched by nucleosynthesis and SN explosions.

GRBs also provide an excellent tool for investigating the cosmic star formation rate (SFR) of the high redshift Universe, and thereby also the rate of large scale structure (LSS) formation, out to (so far) redshifts in the 8–10 range. This is exemplified in Figure 7, which shows the star formation rate determined through various techniques. The LGRBs are the endpoints of the lives of massive stars, and their rate is therefore approximately proportional to the star formation rate in general. However, at high redshifts the rates are very uncertain, and may be subject to various observational biases. There may also be evolutionary biases, such as a dependence of the LGRB formation on the metallicity of the host galaxy, which needs to be taken into account, e.g. Savaglio (2010); Hartmann (2010).

The most distant GRBs may also provide the only possible probes of the era from the first generation of (Population III) stars formed in the Universe (Komissarov & Barkov 2010; Mészáros...
Fig. 7 Cosmic star formation history (Kistler et al. 2009). Shown are the data compiled in Hopkins & Beacom (2006) (light circles) and from Ly-α emitters (LAE) (Ota et al. 2008). Also recent LBG data for two UV LF integrations: down to $0.2L_\odot$ (down triangles, Bouwens et al. 2008), and complete up to $z = 3$ (up triangles). Swift GRB-inferred rates are diamonds, and the shaded band shows the range resulting from varying evolutionary parameters. Also shown is the critical $\rho_\ast$ from Madau et al. (1999) for $C/f_{esc} = 40, 30, 20$ (dashed lines, top to bottom), see also Robertson & Ellis (2012).

& Rees 2010; Toma et al. 2011a; Campisi et al. 2011). These relics of the infant Universe could be the most sensitive probes of the redshift for the start of large scale structure formation, with significant implications for the properties of the dark matter.

12 THE FUTURE

Both Swift and Fermi are likely to be functional and return GRB data for many years to come. They have orbital lifetimes extending beyond 2025 and no critical expendables. The SVOM mission (Paul et al. 2011) is an approved Chinese-French mission to observe GRBs and their afterglow. It has a wide-field instrument to image the bursts and one to study the spectrum. The spacecraft rapidly slews like Swift to point toward X-ray and optical telescopes for afterglow observation. There are other concepts in consideration, such as Lobster (Gehrels et al. 2012), which performs the wide-field observations in the X-ray band suitable for high-redshift GRBs. In addition, combined with such electromagnetic detection programs, increasingly sensitive multi-messenger detection attempts will continue to be pursued using high energy neutrinos (Sect. 8) and gravitational waves (Sect. 7). These expanded observational efforts will require more detailed theoretical interpretation and models, extending well beyond what has been achieved so far. Based on past experience, GRBs may be counted on to provide further exciting surprises during the next decade.

Acknowledgements We are grateful to NASA for support, and to our colleagues for many useful discussions.

References

Ahlers, M., Gonzalez-Garcia, M. C., & Halzen, F. 2011, Astroparticle Physics, 35, 87
Asano, K., & Mészáros, P. 2012, arXiv:1206.0347
Centrella, J. 2011, in American Institute of Physics Conference Series, 1381, eds. F. A. Aharonian,
W. Hofmann, & F. M. Rieger (Washington: AIP), 98
Mészáros, P. 2006, Reports on Progress in Physics, 69, 2259
Michelson, P. F., Atwood, W. B., & Ritz, S. 2010, Reports on Progress in Physics, 73, 074901
Murase, K., & Nagataki, S. 2006, Phys. Rev. D, 73, 063002
Omodei, N. & the Fermi LAT Collaboration, 2011, Talk at Fermi Meeting (Stanford U.)