

The job of a scientist to understand the explosion physics of GRBs is like that of a detective to figure out a crime scene. First of all, one needs to collect as much observational data as possible (similar to footprints, fingerprints, or other evidence a detective may collect from the crime scene). Then one can apply known physical laws and logical reasoning to judge what underlying physical processes might be in operation at the GRB site (quite analogous to what a detective does to picture a crime scene based on the available information). While occasionally a “smoking gun” may be discovered, which carries convincing information to settle a case, most of the time one can only infer the most plausible scenario based on the available information at hand. Before starting to discuss physical models, it is therefore essential to summarize all the available observational clues about GRBs. These are the topics of this chapter. The observational properties of prompt emission and afterglow are summarized in §2.1 and §2.2, respectively; §2.3 discusses the association of a supernova (kilonova) with a long (short) GRB, and §2.4 summarizes the properties of the host galaxies of long/short GRBs. Some global properties of GRBs are presented in §2.5, and some widely discussed empirical correlations of GRB properties are summarized in §2.6. In §2.7, various GRB classification schemes are reviewed.

2.1 Prompt Emission

2.1.1 Definitions

Observationally, the *prompt emission* phase of a GRB is conventionally defined as *the temporal phase during which excessive sub-MeV emission is detected by the GRB triggering detectors above the instrumental background emission level*.

Quantitatively, the duration of a burst is usually defined by the so-called “ T_{90} ”: the time interval between the epochs when 5% and 95% of the total fluence is collected by the detector. Figure 2.1 gives an illustration of how T_{90} is measured. Another less-used parameter is T_{50} : the time interval between the epochs when 25% and 75% of the total fluence is collected.

Such an observation-based definition has some limitations: first, it depends on the energy bandpass of the detector. Since GRB pulses are typically wider at lower energies

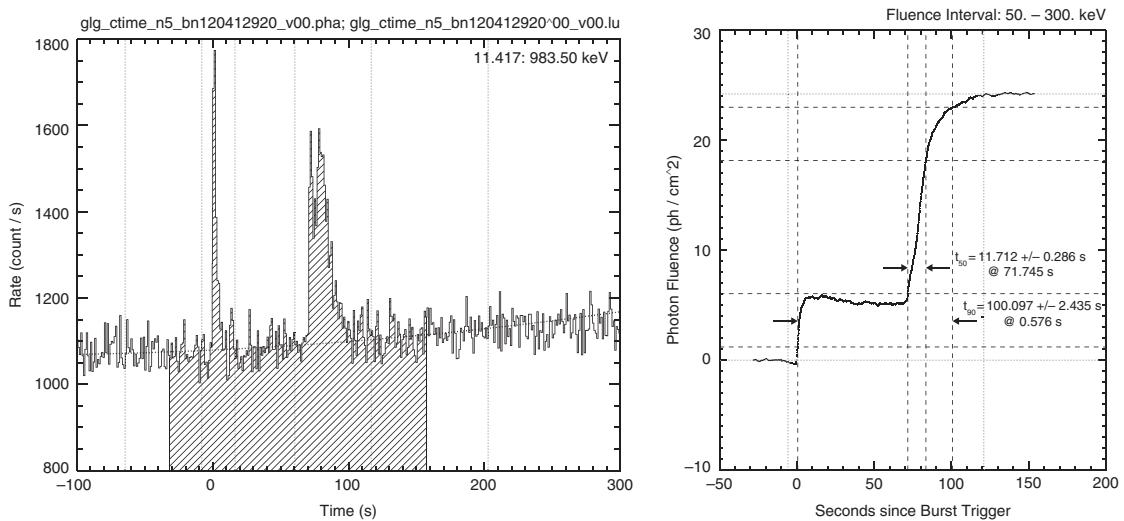


Figure 2.1 An illustration on how T_{90} is defined. *Left:* Lightcurve of GRB 120412A with 1.024 s temporal resolution as detected in *Fermi*/GRB Nal detector 5. Vertical dotted lines define three regions selected for fitting the background. The hatching defines the region selected for duration analysis. *Right:* The determination of T_{90} and T_{50} for GRB 120412A which shows two emission episodes separated by a long quiescent period. Horizontal dotted lines are drawn at 5%, 25%, 75%, and 95% of the total fluence. Vertical dotted lines are drawn at the times corresponding to those fluences, which are used to define the T_{50} and T_{90} intervals. Reproduced from Figures 2 and 3 in von Kienlin et al. (2014) with permission. ©AAS.

(Eq. (2.4)), a detector with a lower energy bandpass would get a longer T_{90} for the same GRB. Second, it is sensitivity dependent. Since the “signals” are defined above the background noise level, a more sensitive detector (e.g. due to a larger collection area) would have a lower background flux level to allow more signal to show up, and hence record a longer T_{90} for the same burst. Finally, some GRBs have clearly separated emission episodes with long quiescent gaps in between. The parameter T_{90} therefore may overestimate the duration of the GRB central engine activity in these cases.

Physically, theorists like to differentiate GRB prompt emission and afterglow based on the physical location where the γ -ray photons are emitted. Prompt emission is related to jet emission from an “internal” site, where the ejecta dissipates energy internally (e.g. through internal shocks or internally triggered magnetic dissipation). Afterglow, on the other hand, is usually considered to be emission from the external shocks (especially the forward shock) due to ejecta–medium interaction. Practically, it is not always easy to differentiate internal emission from external emission. For most GRBs with erratic variability, emission during T_{90} is consistent with having an internal origin.

More generally, one may broadly define “prompt emission” as emission in all wavelengths during the epoch when sub-MeV prompt emission is detected. There have been several GRBs whose prompt GeV, X-ray, and optical emission components were detected during the sub-MeV emission phase. Their temporal profiles generally track that of the sub-MeV emission, but sometimes with noticeable differences (e.g. with spectral lags or offsets). In any case, the rough tracking pattern and the observed rapid variability suggest that most prompt emission in other wavelengths is also of an internal origin.

2.1.2 Temporal Properties

Duration Distribution

The duration T_{90} ranges from milliseconds to thousands of seconds. As detected by BATSE (25–350 keV), the T_{90} distribution includes at least two Gaussian components in logarithmic space with a separation line around 2 seconds in the observer frame (Kouveliotou et al., 1993): a *long-duration class* with T_{90} peaking at 20–30 s, and a *short-duration class* with T_{90} peaking at 0.2–0.3 s. The relative significance of the two components and the peak duration values are energy and sensitivity dependent (e.g. Kouveliotou et al., 1993; Sakamoto et al., 2008b, 2011; Paciesas et al., 2012; Zhang et al., 2012d; Qin et al., 2013). For example, while the short-to-long number ratio in the BATSE data is about 1:3, that in the *Swift* data is only about 1:9 (Sakamoto et al., 2008b, 2011). Qin et al. (2013) showed that when breaking the *Fermi* bandpass into different sub-bands that are consistent with the bandpasses of the previous detectors, similar T_{90} distributions to previous detectors can be reproduced (Fig. 2.2).

Statistically, the long-duration class is on average softer than the short-duration class. The hardness of a burst is usually denoted by its *hardness ratio* (HR), which is the photon count ratio in two fixed observational energy bands. In the T_{90} –HR plane (Fig. 2.3), one can clearly see that the two classes cluster in different regions, with long GRBs typically softer than short GRBs. As a result, the two classes of GRBs are also termed “long/soft” and “short/hard”, respectively. Multi-wavelength observations in the *Swift* era suggest that these two classes of GRBs generally correspond to two physically distinct classes with different progenitor systems, even though the separation between the two physical classes is not unambiguously defined by duration.

Several authors suggested that the T_{90} distribution may include a third, *intermediate-duration group* (e.g. Mukherjee et al., 1998; Horváth, 1998; Hakkila et al., 2003; Horváth et al., 2010; Veres et al., 2010). However, multi-wavelength observations show no clear evidence that these intermediate-duration GRBs have a distinct physical origin.

Lightcurves

GRB *prompt emission lightcurves* are notoriously irregular. Figure 2.4 gives some examples of BATSE GRB lightcurves (denoted by their trigger numbers instead of the dates when they occurred). One can see very different patterns among GRBs. Trigger 332 is a typical one-pulse burst that shows a fast-rising exponential-decay (FRED) profile. Trigger 1989 shows several such pulses in one burst. Triggers 1606 and 1425 show complex multi-episode emission patterns that are very different from each other. Triggers 7994 and 143 show bursts with multi-episodes separated by quiescent gaps. Trigger 219 shows a main burst preceded by a weak precursor emission component. Trigger 8104 is a short-duration GRB, and trigger 1997 is a short-duration GRB followed by extended emission. An extreme example of a GRB was the so-called “double burst”, GRB 110709B, discovered by *Swift* (Zhang et al., 2012c). This burst triggered the *Swift* BAT detector twice (Fig. 2.5), with a gap of about 11 minutes. Analyses ruled out the possibility that the two

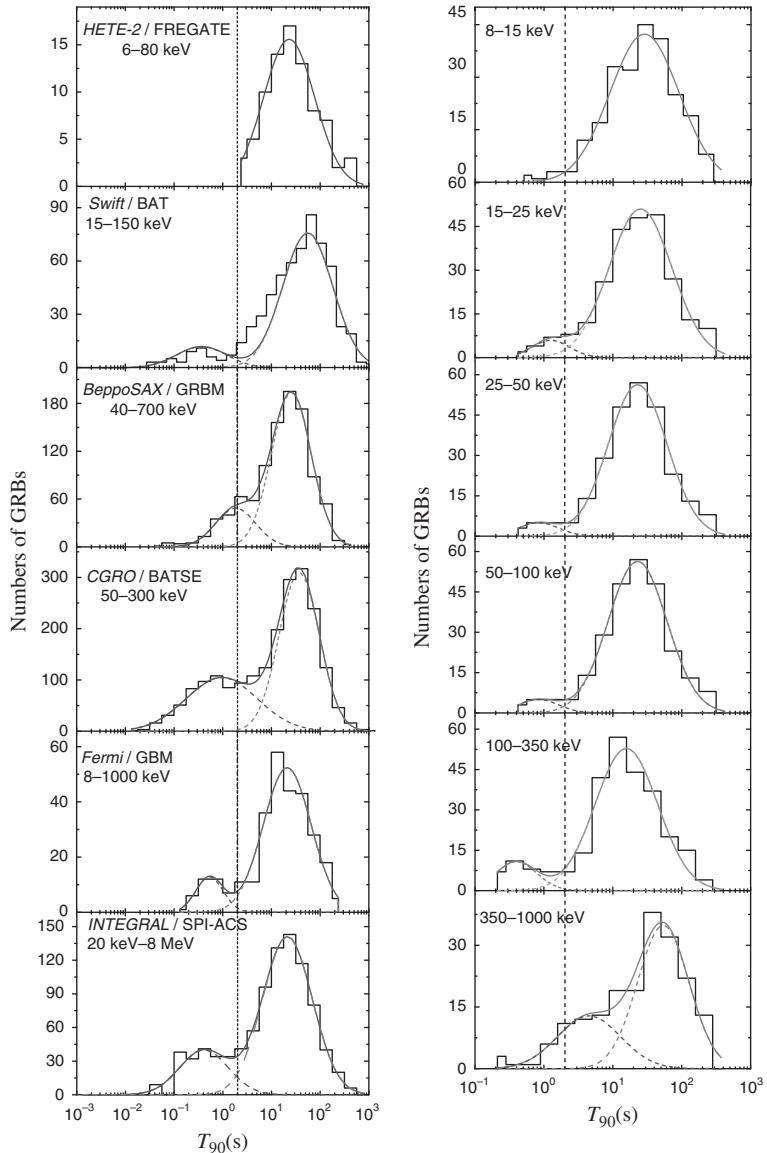


Figure 2.2 *Left:* Detected GRB T_{90} distribution histograms for different detector instruments. *Right:* T_{90} distribution of *Fermi* GRBs in different energy bands. Reproduced from Figures 3 and 5 in Qin et al. (2013) with permission. ©AAS.

episodes were caused by gravitational lensing of the same event, suggesting that the GRB central engine indeed produced two distinct episodes of radiation.

Precursor Emission

A fraction of GRBs have a *precursor emission* component well separated from the main burst. It is typically softer than the main burst. The separation time ranges from several tens

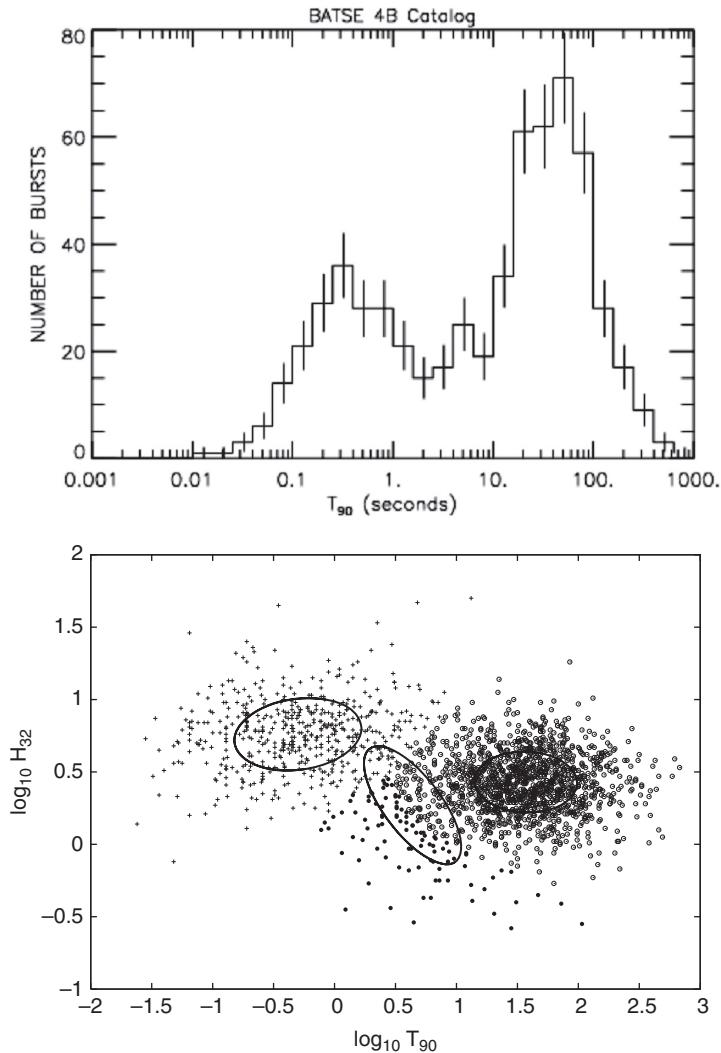


Figure 2.3 *Top:* Duration distribution of GRBs detected by BATSE on board CGRO. Reproduced from the BATSE GRB 4B Catalogs (<https://gamma-ray.nsstc.nasa.gov/batse/grb/duration/>). *Bottom:* Duration–hardness distributions of GRBs detected by BATSE on board CGRO. The three circles in the T_{90} –HR plot show the centers of the long-, short-, and the putative intermediate-duration GRBs. From Horváth et al. (2006).

to hundreds of seconds. There are several different definitions for the “precursor” emission. Koshut et al. (1995) defined a precursor as any case in which the first episode (the precursor episode) has a lower peak intensity than the remaining emission and is separated from the remaining burst emission by a background interval that is at least as long as the remaining emission. They found that $\sim 3\%$ of BATSE GRBs satisfy such criteria. A less stringent definition (e.g. Burlon et al., 2009) defines precursor emission as any emission episode that has a peak flux lower than the following main prompt emission and that is separated from

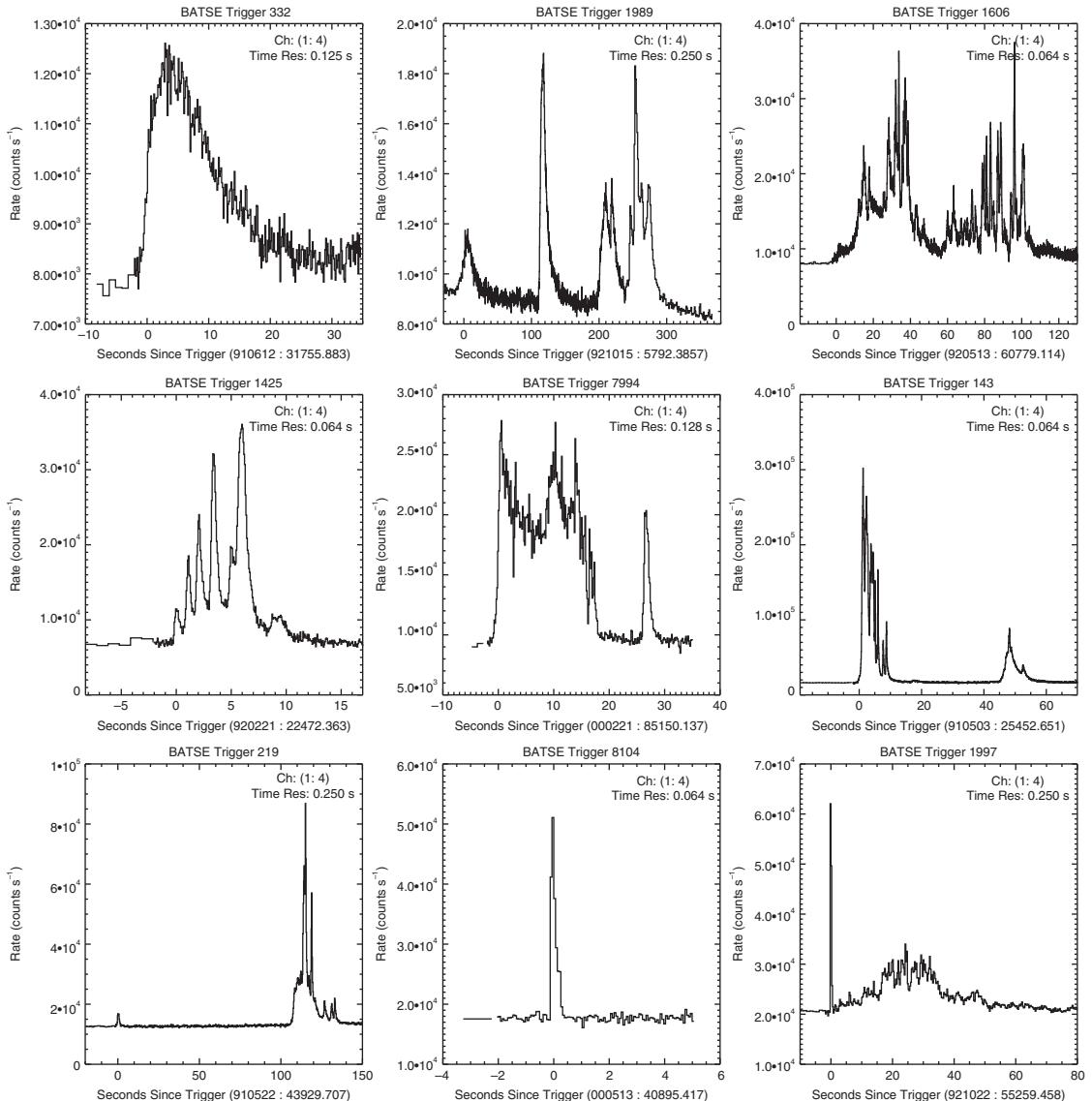


Figure 2.4 Sample lightcurves of BATSE GRBs. Reproduced from the BATSE GRB 4B Catalogs with permission.

the main event by a quiescent period (during which the background-subtracted lightcurve is consistent with being zero). Burlon et al. (2009) found that $\sim 12\%$ of BATSE bursts satisfy these criteria. Figure 2.6 presents some examples of *Swift* GRBs that are triggered by precursor emission (Hu et al., 2014). An interesting claim was that precursor emission also appears in some short GRBs (Troja et al., 2010).

Various statistical studies suggested that the characteristics of the main episode emission are independent of the existence of the precursor emission, and that the properties of the

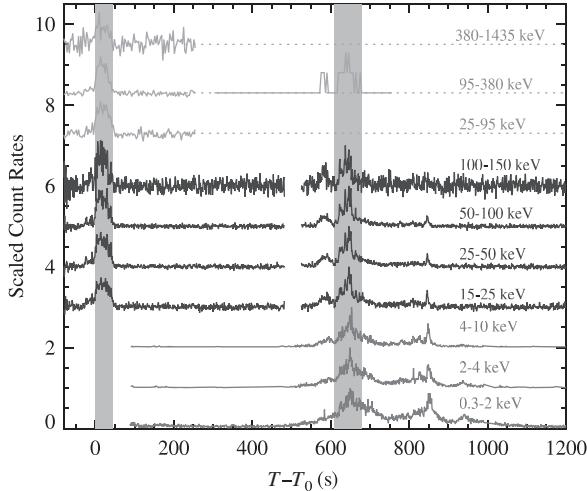


Figure 2.5 Multi-wavelength lightcurves of the “double burst”, GRB 110709B, detected by *Swift*, which shows two distinct emission episodes with an 11-minute gap. From Zhang et al. (2012c).

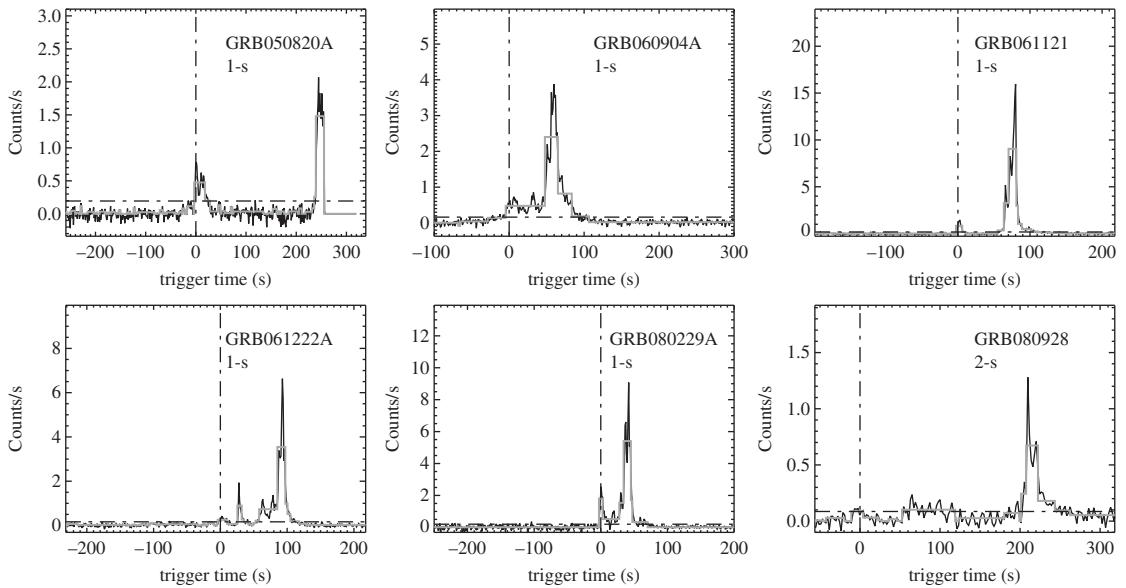


Figure 2.6 Examples of *Swift* GRBs triggered by a precursor. From Hu et al. (2014).

precursor emission in many GRBs are similar to those of the main-episode emission (Lazzati, 2005; Burlon et al., 2008, 2009). In the *Swift* era, a good fraction of GRBs are also found to be followed by softer *X-ray flares* (e.g. Burrows et al., 2005a), see §2.2 below. The simplest interpretation is that the precursors, main emission component, and X-ray flares can all be attributed to the erratic GRB central engine activities, with the central engine switched on or off due to some unknown mechanism. An early, weak activity would correspond to a precursor, while a late, weak activity would correspond to an X-ray flare. In rare

cases, these episodes have similar activity amplitudes, which would correspond to cases such as the double burst GRB 110709B (Zhang et al., 2012c). A systematic study of the properties of precursor emission, main emission, and X-ray flares supports this viewpoint (Hu et al., 2014).

The extremely bright GRB 160625B at $z = 1.406$ had three isolated emission episodes, which correspond to the precursor, main burst, and X-ray flare (with contribution in the BAT band), respectively. Its high luminosity allowed a careful study of the time-resolved spectra of different emission episodes (Zhang et al., 2018b). It was found that the spectra of the first sub-burst (precursor) are essentially thermal, while those of the second sub-burst (main burst) are non-thermal and consistent with a synchrotron origin. This GRB provides evidence of the evolution of GRB jet composition in widely separated emission episodes (Zhang et al., 2018b).

Power Density Spectrum

One can perform a Fourier transform of a GRB lightcurve and study possible features in the corresponding *power density spectrum* (PDS). If a time sequence has a clear period (e.g. lightcurve of a radio pulsar), the PDS would show a sharp feature at the typical frequency along with higher order harmonic spikes. If the time sequence does not have a precise period but has a quasi-periodic behavior (e.g. in X-ray binaries), the PDS would display a broad feature, suggesting a quasi-periodic oscillation (QPO).

PDS analyses of GRB lightcurves revealed null periodicity (Beloborodov et al., 1998, 2000; Guidorzi et al., 2012). This suggests that the GRB central engine does not display an apparent periodic behavior. For individual GRBs, the derived PDSs can be very noisy without a clear feature. This is related to the short durations of the time sequences for GRBs. Nonetheless, when averaging over many GRBs, an interesting feature emerges. For BATSE GRBs, the average PDS of some bright GRBs is a power law with an index of about $-5/3$. The power law extends over almost two decades in frequency, until reaching a sharp break above around 1 Hz (Beloborodov et al., 2000). For *Swift* GRBs, the PDS slope is steeper (-1.7 to -2.0), and no apparent break at 1 Hz is observed (Guidorzi et al., 2012).

The power-law behavior of the PDS suggests a self-similar behavior over a wide range of time scales, suggesting random realizations of the same process on different time scales. It is intriguing that the $-5/3$ index found in BATSE GRBs coincides with the theoretical value of fully developed hydrodynamical turbulence (Kolmogorov, 1941).

Fast and Slow Components

Even though PDS analyses did not reveal temporal features in the lightcurves, visual inspections suggest that GRB lightcurves seem to show the superposition of a *broad (slow) component* and a *rapid (fast) component* (e.g. Fig. 2.4). Such a superposition effect can be revealed through other data analysis methods. Vetere et al. (2006) analyzed the lightcurves of *BeppoSAX* GRBs in three different energy bands (2–5, 5–10, and 10–26 keV), and found that the spikier (fast) component becomes less significant in low

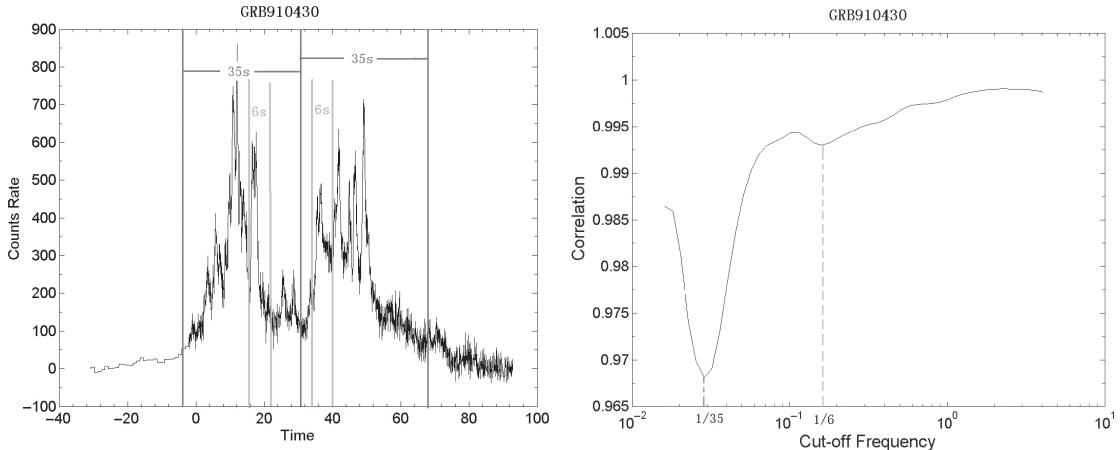


Figure 2.7 The stepwise filter correlation analysis of GRB variabilities. The dips on the correlation curves suggest possible variability components. From Gao et al. (2012).

energies, suggesting that the fast and slow variability components are likely different variability components. Gao et al. (2012) introduced a *stepwise low-pass filter correlation (SFC) method* to study the possible superposition effect. By progressively filtering the high-frequency signals from a lightcurve, they performed a correlation analysis between each adjunct pair of filtered lightcurves to see whether a significant change occurs. They found that there are indeed “dips” in the correlation curves at certain frequencies, which correspond to a typical variability time scale. They found two or more dips in some lightcurves, suggesting that at least some GRB lightcurves are indeed the superposition of fast and slow variability components. One example is given in Fig. 2.7.

Pulses

Most GRB lightcurves can be decomposed into many *pulses* (e.g. Norris et al., 1996, 2005). The shape of the pulses is typically asymmetric, with a sharper rising phase and a shallower decaying phase (Fig. 2.8). On top of these pulses, one usually still sees fast variability, so the pulses can be regarded as tracers of the slow variability component discussed above.

Several functional forms have been proposed to fit the pulse shapes. A common phrase to describe the pulse shape is fast-rising exponential decay, or “FRED”. The pulse function is however not uniquely defined. Norris et al. (1996) suggested an asymmetric exponential-rise, exponential-decay profile

$$I(t) = \begin{cases} A \exp \left[-\left(\frac{|t-t_{\max}|}{\sigma_r} \right)^v \right], & t < t_{\max}, \\ A \exp \left[-\left(\frac{|t-t_{\max}|}{\sigma_d} \right)^v \right], & t > t_{\max}, \end{cases} \quad (2.1)$$

to fit the GRB pulses, where A is the normalization parameter, t_{\max} is the peak time, σ_r and σ_d are the rise and decay time constants, and v measures the sharpness of the pulse. Another pulse model (Kocevski et al., 2003) invokes a five-parameter (I_p, t_p, t_0, r, d) function

$$I(t) = I_p \left(\frac{t+t_0}{t_p+t_0} \right)^r \left[\frac{d}{r+d} + \frac{r}{r+d} \left(\frac{t+t_0}{t_p+t_0} \right)^{r+1} \right]^{-\frac{r+d}{r+1}} \quad (2.2)$$

to define a pulse. Besides the amplitude (I_p), peak time (t_p), and rise and decay time scale (r and d) parameters, a zero time t_0 can be also introduced to allow flexibility to fit the data pulses. Later, Norris et al. (2005) proposed a pulse model invoking four parameters (A , λ , τ_1 , τ_2):

$$I(t) = \frac{A\lambda}{\exp\left(\frac{\tau_1}{t} + \frac{t}{\tau_2}\right)}, \quad t > 0, \quad (2.3)$$

with $\lambda = \exp(2\mu)$, and $\mu = (\tau_1/\tau_2)^{1/2}$. Since λ is a function of τ_1 and τ_2 , there are three independent parameters.

Fitting the pulse models to the data, the pulses can essentially catch the “slow” component of the lightcurves, and there are usually additional “fast” spikes overlapping on the pulses (Fig. 2.8). Some careful analyses suggested that the simple pulse models are inadequate to delineate all the pulses, with the residues sometimes displaying further features (Hakkila and Preece, 2014).

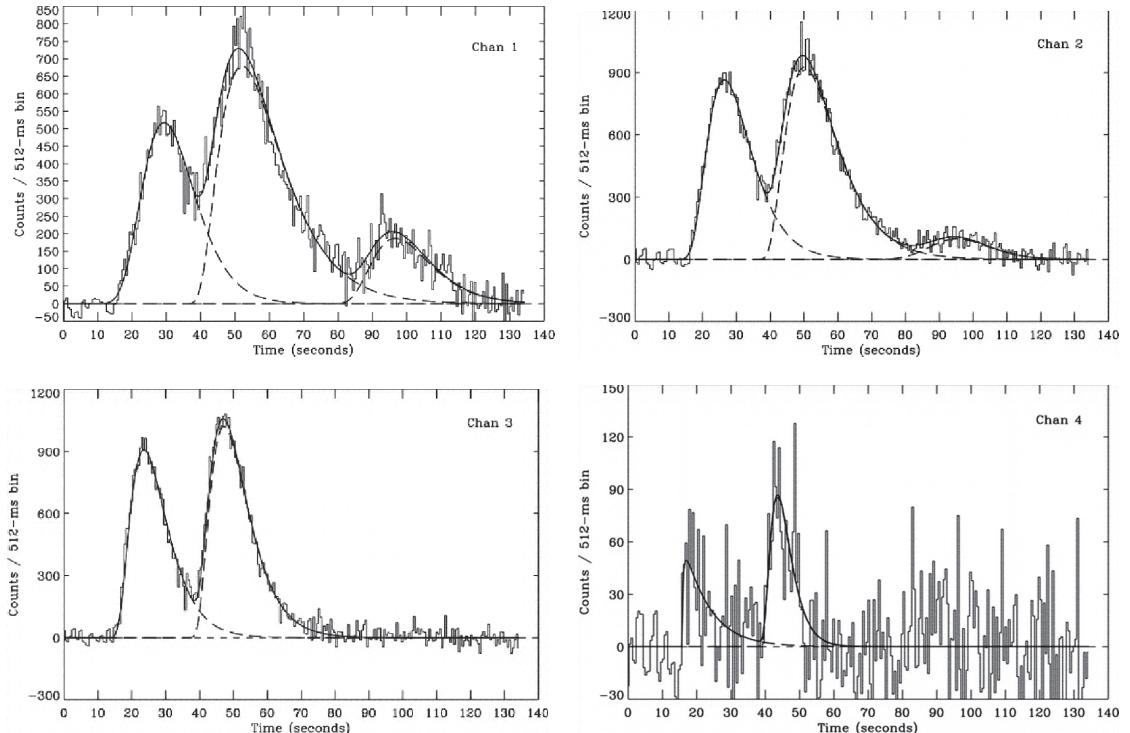


Figure 2.8 An example of the decomposed pulses in the lightcurves of GRB 000323 in different energy channels of BATSE. Reproduced from Figure 26 in Norris et al. (2005) with permission. ©AAS.

Energy Dependence

GRB lightcurves vary with energy. Figures 2.9 and 2.10 show lightcurves of two bright *Fermi* GRBs (080916C and 090902B) in different energy bands.

In general (but not always), pulses tend to be narrower in harder energy bands, and wider in softer energy bands. One can fit the energy-dependent pulse width $w(E)$ as a power-law function of energy E :

$$w(E) \propto E^{-\alpha}. \quad (2.4)$$

The typical value of the index is $\alpha \sim (0.3\text{--}0.4)$ (Norris et al., 2005; Liang et al., 2006b). See lower left panel of Fig. 2.11.

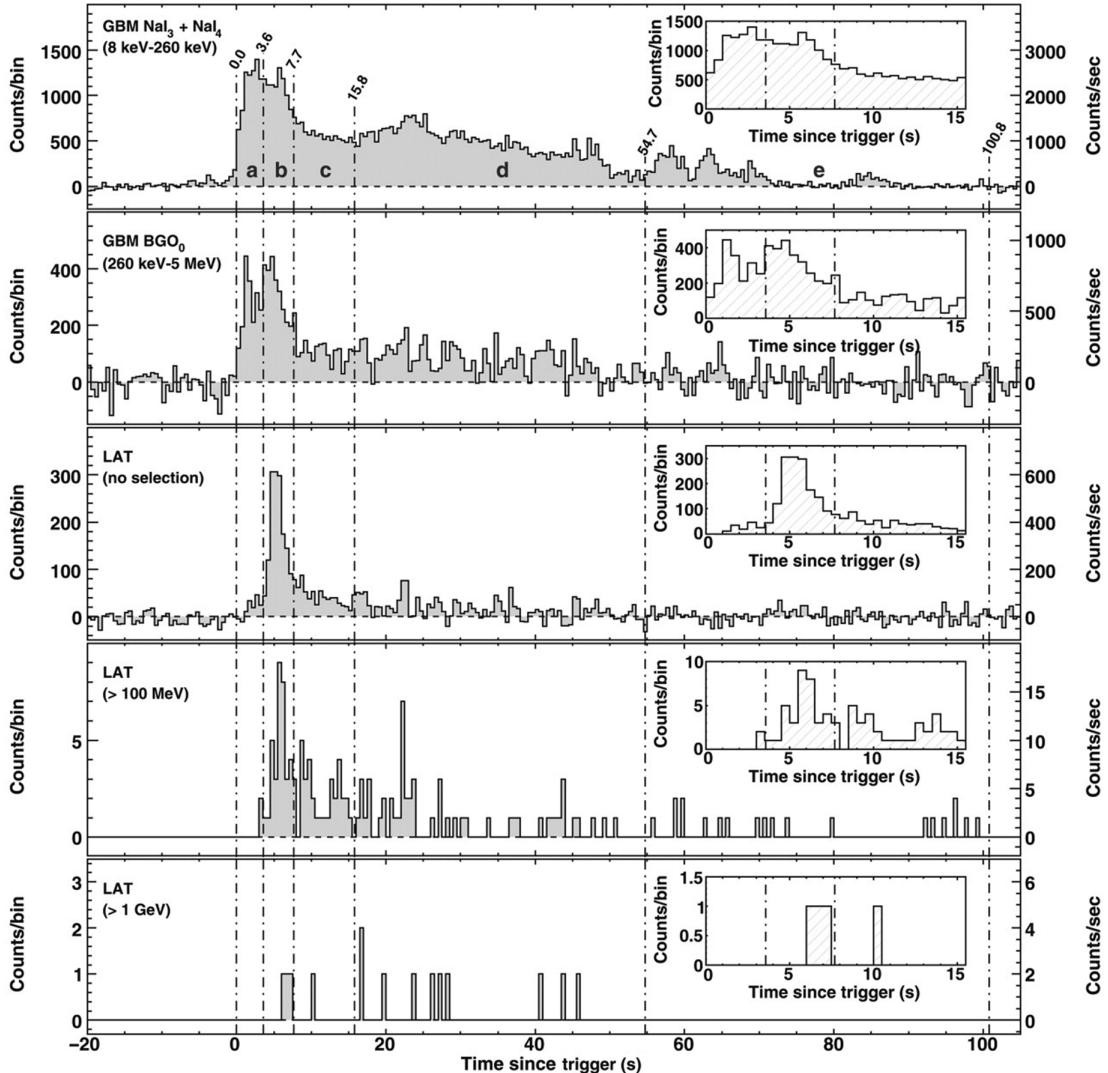


Figure 2.9 Multi-wavelength lightcurves of GRB 080916C as observed with *Fermi* GBM and LAT. From Abdo et al. (2009c).

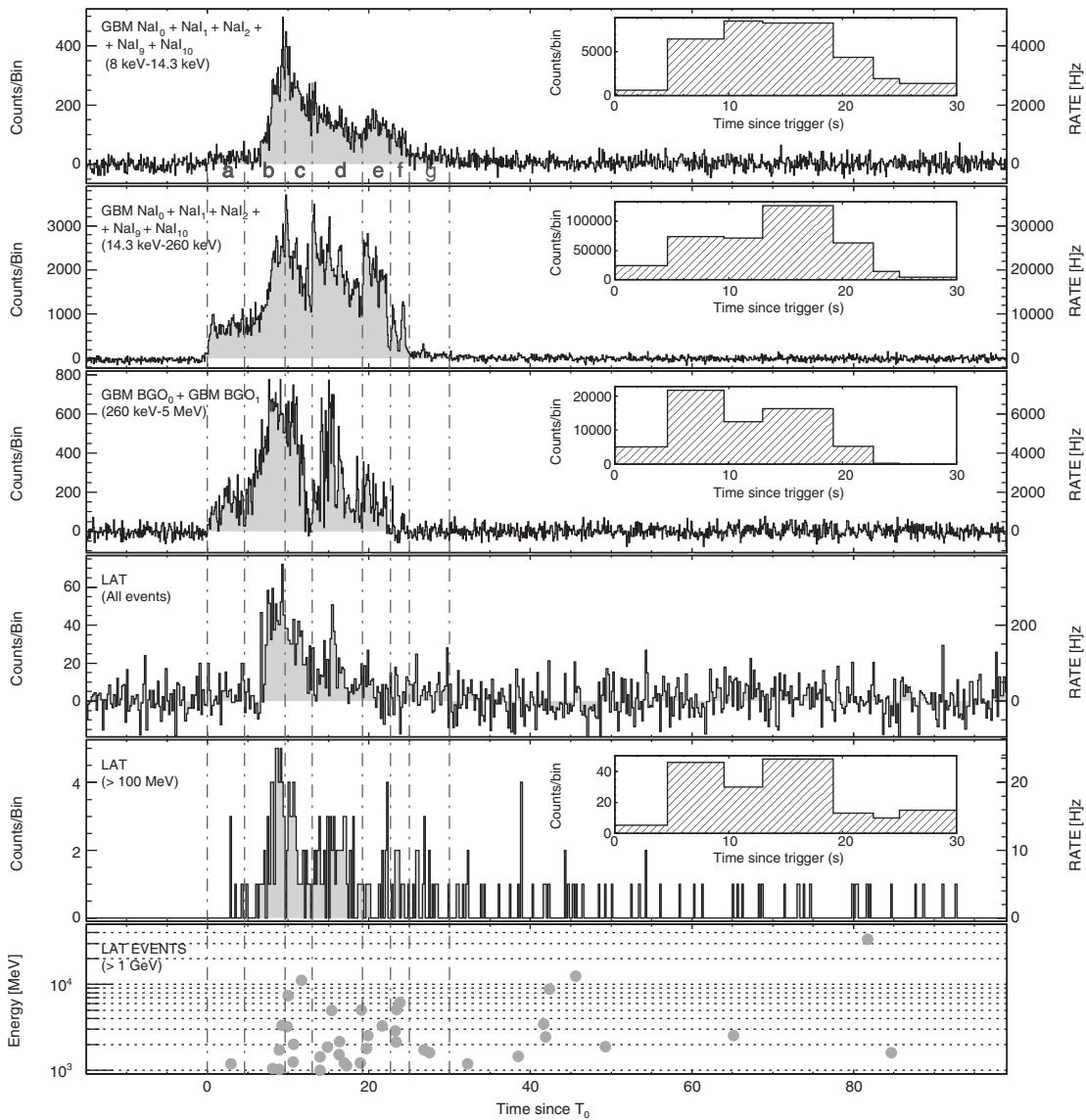


Figure 2.10 Multi-wavelength lightcurves of GRB 090902B as observed with *Fermi* GBM and LAT. Reproduced from Figure 1 in Abdo et al. (2009b) with permission. ©AAS.

Spectral Lag

In the energy range below 10 MeV, the arrival time of a pulse in a softer band is typically delayed (or “lagged”) with respect to the arrival time in a harder band (Norris et al., 2000; Norris, 2002; Norris et al., 2005). Such a *spectral lag* may be visually inspected as the delay time of the pulse peaks in different energies, but can be more rigorously calculated using quantitative methods, such as the *cross-correlation function (CCF) method*: one may

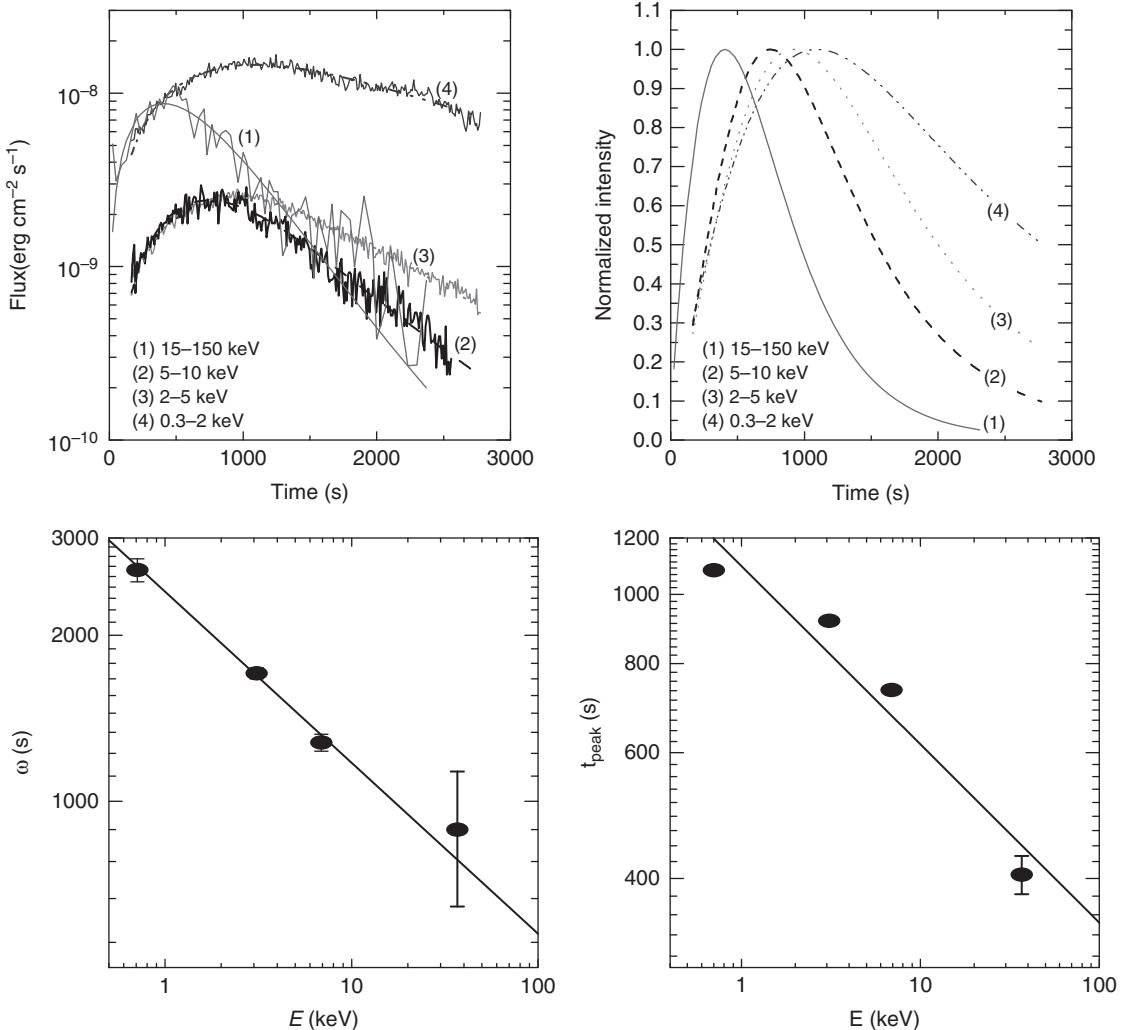


Figure 2.11 GRB 060218 as an example of energy-dependent pulse width and spectral lags. *Upper left:* Photon count lightcurves in different energy bands. *Upper right:* Normalized lightcurves to show clear spectral lags in different bands. *Lower left:* Energy-dependent pulse width w . *Lower right:* Energy-dependent peak time t_p . From Liang et al. (2006a).

progressively shift the lightcurve in a particular band and cross correlate with the lightcurve in another band. The lag time can be measured when the two lightcurves reach the best correlation. An example of GRB 060218 (Liang et al., 2006a) is presented in Fig. 2.11.

Even though commonly observed in many GRB pulses, there are pulses that do not show significant lags. In general, significant spectral lags are more commonly measured in long-duration GRBs. Short GRBs do not show significant spectral lags (Norris and Bonnell, 2006). A fraction of short GRBs even show “negative” lags, i.e. high-energy pulses lag behind low-energy pulses (Yi et al., 2006).

In the >100 MeV regime, the onset of high-energy emission typically lags behind the low-energy (keV–10 MeV) emission (Abdo et al., 2009c,a; Zhang et al., 2011), see

Figs. 2.9 and 2.10. This is an opposite trend to the low-energy lag behavior, and likely has a different astrophysical origin.

2.1.3 Spectral Properties

How to Describe a Spectrum

GRB spectra are non-thermal. A *thermal spectrum* is produced by a gas in thermal equilibrium, which is defined by a temperature. There is an exponential decrease of photon flux at high energies. A *non-thermal spectrum*, on the other hand, extends to much higher energies, typically with a power-law distribution in flux. This requires the electrons to be accelerated in a non-equilibrium environment (e.g. shocks, magnetic reconnection sites).

In order to extract a spectrum, one needs to accumulate photons over a certain interval of time. Since GRBs are highly variable events, ideally one should extract spectra in time bins as small as possible, so that the evolution of spectral properties can be studied. Practically, the smallest time bin is limited by photon statistics; namely, the number of photons in the time bin should be large enough to allow a reasonable spectral fit to test various plausible spectral models.

Before moving on to introduce the mathematical forms of various spectral models, it is necessary to introduce the conventions of expressing a spectrum. In high-energy astronomy, a spectrum is usually displayed in the form of intensity as a function of frequency (or energy) rather than wavelength (which is more commonly used in optical astronomy) on a logarithmic scale. There are four ways to display the same spectrum:

- The first is the raw photon count spectrum $C(E)$, in units of $\text{cts} \cdot \text{s}^{-1} \cdot \text{keV}^{-1}$. This is to display the detected photon number counts as a function of energy bin. Since a detector usually has different sensitivities in different energy bins, the count spectrum is heavily affected by the detector's instrumental response function, and shows an irregular shape even if the intrinsic spectrum is regular (e.g. power law or thermal).
- The second way to describe a spectrum is the “photon number” spectrum $N(E)$, in units of $\text{photons} \cdot \text{cm}^{-2} \cdot \text{s}^{-1} \cdot \text{keV}^{-1}$. By extracting such a spectrum, one has already corrected for instrumental response function effects so that the spectrum mimics the true specific photon flux detected from Earth. In this notation, $N(E)dE$ is the number of photons in the energy bin dE . In X-ray and γ -ray astronomy, photons are counted individually, so that the photon number spectrum is the most straightforward to obtain.
- The third is to display the “specific flux density” spectrum, usually expressed as F_ν (e.g. in units of $\text{erg} \cdot \text{cm}^{-2} \cdot \text{s}^{-1} \cdot \text{Hz}^{-1}$), or $EN(E)$ (e.g. in units of $\text{erg} \cdot \text{cm}^{-2} \cdot \text{s}^{-1} \cdot \text{keV}^{-1}$). Such a spectrum is usually used in IR/optical/UV astronomy, when individual photons cannot be counted directly. Instead, the photon energy per unit frequency (or unit energy) is measured and displayed. As a result, $F_\nu d\nu$ or $EN(E)dE$ is the photon energy in the frequency bin $d\nu$ or energy bin dE .
- The last one is to display the “energy” spectrum, usually expressed as νF_ν or $E^2 N(E)$ (e.g. in units of $\text{erg} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}$). This is also called a “spectral energy distribution” (SED).

By displaying such a spectrum, one can immediately see how the bolometric energy of the source is distributed in frequency or energy. Such a spectrum is of theoretical interest the most.

For GRBs, $N(E)dE$ spectra are usually constructed first. Several spectral models have been applied to fit such spectra.

“Band” Function

When the detector’s energy band is wide enough, a GRB spectrum can usually be fit with a smoothly joint (in an exponential form) broken power law known as the *Band function* or *GRB function* (Band et al., 1993). The photon number spectrum in this model reads

$$N(E) = \begin{cases} A \left(\frac{E}{100 \text{ keV}} \right)^\alpha \exp \left(-\frac{E}{E_0} \right), & E < (\alpha - \beta)E_0, \\ A \left[\frac{(\alpha - \beta)E_0}{100 \text{ keV}} \right]^{\alpha - \beta} \exp(\beta - \alpha) \left(\frac{E}{100 \text{ keV}} \right)^\beta, & E \geq (\alpha - \beta)E_0, \end{cases} \quad (2.5)$$

where A is the normalization of the spectrum, E_0 is the break energy in the spectrum, α and β (both negative) are the *low-energy* and *high-energy photon spectral indices*, respectively.¹ The two spectral regimes are separated by the *break energy* E_0 . The peak energy in the $E^2N(E)$ spectrum is called the *E peak* (E_p), which is related to E_0 through

$$E_p = (2 + \alpha)E_0. \quad (2.6)$$

Figure 2.12 gives an example of GRB 990123 whose time-integrated spectrum is well fit by the Band function (Briggs et al., 1999a).

The E_p distribution of GRBs covers at least 3 orders of magnitude. While bright BATSE GRBs (a sample of 156 bursts with 5500 spectra) have E_p clustered around the 200–300 keV range (Preece et al., 2000; Goldstein et al., 2013), bursts with lower E_p were also observed with softer detectors such as *HETE-2* and *Swift*. The distribution of E_p seems to form a continuum from several keV to multi-MeV (e.g. Gruber et al., 2014; Bošnjak et al., 2014). From hard to soft, bursts are sometimes vaguely classified as gamma-ray bursts (GRBs, $E_p > 50$ keV), X-ray rich GRBs (XGRBs, $30 \text{ keV} < E_p < 50 \text{ keV}$), and X-ray flashes (XRFs, $E_p < 30$ keV), with no clear boundary in between (Sakamoto et al., 2008a). For the bright BATSE sample, the two spectral indices have a distribution of $\alpha \sim -1 \pm 1$ and $\beta \sim -2^{+1}_{-2}$ (Preece et al., 2000). Such distributions were confirmed for the GRBs detected by other detectors such as *Fermi* and the *INTERnational Gamma-Ray Astrophysics Laboratory (INTEGRAL)* (Zhang et al., 2011; Nava et al., 2011b; Goldstein et al., 2012; Gruber et al., 2014; Bošnjak et al., 2014).

¹ Within the GRB afterglow context, another convention for the notations α and β is used: $F_\nu \propto t^{-\alpha} \nu^{-\beta}$, where α and β are the temporal decay index and flux density spectral index of the afterglow, respectively. In this book, we do not differentiate these notations and still keep the conventions widely adopted in the community. The physical meanings of the notations are usually self-evident within the context of the book, but we alert readers to pay special attention to the notations to avoid possible confusion.

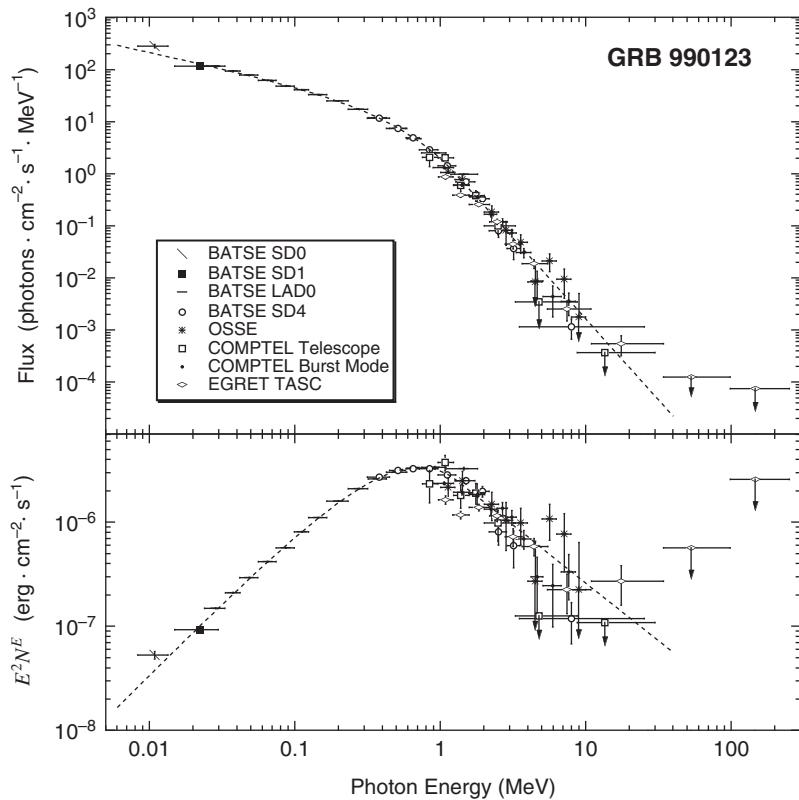


Figure 2.12 An example Band-function spectrum in GRB 990123. Reproduced from Figure 2 in Briggs et al. (1999a) with permission. ©AAS.

Cutoff Power Law and Power Law

If a detector's energy band is not wide enough or a GRB is not bright enough, the spectrum of the GRB sometimes can be fit by a *cutoff power law*, in the form of

$$N(E) = A \left(\frac{E}{100 \text{ keV}} \right)^\alpha \exp \left(-\frac{E}{E_c} \right). \quad (2.7)$$

This is essentially the first portion of the Band function, with the break energy E_0 being replaced by the cutoff energy E_c . Similar to the Band function, the peak energy in the $E^2 N(E)$ spectrum of this model is

$$E_p = (2 + \alpha)E_c. \quad (2.8)$$

This function has been used to fit the prompt emission spectra of many *HETE-2*, *Swift*, and *Fermi*/GBM GRBs (Sakamoto et al., 2005, 2008b; Paciesas et al., 2012). However, this is mainly due to the narrow bandpass of the detectors or low statistics of the high-energy photon counts, so that the high-energy photon index β of the putative Band function is not

well constrained. In fact, when a *Swift* burst is co-detected by another detector with high-energy band coverage (e.g. *Wind*/KONUS, *Fermi*/GBM), in most cases the global spectrum can still be fit by a Band function.

For historical reasons, this functional form is sometimes called a *Comptonized model* (e.g. Gruber et al., 2014). This was because historically people introduced a Comptonized model (e.g. Brainerd, 1994; Liang, 1997) to interpret GRB spectra with a cutoff power-law form. In general, physical mechanisms other than Comptonization (e.g. synchrotron radiation with an intrinsic cutoff in the electron energy distribution) can also account for a cutoff power-law spectrum, and Comptonization can also give rise to spectra different from the cutoff power-law form (§5.2.5). Therefore, the term “Comptonized model” carries misleading information, and should be avoided.

Even though most GRBs with a cutoff power-law spectrum may have an intrinsic Band spectrum whose high-energy spectral index is not well constrained, an intrinsic, cutoff power-law model was found to correctly describe the joint *Swift* BAT/XRT time-dependent prompt emission spectra of the nearby low-luminosity GRB 060218 (Campana et al., 2006). The E_p of this burst rapidly evolved with time from ~ 80 keV to 5 keV. Since GRB 060218 is special in many aspects (e.g. nearby, low luminosity, supernova association, extremely long duration, existence of a thermal X-ray component that might be of a shock breakout origin), the prompt emission of this burst (and probably also of other nearby low-luminosity GRBs) may have a different mechanism than most high-luminosity GRBs.

For a narrow detector’s bandpass and a faint GRB (e.g. *Swift* BAT GRBs near the detection threshold), a GRB spectrum sometimes can only be fit with a *simple power law* (e.g. Sakamoto et al., 2008b, 2011):

$$N(E) = A \left(\frac{E}{100 \text{ keV}} \right)^{-\hat{\Gamma}}, \quad (2.9)$$

where $\hat{\Gamma}$ is the photon index, which is positive by definition. The intrinsic spectrum is likely curved (e.g. Band function), but the photon number is too small to constrain the parameters in more complicated models. Indeed, if the same burst was also observed by other instruments (e.g. *Fermi*/GBM, or *Wind*/KONUS) with a wider spectral window, usually the spectrum could be fit with a Band function or a cutoff power law with a measured E_p . For *Swift* GRBs, it was found that the measured E_p is (crudely) correlated with the BAT-band photon index $\hat{\Gamma}_{\text{BAT}}$ (Sakamoto et al., 2009; Zhang et al., 2007b). A systematic analysis of joint *Swift*/*Fermi* GRBs refined these correlations, which read (Virgili et al., 2012)

$$\log E_p = (4.40 \pm 0.51) - (1.31 \pm 0.15) \hat{\Gamma}_{\text{BAT}}, \quad (2.10)$$

or

$$\log E_p = (3.05 \pm 0.36) - (3.79 \pm 0.55) \log \hat{\Gamma}_{\text{BAT}}. \quad (2.11)$$

These correlations may be used to roughly estimate E_p for *Swift* GRBs when they are not jointly detected by other detectors but their E_p ’s are needed for other purposes.

Thermal Component

Even though the main spectral component of GRB spectra is non-thermal and Band-like, a (*quasi-)*thermal component is found to contribute to the observed spectra of a good fraction of GRBs. In the pre-*Fermi* era, it was suggested (Ryde, 2005; Ryde and Pe'er, 2009) that the observed prompt GRB spectra of some BATSE GRBs are the superposition of a thermal (blackbody) component and a non-thermal (power-law) component. Within such a picture, the observed E_p is interpreted as the peak of the thermal component defined by its temperature. The spectra of some BATSE GRBs can be fit with such a “hybrid” model, which, within the BATSE window, may mimic a Band-function spectrum. This model however over-predicts the flux in the X-ray range for most GRBs, which violates the observational constraints by *BeppoSAX* for some BATSE bursts (Ghirlanda et al., 2007; Frontera et al., 2013). The predicted X-ray excess was indeed observed in the *Fermi* GRBs 090902B (Abdo et al., 2009b) and 090510 (Ackermann et al., 2010), suggesting that the superposition model is valid for at least some bursts. On the other hand, these cases seem uncommon.

Fermi, with both GBM and LAT on board, significantly widened the observational spectral window, and allowed a systematic search for the thermal component in GRB spectra. The first bright LAT GRB, GRB 080916C, has a series of time-resolved spectra that are adequately described by a Band function that covers 6–7 orders of magnitude (Abdo et al., 2009c) (top left panel of Fig. 2.13). A thermal component, if any, must be sub-dominant.² Soon after, a very different burst, GRB 090902B (Abdo et al., 2009b), was discovered. It has a narrow Band component (with unusually hard α index and soft β index) superposed on an underlying power-law segment. The narrow Band component is found to be consistent with a multi-color quasi-thermal spectrum (Ryde et al., 2010; Zhang et al., 2011) (top right panel of Fig. 2.13). More interestingly, this component becomes narrower as the time bin becomes smaller (Zhang et al., 2011) (Fig. 2.14). This suggests that this narrow Band component is likely the thermal component. The time-resolved spectra of GRB 080916C, on the other hand, do not show such narrowing as the time bin becomes smaller (Zhang et al., 2011) (Fig. 2.14). This suggests that these two GRBs present two types of GRB spectra, non-thermal-dominated (GRB 080916C-like) and thermal-dominated (GRB 090902B-like) ones. It also suggests that both a non-thermal component and a thermal component can define E_p in a GRB. A systematic analysis of 17 *Fermi*/LAT GRBs suggests that the first type (GRB 080916C-like) is more common (14/17), while the second type (GRB 090902B-like) is relatively rare (2/17) (Zhang et al., 2011). Another example of the thermal-dominated case may be the short-duration LAT GRB 090510 (lower left panel of Fig. 2.13).

One would naturally expect some intermediate types of spectra in which both non-thermal and thermal components co-exist. Later *Fermi* observations indeed show several examples that display a sub-dominant thermal component appearing at the left shoulder of the Band component, such as GRB 100724B (Guiriec et al., 2011), GRB 110721A

² A later more detailed analysis revealed the existence of a weak thermal component (Guiriec et al., 2015), which is indeed sub-dominant.

2.1 Prompt Emission

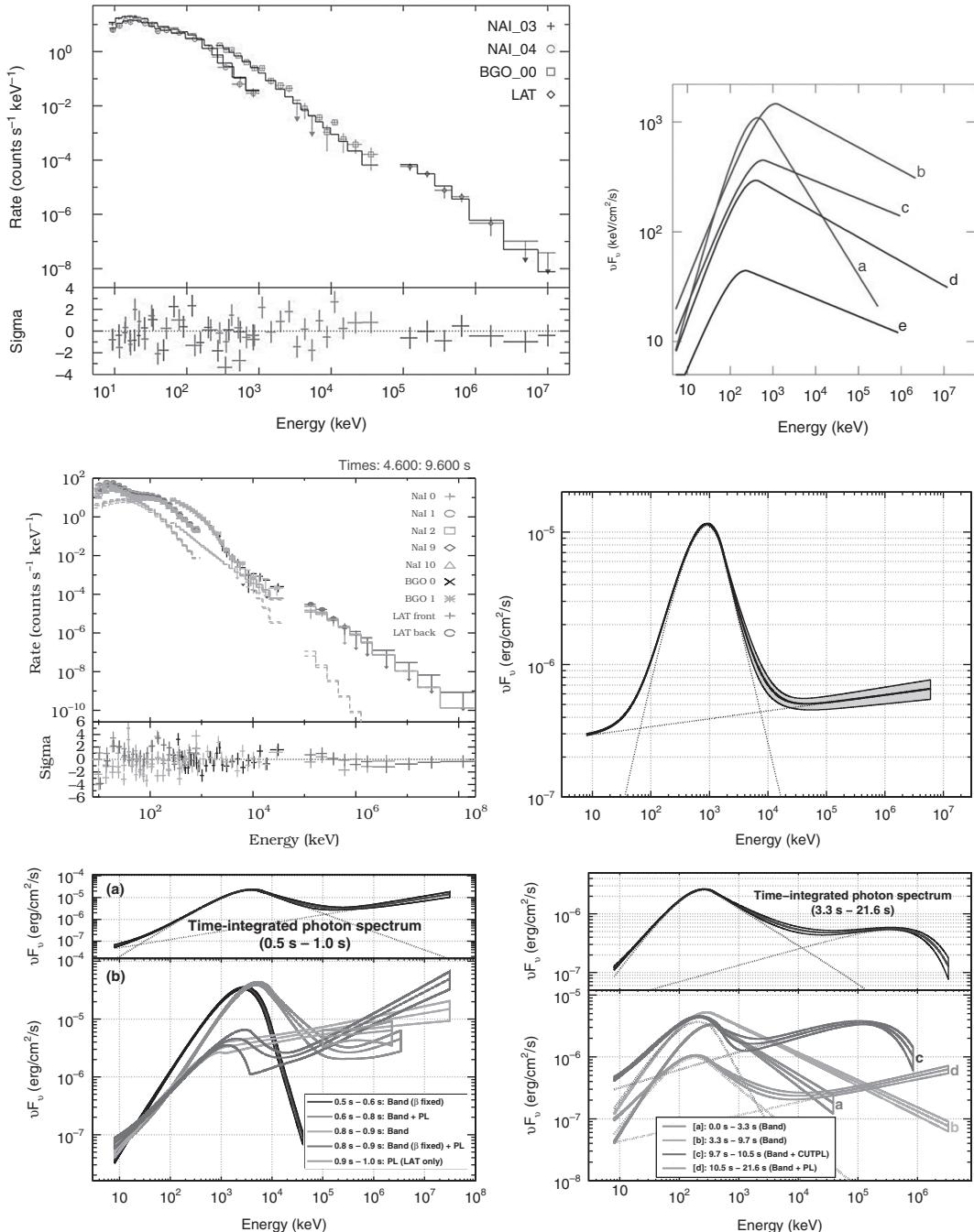


Figure 2.13 The best fit spectral models for the time-resolved spectra of four bright GRBs detected by *Fermi* GBM and LAT. First row: GRB 080916C. From Abdo et al. (2009c). Second row: GRB 090902B. Reproduced from Figure 3 in Abdo et al. (2009b) with permission. ©AAS. Lower. Lower left: Short GRB 090510. Reproduced from Figure 5 in Ackermann et al. (2010) with permission. ©AAS. Lower right: GRB0909026. Reproduced from Figure 5 in Ackermann et al. (2011) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

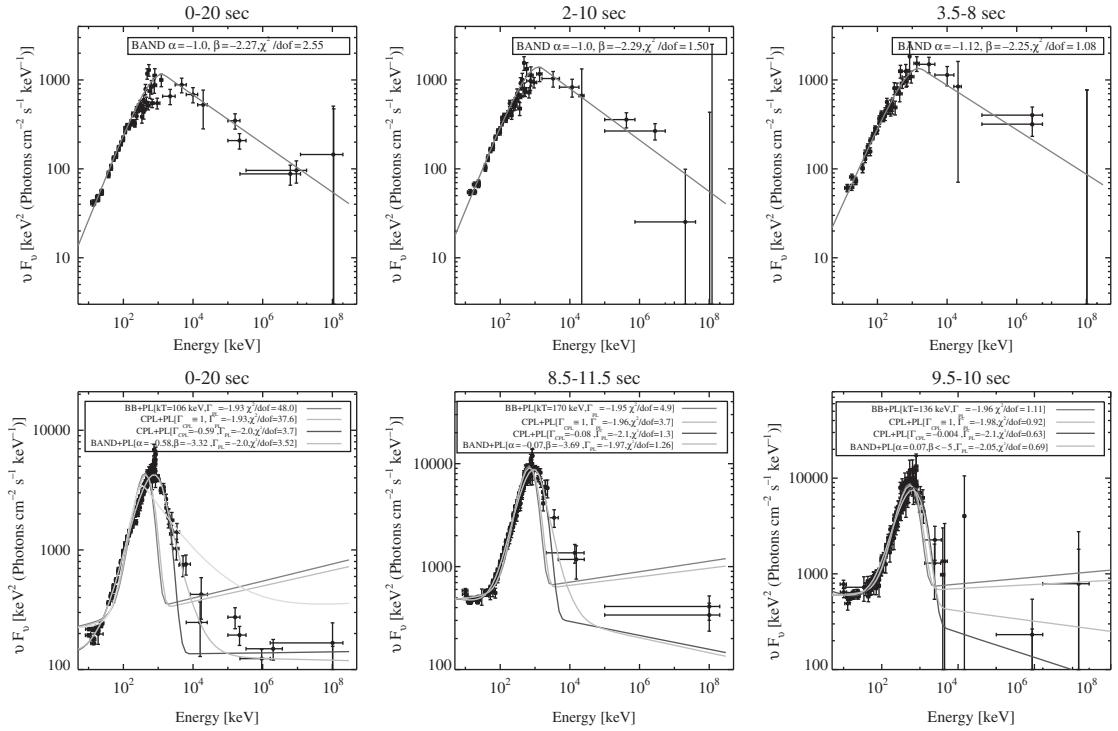


Figure 2.14 A comparison between GRB 080916C, which shows no evidence of spectral narrowing with reducing time bin, and GRB 090902B, which shows clear spectral narrowing with reducing time bin. This suggests that the former is dominated by a non-thermal spectral component, while the latter is dominated by a thermal spectral component. From Zhang et al. (2011).

(Axelsson et al., 2012), and GRB 120323A (Guiriec et al., 2013) (see Fig. 2.15 for two examples).

High-Energy Component

Besides the Band and thermal components, another *high-energy spectral component* is required to fit the broad-band spectra of some GRBs. Hints of the existence of such a component came from the EGRET GRB 941017 (González et al., 2003). Later, this component was clearly detected in several *Fermi* LAT GRBs (e.g. GRB 090902B, 090510, and 090926A (Abdo et al., 2009a; Ackermann et al., 2010, 2011)). This is a power-law component extending to the *Fermi* LAT band in high energies (above 100 MeV), but sometimes also extending to low energies (in the X-ray band). The slope is usually positive in the νF_ν spectrum. In order to avoid a divergence in energy there must be a turnover at high energies (say, in the 1–100 GeV range in view of the non-detections by the ground-based TeV detectors), which would define a second E_p . Such a second E_p was indeed inferred from the spectral analysis of GRB 090926A (Ackermann et al., 2011) (see lower right panel of Fig. 2.13).

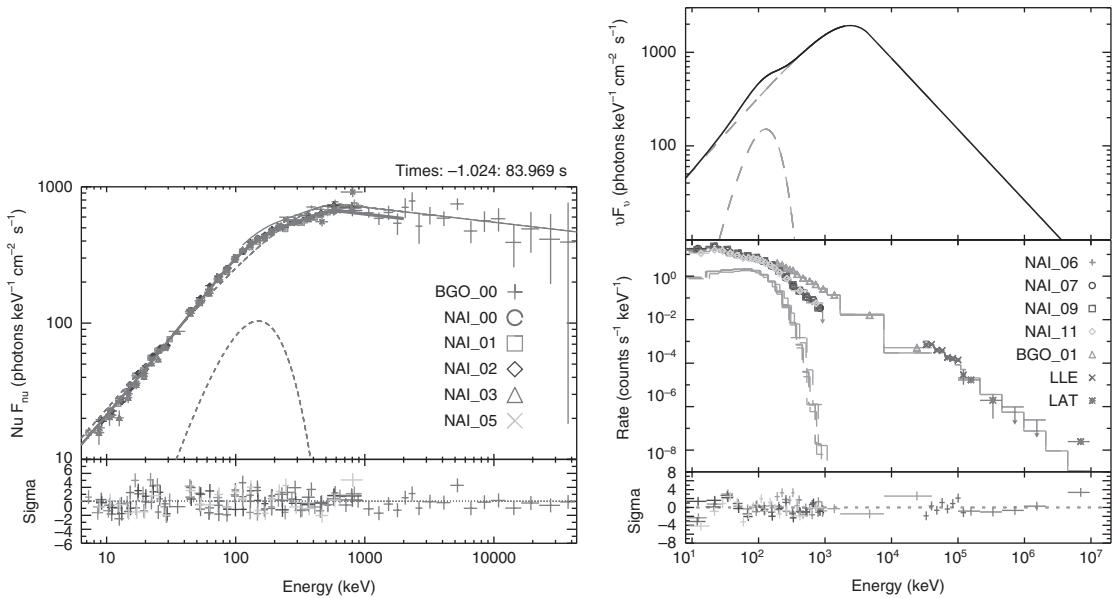


Figure 2.15 Spectral fits and residuals to the time-integrated spectra of two GRBs that show superposition of a thermal (blackbody) component on a non-thermal (Band) component. *Left:* GRB 100724B. Reproduced from Figure 2 in Guiriec et al. (2011) with permission. ©AAS. *Right:* GRB 110721A. Reproduced from Figure 2 in Axelsson et al. (2012) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

Elemental Spectral Components

One may speculate a synthesized prompt emission spectrum of GRBs, which may include three *elemental spectral components* (Zhang et al., 2011): (I) a non-thermal Band component; (II) a quasi-thermal component; and (III) another non-thermal component extending to high energies (Fig. 2.16). The significance of different spectral components may vary among GRBs. Usually component I is the dominant component. The superposition of components I and II has been seen in 100724B (Guiriec et al., 2011), 110721A (Axelsson et al., 2012), 120323A (Guiriec et al., 2013), and several other GRBs; while the superposition between I (or II) and III has been seen in GRB 090926A (Ackermann et al., 2011). GRB 090902B is likely an example of the superposition between components II and III. It is even possible that all three components exist in at least some GRBs (Guiriec et al., 2015). While some correlations between components I and II have been reported (e.g. Burgess et al., 2014), component III seems to evolve independently, and usually emerges at a later epoch than the other two components (e.g. Ackermann et al., 2011).

The physical origins of the three elemental spectral components have not been fully identified. A plausible picture attributes the thermal component (II) to the photosphere emission from the ejecta, and the Band component (I) to the non-thermal synchrotron radiation in the optically thin region. Some argue that both components are quasi-thermal emission from the photosphere. Component III is mysterious. Rapid variability associated with this

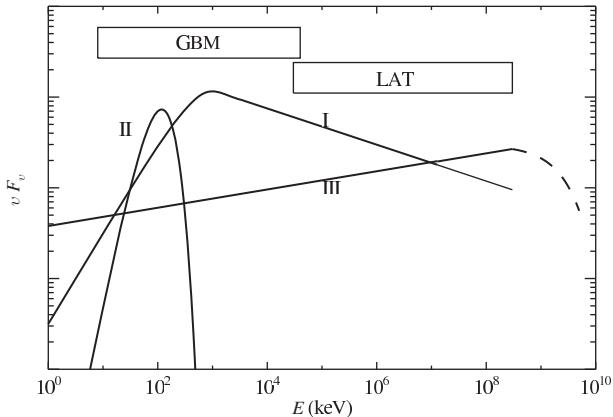


Figure 2.16 The three possible elemental spectral components that shape the observed time-resolved spectra of GRBs. Some components can be suppressed in some GRBs. Adapted from Zhang et al. (2011).

component suggests that it may not originate from the external shock region. However, its exact physical origin is subject to debate, even though some sort of inverse Compton scattering processes are likely at play.

Other Spectral Models

Other spectral forms have been suggested to fit GRB spectra. One is a more general form of the smoothly broken power-law model (e.g. Kaneko et al., 2006; Gruber et al., 2014). It is written in the form

$$N(E) = A \left(\frac{E}{100 \text{ keV}} \right)^b 10^{(a-a_{\text{piv}})}, \quad (2.12)$$

where

$$\begin{aligned} a &= m\Lambda \ln \left(\frac{e^q + e^{-q}}{2} \right), \\ a_{\text{piv}} &= m\Lambda \ln \left(\frac{e^{q_{\text{piv}}} + e^{-q_{\text{piv}}}}{2} \right), \\ q &= \frac{\log(E/E_b)}{\Lambda}, \\ q_{\text{piv}} &= \frac{\log(E_{\text{piv}}/E_b)}{\Lambda}, \\ m &= \frac{\lambda_2 - \lambda_1}{2}, \\ b &= \frac{\lambda_1 + \lambda_2}{2}, \end{aligned} \quad (2.13)$$

with $E_{\text{piv}} = 100 \text{ keV}$.

Compared with the four-parameter Band function, this model has five parameters: the amplitude parameter A , the low-energy and high-energy spectral indices λ_1 and λ_2 , the

break energy E_b , and another free parameter Λ that describes the sharpness of the broken power law (a smaller value corresponds to a sharper break). Practically, this fifth parameter Λ cannot be well constrained.³ An appropriate value is $\Lambda \sim 0.3$ in order to fit the data (Kaneko et al., 2006; Gruber et al., 2014).

Another proposed function is the *log-parabolic model* (Massaro et al., 2010)

$$N(E) = A \left(\frac{E}{E_0} \right)^{-a-b \log(E/E_0)}. \quad (2.14)$$

Such a model is globally curved, and does not have asymptotic power-law indices in both low- and high-energy regimes. Massaro et al. (2010) argued that it can fit some BATSE GRBs. The high-quality, broad-band data of *Fermi* GRBs seem not to support this model.

Spectral Evolution

For bright bursts, significant spectral evolution is usually observed, which provides more clues about the GRB prompt emission mechanism.

One interesting feature is the correlation between E_p and flux in individual GRB pulses. It is found that in general there are two types of evolution patterns. The first type shows a *hard-to-soft evolution* pattern, which means that E_p decreases throughout the pulse, even during the rising phase of the pulse (Norris et al., 1986). The second type shows an *intensity tracking* pattern: E_p tracks the intensity, so that it increases during the rising phase and decreases during the falling phase (Golenetskii et al., 1983). Figure 2.17 shows some examples of E_p evolution patterns (Lu et al., 2012). Clear hard-to-soft evolution patterns are seen in GRBs 081125 and 081224, while clear tracking patterns are seen in GRBs 081207 and 081222. Observationally, both types of behavior can be seen in the same burst (Lu et al., 2010, 2012), see GRB 081221 in Fig. 2.17. The hard-to-soft evolution pattern is more common in the first pulse of the lightcurves, while the intensity tracking pattern is more common in later pulses. Simulations show that the superposition of hard-to-soft evolution pulses may give rise to an apparent tracking behavior (Lu et al., 2012), and arguments have been made that all pulses could be consistent with having a hard-to-soft evolution pattern (Hakkila and Preece, 2011). However, the fact that tracking patterns exist in the first pulse or isolated pulses of some GRBs suggests that both evolution patterns are likely intrinsic.

The first pulse of the nearby bright GRB 130427A, thanks to its extremely high flux and fluence, can be used to study pulse properties. The data show a clear hard-to-soft evolution pattern, consistent with a synchrotron emission origin (Preece et al., 2014).

A fraction of (but not all) LAT GRBs show a delayed onset of GeV emission with respect to MeV emission as shown in Figs. 2.9 and 2.10 (Abdo et al., 2009c,b; Ackermann et al., 2010; Zhang et al., 2011; Ackermann et al., 2013). For at least some GRBs (e.g. 090902B and 090926A), this may be related to the delayed onset of spectral component III.

³ This suggests that the smoothness of the break is not well constrained by the data. The specific exponential transition invoked in the Band function is just one of many ways to connect the two asymptotic power-law segments. This has implications for understanding the physical origin of GRB spectra. See details in §9.10.

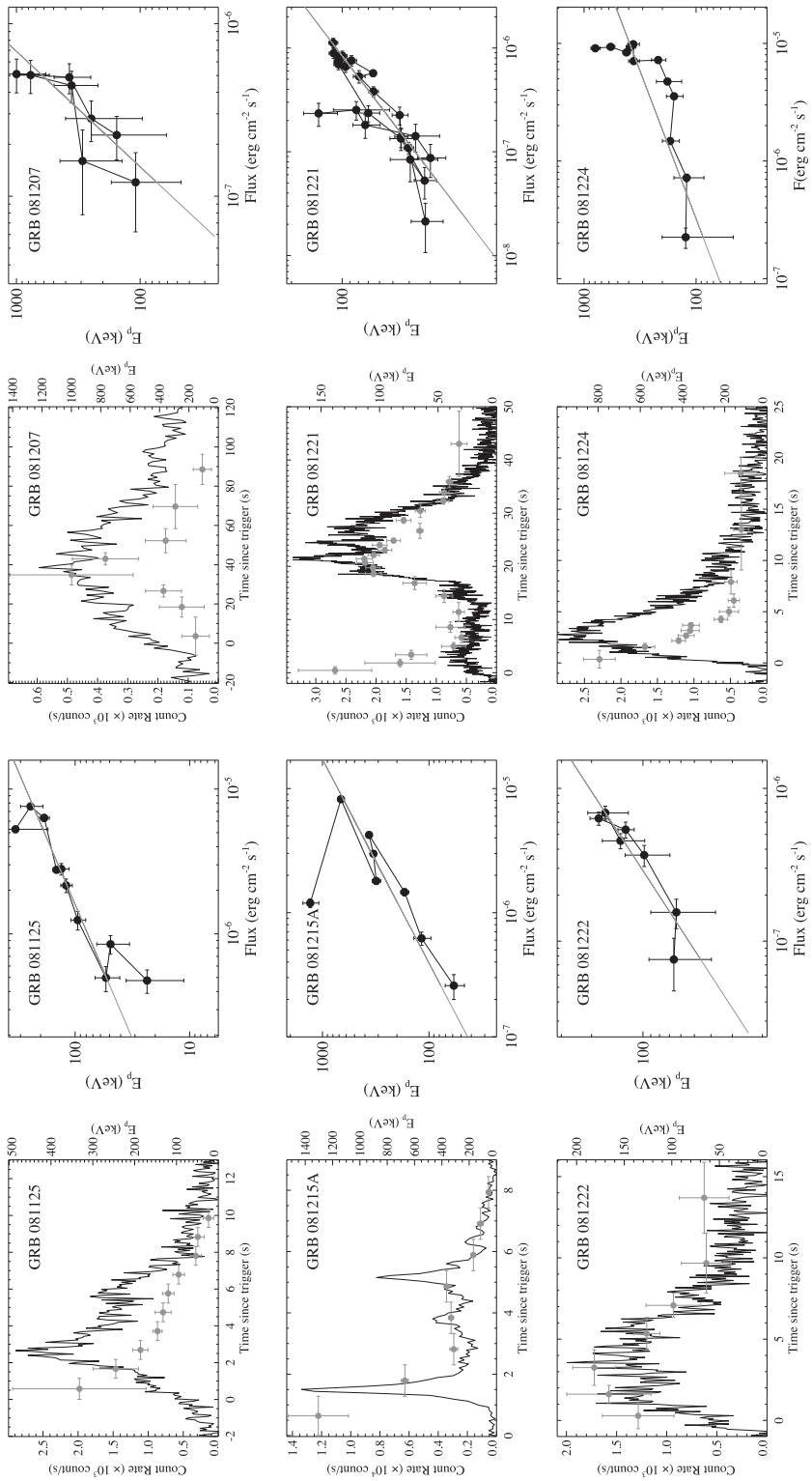


Figure 2.17 Some examples of E_p evolution patterns in GRBs. From Liu et al. (2012).

2.1.4 Broad-Band Emission

During the prompt emission phase, emission outside the bandpass window of the triggering detectors is naturally expected. Observationally, lacking wide-field, broad-band telescopes staring at the same position as the GRB detectors when a GRB randomly shows up in the sky, it is very challenging to obtain a broad-band prompt emission spectrum. Nonetheless, a sparse picture is revealed by the available observations.

In the high-energy regime, *Fermi*/LAT observations so far suggest that most GRBs do not have significant emission beyond 100 MeV. Most of the LAT detections or upper limits are either consistent with the extension of a Band-function spectrum to the GeV regime (Zhang et al., 2011; Ackermann et al., 2013), or require a spectral cutoff between the GBM and LAT band (Ackermann et al., 2012). As discussed earlier, a small group of GRBs (e.g. GRBs 941017, 090510, 090902B, and 090926A) do show a hard component (component III in Fig. 2.16), which sets in later. These sources have significant emission in the LAT band and beyond, which are ideal targets for ground-based 100 GeV – TeV detectors (e.g. Kakuwa et al., 2012; Inoue et al., 2013).

In the low-energy regime, prompt optical observations were made for a small sample of GRBs. Some of these GRBs had a precursor or a very long duration, so that *Swift* XRT and UVOT were able to slew to the source before the main burst finished or even before it started. Examples include GRB 060124 (Romano et al., 2006), GRB 060218 (Campana et al., 2006), and GRB 061121 (Page et al., 2007). For some other bursts, early optical observations were carried out by ground-based robotic telescopes during the prompt emission phase. Examples include GRB 990123 (Akerlof et al., 1999), GRB 041219A (Blake et al., 2005; Vestrand et al., 2005), GRB 050820A (Vestrand et al., 2006), GRB 080319B (Racusin et al., 2008; Beskin et al., 2010), and GRB 110205A (Zheng et al., 2012; Cucchiara et al., 2011b; Gendre et al., 2012).

In rare cases, such as the nearby very bright GRB 130427A at $z = 0.34$ (Maselli et al., 2014), both a prompt optical flash (Vestrand et al., 2014) and a GeV flash (Ackermann et al., 2014) were jointly detected during the prompt emission phase, which roughly coincided in time (Vestrand et al., 2014) (Fig. 2.18). More cases are needed to see whether this is a common feature among GRBs.

There are at least three patterns for the relationship between prompt optical emission and prompt sub-MeV emission (Fig. 2.19). The first pattern shows a clear mis-match between the optical flux peak and γ -ray flux peaks. An example is GRB 990123, which showed an optical peak after all the γ -ray peaks (Akerlof et al., 1999). This suggests a different physical origin for the two components (e.g. an internal emission site for γ -rays and an external reverse shock emission site for the optical emission). The second pattern shows a roughly tracking behavior between the optical and γ -ray lightcurves. It was seen in GRB 041219B with sparse time resolution in the optical data (Vestrand et al., 2005), but was exemplified by the “naked-eye” GRB 080319B with high-quality optical and γ -ray data (even though there is a systematic ~ 3 s lag between optical and γ -ray photons) (Racusin et al., 2008; Beskin et al., 2010). Spectroscopically, although the optical flux of GRB 041219B is consistent with the spectral extension of the γ /X-ray flux to the optical band, that of GRB

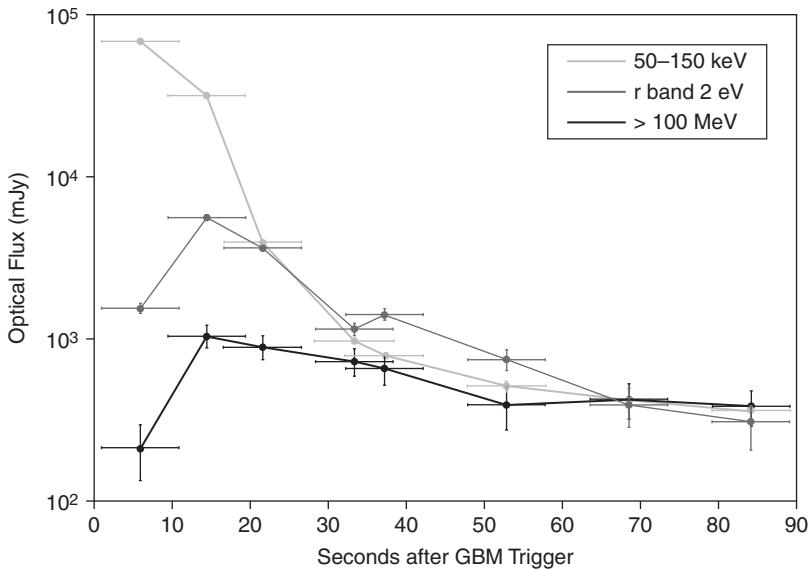


Figure 2.18 Multi-wavelength lightcurves of the nearby bright GRB 130427A, which show a coincident optical and GeV flash. From Vestrand et al. (2014). A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

080319B clearly stands above the spectral extension of the γ /X-ray flux, suggesting a distinct spectral origin of the optical emission (Racusin et al., 2008). The third pattern shows a mix of both (mis-match and tracking) components, as evidenced in GRB 050820A (Vestrand et al., 2006) and GRB 110205A (Zheng et al., 2012). Multiple emission sites have to be invoked to generate these components.

So far, no detection of GRB prompt emission has been made in the radio band. This is partially due to the slow slewing speed of large radio telescopes, and partially due to the lack of theoretical motivation: during the prompt emission phase, radio flux is expected to be strongly suppressed due to synchrotron self-absorption, unless strong coherent emission can be released during the prompt emission phase.

2.1.5 Polarization

Several claims have been made suggesting that the prompt γ -ray emission is linearly polarized with a large degree of polarization Π . Coburn and Boggs (2003) analyzed the *RHESSI* data of GRB 021206 and suggested a polarization degree $\Pi = 80 \pm 20\%$. However, the conclusion was not confirmed by a later independent analysis (Rutledge and Fox, 2004). Later, using the BATSE Albedo Polarimetry System (BAPS) data, Willis et al. (2005) claimed the discovery of linear polarization with $\Pi > 35\%$ and $\Pi > 50\%$ for GRB 930131 and GRB 960924, respectively. Two analyses of the *INTEGRAL* data of GRB 041219A led to evidence of linear polarization, but the significance is only marginal (Kalemci et al., 2007; McGlynn et al., 2007). Later, Yonetoku et al. (2011) claimed the detection of

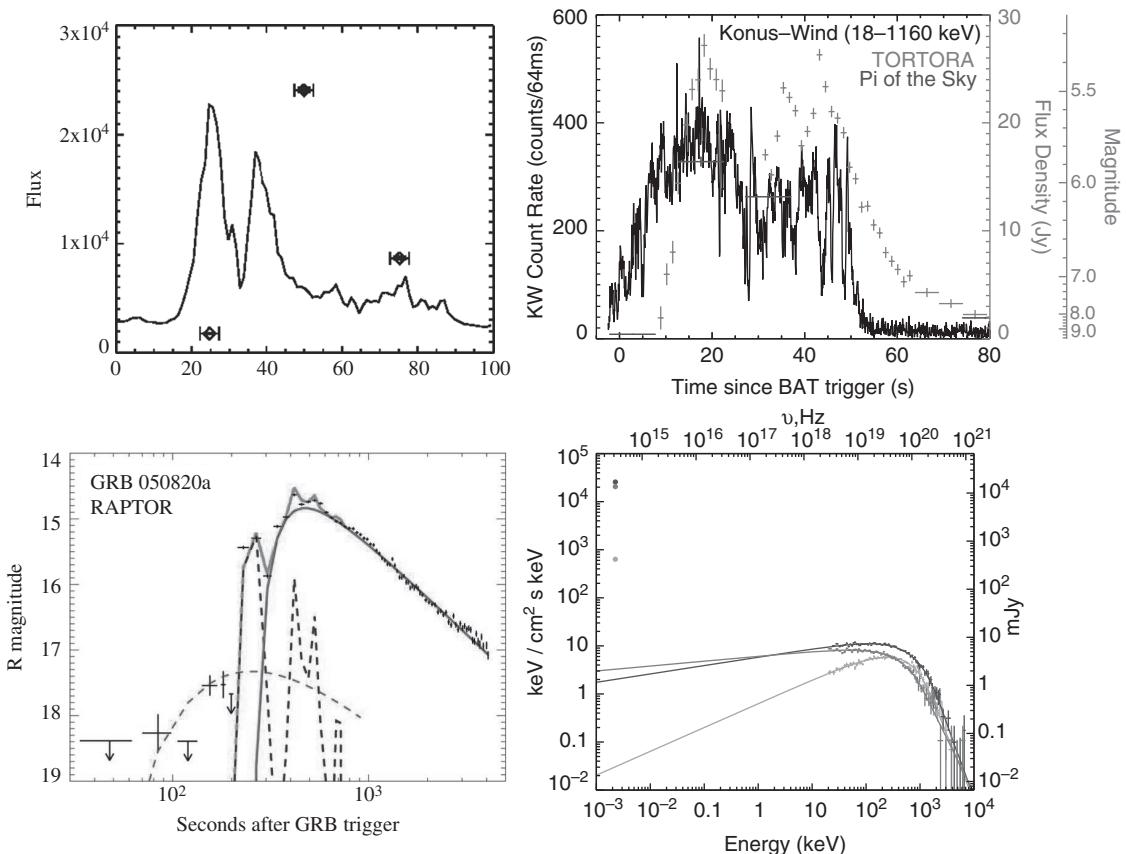


Figure 2.19 Examples of prompt optical emission that show three patterns with respect to the γ -ray emission. *Top left:* GRB 990123 shows an offset of optical peak with respect to the γ -ray emission peak. From Akerlof et al. (1999). *Top right:* GRB 080319B (the “naked-eye” GRB) shows a clear tracking behavior between optical and γ -rays. From Racusin et al. (2008). *Lower left:* GRB 050820A shows the “hybrid” pattern. From Vestrand et al. (2006). *Lower right:* The optical emission of the naked-eye GRB has a distinct spectral component from the γ -rays. From Racusin et al. (2008). The three dots in the optical band (upper left region in the plot) from top to bottom are related to the three curves in the γ -ray band, respectively, with the same top-to-bottom order in terms of the peak flux of the curves. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

$\Pi = 27 \pm 11\%$ with 2.9σ significance during the prompt emission of GRB 100826A using a GAMMA-ray burst Polarimeter (GAP) on board a small Japanese solar-power-sail demonstrator, *Interplanetary Kite-craft Accelerated by Radiation Of the Sun (IKAROS)*. Later, the same team reported detections of high polarization degrees for another two bright GRBs: $\Pi = 70 \pm 22\%$ for GRB 110301A (3.7σ) and $\Pi = 84^{+16}_{-28}\%$ for GRB 110721A (3.3σ) (Yonetoku et al., 2012). For a review of GRB polarization observations, see McConnell (2017).

Even though more detections with a higher confidence level are needed to make a robust claim, these preliminary observational reports hint at the following picture: at

least for some bright GRBs, the prompt γ -ray emission likely carries a high degree of linear polarization. This has profound implications for understanding the unknown jet composition and radiation mechanism of GRB prompt emission. Even though different models can generate polarized γ -rays, a statistical analysis of the polarization properties of several tens of GRBs would provide great constraints on the underlying models (Toma et al., 2009).

An observational campaign of GRB 160625A revealed a detection of significant (up to $\Pi = 8.0 \pm 0.5\%$) and variable linear polarization for the prompt optical emission of the burst (Troja et al., 2017). Since prompt optical emission is believed to have an internal origin similar to γ -rays, this result is consistent with the high polarization degrees claimed in previous γ -ray observations.

2.1.6 Summary

In summary, GRB prompt emission has rich observational features, most of which are not properly understood. Even though GRB prompt emission was discovered starting from the first detected GRB, due to its short duration, it is still a great challenge to observe this phase in all wavelengths (from radio to TeV), which is essential to diagnose many important physical questions related to prompt emission, such as the composition of the jet, the processes of energy dissipation and particle acceleration, and the mechanism of radiation. An ideal observational campaign would be to have wide-field telescopes in all wavelengths, with a large field of view similar to that of the GRB-triggering detectors and with polarization measurement capabilities, to watch the same direction of the sky simultaneously until a bright GRB is detected. Ideally this comprehensive data set should also be supplemented by the detection data or meaningful upper limits provided from the high-energy neutrino and gravitational wave detectors. One would then get a complete picture of GRB prompt emission. Such a “jumbo” observational campaign may not be accomplished in the near future.

2.2 Afterglow

Observationally, the *afterglow* phase of a GRB may be conventionally defined as the *temporal phase after the end of the prompt sub-MeV emission*.

The afterglow emission was actually *predicted* before its discovery (Paczynski and Rhoads, 1993; Katz, 1994; Mészáros and Rees, 1997a). The basic argument is the following: whatever the central engine is, a GRB suddenly delivers a huge amount of energy in a small volume of space, resulting in a *fireball* moving at a relativistic speed if baryon loading is not heavy. Wherever in the universe, there must be a *circumburst medium* (even though the density can be low), which will decelerate this relativistic ejecta, generally through a strong *forward shock* propagating into the medium, but early on also through a *reverse shock* penetrating the ejecta itself. Electrons (and protons) would be accelerated in the shocks, giving rise to bright broad-band non-thermal emission through synchrotron

radiation (and also *synchrotron self-Compton (SSC) emission* in the high-energy band). As the fireball slows down, the strength of the shock reduces, so that the emission softens and fades with time. As a result, a fading afterglow is expected from a GRB.

Therefore, from the theoretical point of view, an afterglow can be defined as *broad-band emission released during the interaction between the fireball ejecta and the circumburst medium*. The emission is supposed to come either from the external forward shock, or from the external reverse shock during the reverse shock crossing phase. However, *Swift* observations suggested that not all observationally defined afterglow emission (i.e. low-frequency emission observed after the prompt emission phase) can be attributed to emission from the external shocks. For example, X-ray flares and “internal” plateaus (see §2.2.2 for details) are likely emission from an “internal” site in the jet, which is powered by *late central engine activities*. In some other cases, it is not straightforward to judge whether the emission is from the external shocks or from an internal emission site. Throughout the book, we will therefore apply the observational definition for the afterglow.

Interestingly, the first X-ray and optical afterglows were discovered on 28 February 1997 for GRB 970228 (Costa et al., 1997; van Paradijs et al., 1997), 18 days after the publication (on 10 February 1997) of the seminal paper by Mészáros and Rees (1997a), who provided detailed, self-consistent predictions for the broad-band afterglow based on the external shock model. The first radio afterglow was discovered for GRB 970508 (Frail et al., 1997) less than 3 months later. Afterglow observations are routinely carried out nowadays, and a great amount of data has been accumulated. Below we summarize the observational properties of the afterglow emission.

2.2.1 Multi-Wavelength Afterglow: GRB 130427A as an Example

The ubiquitous property of a GRB afterglow is its “multi-wavelength” nature. As predicted from the synchrotron/SSC external shock model (e.g. Mészáros and Rees, 1997a; Sari et al., 1998; Dermer et al., 2000a; Sari and Esin, 2001; Zhang and Mészáros, 2001b), the afterglow should cover a very wide frequency range, from low-frequency radio to the TeV range. At any instant, the broad-band afterglow spectrum is supposed to be a broken power law. Fixing a particular frequency, the lightcurve should also be a multi-segment broken power law. Since the strength of the shock reduces as the blastwave decelerates, the lightcurves in all wavelengths are expected to decay (as power laws) at late times (after an initial rising phase). As a result, the afterglow flux density can usually be characterized by

$$F_\nu(t, \nu) \propto t^{-\alpha} \nu^{-\beta}, \quad (2.15)$$

a convention we will adopt throughout the book. One should pay attention to the negative signs before the *temporal decay index* α and *spectral index* β in the exponents, so that the typical values of α and β are usually positive.⁴

Whether an afterglow can be detected in a certain wavelength for a certain GRB depends on the brightness of the afterglow emission (which is related to its luminosity and redshift)

⁴ Recall that in the prompt emission context, α and β are adopted to denote the low- and high-energy spectral indices of the Band function (§2.1.3). Their values are negative.

and the detector’s sensitivity. GRB 130427A, a very luminous GRB at $z = 0.34$ termed a “nearby ordinary monster” (Maselli et al., 2014), was the brightest GRB detected in the afterglow era, which allowed accumulation of the richest multi-wavelength afterglow data set as of the year 2018 (e.g. Maselli et al., 2014; Vestrand et al., 2014; Ackermann et al., 2014; Laskar et al., 2013; Perley et al., 2014; Kouveliotou et al., 2013; van der Horst et al., 2014). Figure 2.20 shows the broad-band afterglow data collected by Perley et al. (2014). From radio all the way to GeV energies, the lightcurves show “well-behaved” (broken) power-law decays (upper panel). The bump feature in the radio band at early epochs may originate from a different emission component, e.g. the external reverse shock. The broad-band *spectral energy distributions* (SEDs, lower panel) at different epochs are well consistent with the broken power-law nature of the afterglow as predicted by theory. The bump feature in the low-frequency range again suggests the existence of another component, e.g. the reverse shock.

In the following we discuss the afterglow properties in different energy bands in turn.

2.2.2 X-ray Afterglow

The late-time X-ray afterglow of some GRBs was observed by *BeppoSAX*. It was not until the launch of *Swift* that the GRB X-ray afterglow was routinely observed. The XRT on board *Swift* can slew to the target within tens of seconds, allowing regular observations of the early X-ray afterglow phase for most GRBs.

Some example X-ray afterglow lightcurves are presented in Fig. 2.21 (from Nousek et al., 2006). The data can be summarized as a *canonical lightcurve* composed of five components (Zhang et al., 2006, see Fig. 2.22). This canonical lightcurve is a “synthetic” one, which means that not every GRB has all five components. On the other hand, for most GRBs the lightcurves may be decomposed into multiple components, each of which can be categorized as one of these five components.

Below we discuss the observational properties of these five components as well as their possible physical origins. Several comprehensive studies of the properties of X-ray afterglows can be found in, e.g. Nousek et al. (2006); O’Brien et al. (2006); Willingale et al. (2007); Zhang et al. (2007c); Liang et al. (2007b, 2008a, 2009); Evans et al. (2007, 2009); Margutti et al. (2013).

I. Steep Decay Phase

This component is the earliest power-law decay segment, commonly observed in GRBs (Tagliaferri et al., 2005). The temporal decay slope is steep, typically in the range of ~ -3 to ~ -10 . When joint XRT/BAT observations are available, it is usually found that the extrapolation of this phase is smoothly connected to the end of the prompt emission (Barthelmy et al., 2005b). This suggests that this phase is the natural “tail” of the prompt emission. A time-resolved spectral analysis (Zhang et al., 2007c) of this segment suggested that a good fraction of GRBs showed a clear hard-to-soft evolution during the steep decay phase.

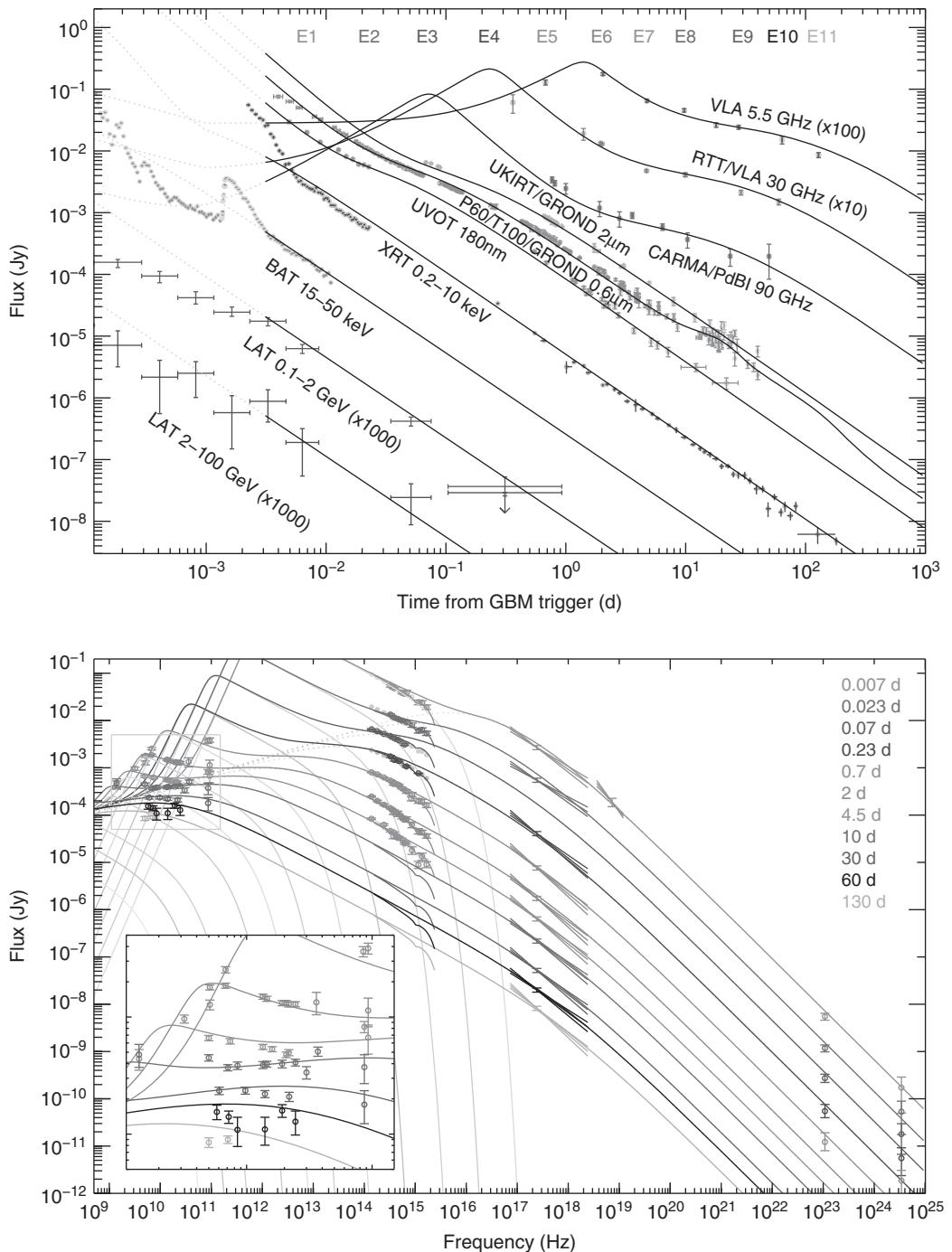


Figure 2.20 Multi-wavelength afterglow lightcurves (*upper panel*) and multi-epoch spectral energy distributions (*lower panel*) of GRB 130427A. Reproduced from Figures 10 and 11 in Perley et al. (2014) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

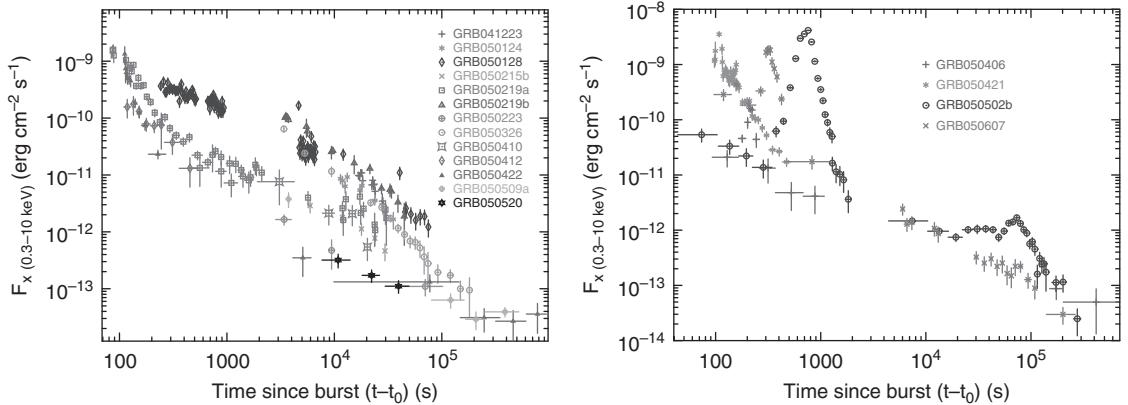


Figure 2.21 Some examples of X-ray afterglow lightcurves detected with *Swift* XRT. Reproduced from Figure 2 in Nousek et al. (2006) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

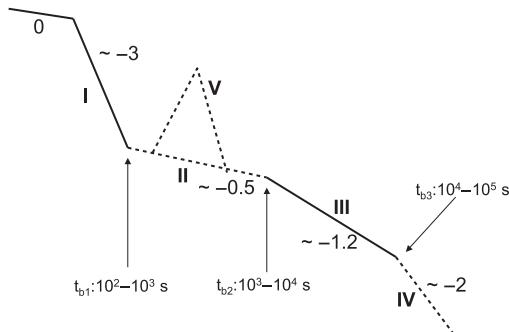


Figure 2.22 A canonical X-ray afterglow lightcurve showing five distinct temporal components: I. steep decay phase; II. shallow decay phase (or plateau if the decay slope is close to 0); III. normal decay phase; IV. post-jet-break phase; V. flares. The segment “0” denotes the prompt emission phase. From Zhang et al. (2006).

Before *Swift*, there was a debate regarding the emission site of GRB prompt emission, i.e. the external shock (Rees and Mészáros, 1992; Mészáros et al., 1993; Dermer and Mitman, 1999) vs. a site “internal” to the jet (the leading candidate being internal shocks, Rees and Mészáros 1994, but alternatives could be the photosphere of the jet, or a site of magnetic dissipation). An efficiency argument (i.e. the external shock model is too inefficient to produce highly variable GRBs since it needs to invoke a clumpy medium with a low filling factor) was raised by Sari and Piran (1997) in favor of the internal shock model. The fact that this steep decay phase is smoothly connected to prompt emission but breaks to a shallower decay phase (which is most likely of the external shock origin) suggests that prompt emission and afterglow are indeed from different emission sites. This settles the debate: since the afterglow has an external shock origin, the prompt GRB emission must arise from an internal emission site (Zhang et al., 2006).

The simplest explanation of the steep decay “tail” emission is that the GRB central engine stops abruptly at the end of prompt emission or turns off with a steeper temporal slope than the observed decline slope. The observed flux is therefore controlled by the so-called “curvature effect”, namely, progressively fainter emission from progressively higher latitudes with respect to the observer’s line of sight arrives at progressively later observational times. Such a model has a well-predicted behavior, i.e. $F_v \propto t^{-(\beta+2)}v^{-\beta}$ (Kumar and Panaiteescu, 2000a; Dermer, 2004) (see §3.4.4 for details). Such a prediction is roughly consistent with the data if one considers the change of decay slope due to an improper choice of the zero time to plot the lightcurve in log-log space (Zhang et al., 2006; Liang et al., 2006b). For this model to work, the GRB emission site should be at a relatively large distance from the central engine (e.g. $R \geq 10^{15}$ cm, Kumar et al. 2007; Hascoët et al. 2012a). Alternatively, if the emission radius is small, the steep decline may be attributed to the intrinsic fading behavior of the GRB central engine power (Fan and Wei, 2005; Barniol Duran and Kumar, 2009).

The strong spectral softening during the steep decay phase (Zhang et al., 2007c) is not expected in the simplest version of the high-latitude curvature effect model, but can be accounted for if the instantaneous spectrum at the end of prompt emission is characterized by a curved spectrum (e.g. a cutoff power-law spectrum or a Band function). Detailed modeling of a sample of GRBs suggests that the high-latitude curvature effect model invoking a curved spectrum can indeed explain the steep decay phase with spectral evolution of at least some GRBs (Zhang et al., 2009b; Genet and Granot, 2009; Mangano and Sbarufatti, 2011; Zhang et al., 2012b).

II. Shallow Decay Phase or Plateau

The *shallow decay phase* typically has a slope from ~ 0 to ~ -0.7 , occasionally with a slight rise early on. It is usually followed by a normal decay phase III ($\sim t^{-1}$). Spectral analyses on segments II and III suggest that there is essentially no spectral evolution across the break (Vaughan et al., 2006; Liang et al., 2007a), suggesting that the break is hydrodynamical or geometrical, but not spectral (i.e. due to the crossing of a spectral break in the X-ray band, which suggests that the spectral indices before and after the break are different). If the slope is close to 0, it is also called a *plateau*. In rare cases, an X-ray plateau can be followed by a very steep decay (e.g. t^{-8} in GRB 070110, Troja et al. 2007, and in some short GRBs as well, Rowlinson et al. 2010, 2013, see Fig. 2.23).

The shallow decay segment followed by the normal decay segment can be interpreted within the standard external forward shock model (Zhang et al., 2006; Nousek et al., 2006; Panaiteescu et al., 2006b), by invoking a *continuous energy injection* into the blastwave, due to either a long-lasting central engine (Dai and Lu, 1998c; Zhang and Mészáros, 2001a), or a stratification of the ejecta Lorentz factor in an impulsively ejected fireball (Rees and Mészáros, 1998; Sari and Mészáros, 2000; Uhm et al., 2012). If a long-lasting reverse shock can outshine the forward shock emission under certain conditions, it can also account for the observed features (Uhm and Beloborodov, 2007; Genet et al., 2007; Uhm et al., 2012; Uhm and Zhang, 2014a). The plateaus followed by a very steep decay phase (Troja et al., 2007; Liang et al., 2007a; Lyons et al., 2010; Rowlinson et al., 2010, 2013; Lü and Zhang, 2014) cannot be interpreted within the framework of the external shock models, and

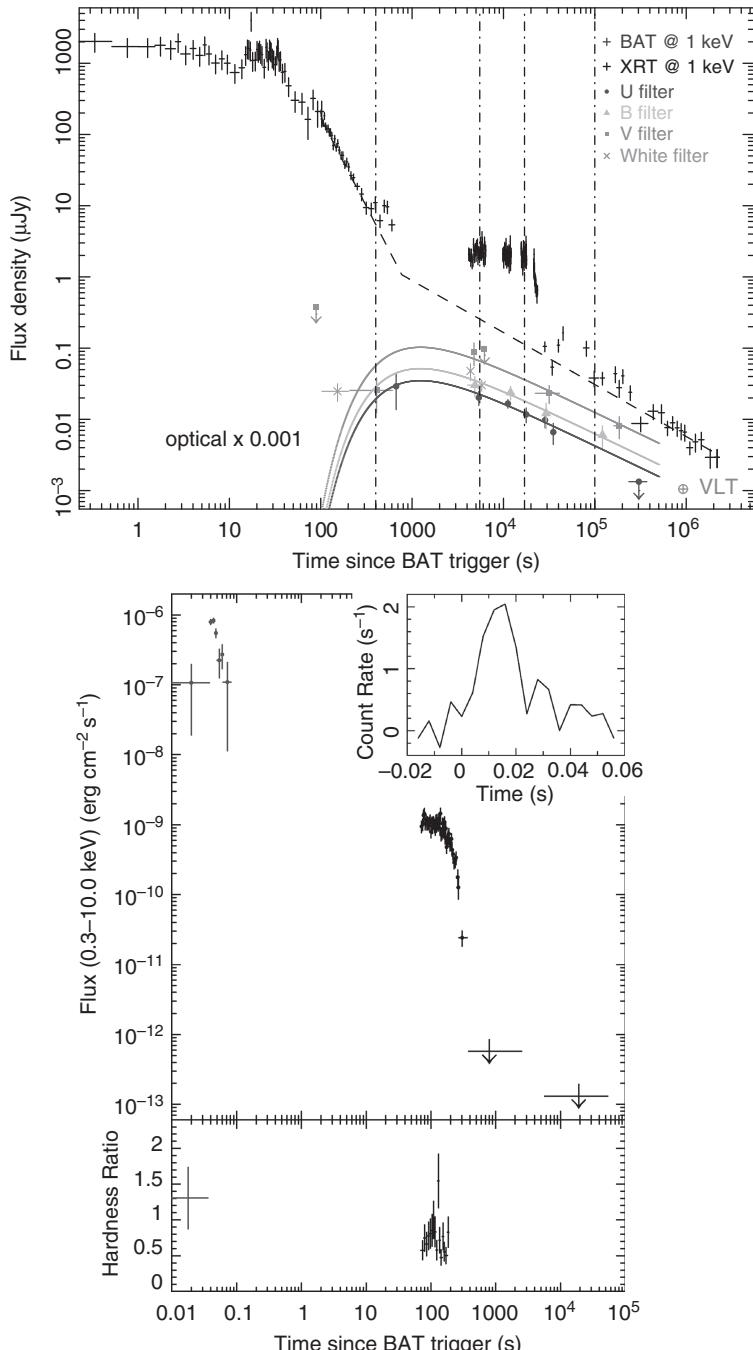


Figure 2.23 The “internal” X-ray plateaus observed in the long GRB 070110 (upper panel; from Troja et al. 2007) and in the short GRB 090515 (lower panel; from Rowlinson et al. 2010). The decay slope after the plateau is too steep to be interpreted within the framework of the external shock model, so the plateau should be of an “internal” origin. This suggests that the central engine launches a long-lasting outflow with steady dissipation until the engine suddenly shuts off.

have to invoke an internal dissipation process (e.g. dissipation of a millisecond magnetar wind). Such a plateau is usually called an *internal plateau* (e.g. Lyons et al., 2010).

III and IV. Steeper Decay Phases Following the Plateau

Segment III has a decay slope of ~ -1 , which is the typical value predicted in the standard external forward shock model. It is therefore considered as “normal”, so the segment is called the *normal decay phase*. Often this normal decay segment steepens to segment IV, with a decay slope of ~ -2 or steeper. This is also expected in the external forward shock model due to the so-called *jet break* effect (§8.4.2). Therefore segment IV may be considered the *post-jet-break phase*.

The internal plateaus (Fig. 2.23, Troja et al. 2007; Rowlinson et al. 2010) firmly suggest that a long-lasting central engine can produce steady emission due to internal dissipation of the outflow energy. Even though the normal segments III and IV can be interpreted as the external shock emission, some authors raised the possibility that the entire X-ray afterglow (including the plateau phase and the following decay phases) can be powered by internal dissipation of a *long-lasting central engine wind* (e.g. Ghisellini et al., 2007; Kumar et al., 2008a,b; Cannizzo and Gehrels, 2009; Lindner et al., 2010). The fact that some GRBs have *chromatic X-ray and optical afterglow* lightcurves (i.e. the lightcurves in the X-ray and optical bands do not show a temporal break at the same time) strengthens such a possibility (e.g. Panaiteescu et al. 2006b; Liang et al. 2007a, 2008a; see §2.2.3 for more discussion). However, detailed studies (e.g. Wang et al., 2015b; Li et al., 2015) suggested that the external shock models can in fact account for the X-ray afterglow emission of most GRBs, although some GRBs (especially those with internal plateaus) indeed require an additional emission component directly powered by the central engine.

V. X-ray Flares

X-ray flares have been discovered in a good fraction of GRBs (Burrows et al., 2005a; Chincarini et al., 2007; Falcone et al., 2007; Chincarini et al., 2010; Margutti et al., 2010). Their lightcurves typically show rapid rise and fall with steep rising and decaying indices. They are “superposed” on a background power-law decay component, which usually shows the same temporal decay index before and after the flare (Fig. 2.24). This suggests that the X-ray flares have a different emission site from the power-law decay segment, and should have an internal origin. Flares are typically narrow. Fitting flares with a Gaussian function, the average value of width-to-peak-time ratio is $\langle \delta t/t \rangle \sim 0.1$ (Chincarini et al., 2007). Compared with the background emission, flares are typically harder, but show a hard-to-soft evolution within the flares (Margutti et al., 2010). For GRBs with multiple flares, the flare peak luminosity also decays with time, on average with $\sim t^{-2.7}$ (Margutti et al., 2011). The number of well-defined flares per GRB ranges from 0 to around 10. Nearly half of *Swift* GRBs have at least one flare, and the average number of flares for GRBs with flares is ~ 2.5 . Most flares happen early (hundreds to thousands of seconds after trigger), but some flares can be very late

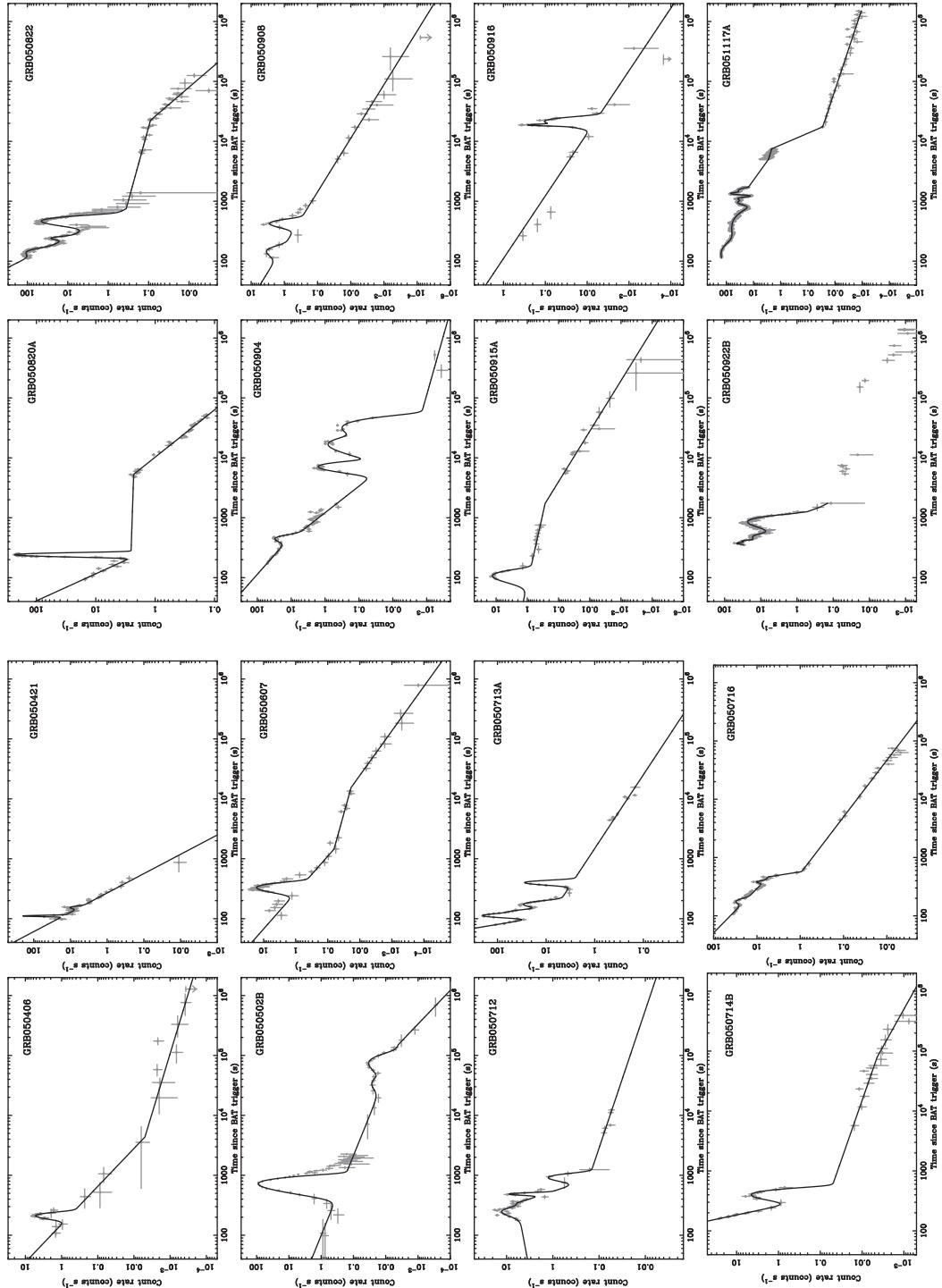


Figure 2.24 A gallery of X-ray flares detected in *Swift* GRBs. From Chincarini et al. (2007).

(e.g. as late as 10^6 s, Falcone et al. 2006).⁵ In several cases, the soft γ -ray counterparts of X-ray flares are detected by *Swift*/BAT. Compared with the fluence of prompt γ -ray emission, the fluence of a flare ranges from less than 1% to comparable or even slightly larger (e.g. in the giant flare in GRB 050502B, Falcone et al. 2006). Figure 2.24 shows a gallery of X-ray flares detected in *Swift* GRBs. One can see a variety of patterns. For example, some bursts (e.g. GRBs 050406, 050607, 050714B, 050915A) have one single, clearly defined flare, which is superposed on either the steep decay phase or the shallow decay phase. This suggests that flares are independent internal events, and have no “knowledge” about the development of the external shock. Some GRBs (e.g. 050712, 050713A, 050822, 050908) have multiple flares with different amplitudes and profiles. Occasionally, the X-ray afterglow is flare dominated (e.g. for the high-redshift GRB 050904) without clear evidence of an underlying power-law afterglow component.

Temporal and spectral analyses of X-ray flares reveal many properties analogous to prompt emission. Margutti et al. (2010) revealed a *luminosity–spectral-lag relation* for both prompt γ -ray emission and X-ray flares, suggesting a direct link between X-ray flares and prompt emission. Peng et al. (2014) showed that many X-ray flares are jointly detected by BAT, which should be regarded as part of prompt emission. A joint spectral analysis with BAT and XRT data by these authors suggests that the Band function can also well describe the spectra of X-ray flares. In some cases, a thermal component with a \sim keV temperature is revealed.

All the evidence suggests that X-ray flares are directly powered by the GRB central engine, similar to prompt emission. They are the extension of prompt emission, delayed, and with reduced amplitudes (Burrows et al., 2005a; Zhang et al., 2006; Fan and Wei, 2005). Direct support for this interpretation comes from the following data analyses. Assuming that the decay phase of X-ray flares is dominated by high-latitude emission, Liang et al. (2006b) searched for the zero point time (T_0) to allow for the simple prediction, $F_\nu \propto t^{-(2+\beta)}\nu^{-\beta}$, from the high-latitude curvature effect model (Kumar and Panaitescu, 2000a), to be satisfied (§3.4.4). They found that the required T_0 are usually associated with the X-ray flares. This is direct evidence that the central engine “restarts the clock” when new outflows are ejected from the engine. Detailed theoretical modeling (Wu et al., 2006; Lazzati and Perna, 2007; Maxham and Zhang, 2009) also supports such an interpretation. Other ideas for the origin of X-ray flares include delayed magnetic dissipation activity as the ejecta decelerates (Giannios, 2006) and anisotropic emission in the blastwave comoving frame (Beloborodov et al., 2011). However, these models may not account for the extremely high luminosity/energy of some flares (e.g. in GRB 050502B, Falcone et al. 2006), and do not straightforwardly account for the T_0 effect revealed by Liang et al. (2006b).

The existence of X-ray flares and internal plateaus suggests that the duration of a GRB is usually (much) longer than what T_{90} records (e.g. Zhang et al., 2014). These cosmic explosions harbor a “dying hard” *long-lasting central engine*, which lasts much longer

⁵ These late “flares” usually do not show very steep rise and decay, so that they may still be accounted for within the external shock model (e.g. Falcone et al., 2006).

than previously believed. By taking into account the duration of the X-ray flares, Zhang et al. (2014) found that the distribution of GRB central engine durations, t_{burst} (defined as the duration from the burst trigger to the end of the last observed X-ray flare), peaks at around several hundred seconds, with more than 10% of bursts having t_{burst} longer than 10^4 s.

A Two-Component Phenomenological Model

An alternative way to describe the GRB X-ray afterglow is a two-component phenomenological model proposed by O'Brien et al. (2006) and Willingale et al. (2007). These authors found that, if one removes X-ray flares and extrapolates the BAT-band γ -ray emission to the X-ray band, most X-ray lightcurves can be fit with a model invoking one or two components, both of which have the same functional form:

$$f_c(t) = \begin{cases} F_c \exp\left(\alpha_c - \frac{t\alpha_c}{T_c}\right) \exp\left(\frac{-t_c}{t}\right), & t < T_c, \\ F_c \left(\frac{t}{T_c}\right)^{-\alpha_c} \exp\left(\frac{-t_c}{t}\right), & t \geq T_c. \end{cases} \quad (2.16)$$

According to this method, the X-ray afterglow can usually be decomposed into a “prompt” component (the prompt emission phase and the subsequent rapid decay phase) and an “afterglow” component (the plateau, normal decay and the late rapid decay). Although no theoretical model predicts the specific mathematical form of these two components, this phenomenological model seems to work well in fitting the X-ray afterglow lightcurves of many *Swift* GRBs and identifying extra features (e.g. X-ray flares or internal plateaus) that demand additional central engine activities (e.g. Lyons et al., 2010).

Polarization

It has been speculated that the X-ray emission components directly powered by the central engine (e.g. X-ray flares and internal plateaus) may be linearly polarized, with a moderately high polarization degree (e.g. Fan et al., 2005c). Lacking an X-ray polarimeter with rapid slewing capability, no X-ray polarization observation of GRBs has been conducted as of 2018.

2.2.3 Optical Afterglow

Late-Time Optical Afterglow

The late-time (later than a couple of hours after the GRB trigger) optical afterglow lightcurves are relatively “regular”, typically having a single power-law decay with a decay index of ~ -1 . If the afterglow is bright enough, one may see a steepening break at a later time. The lightcurve can then be fit with a two-segment broken power law: from a normal decay $t^{-\alpha_1}$ breaking to a steeper decay $t^{-\alpha_2}$, with $\alpha_1 \sim 1$, $\alpha_2 \sim 2$ (e.g. in GRB 990510, see upper left panel of Fig. 2.25, Harrison et al. 1999). This behavior is consistent with the prediction of the standard external forward shock model, with the temporal break defined as the “jet break”.

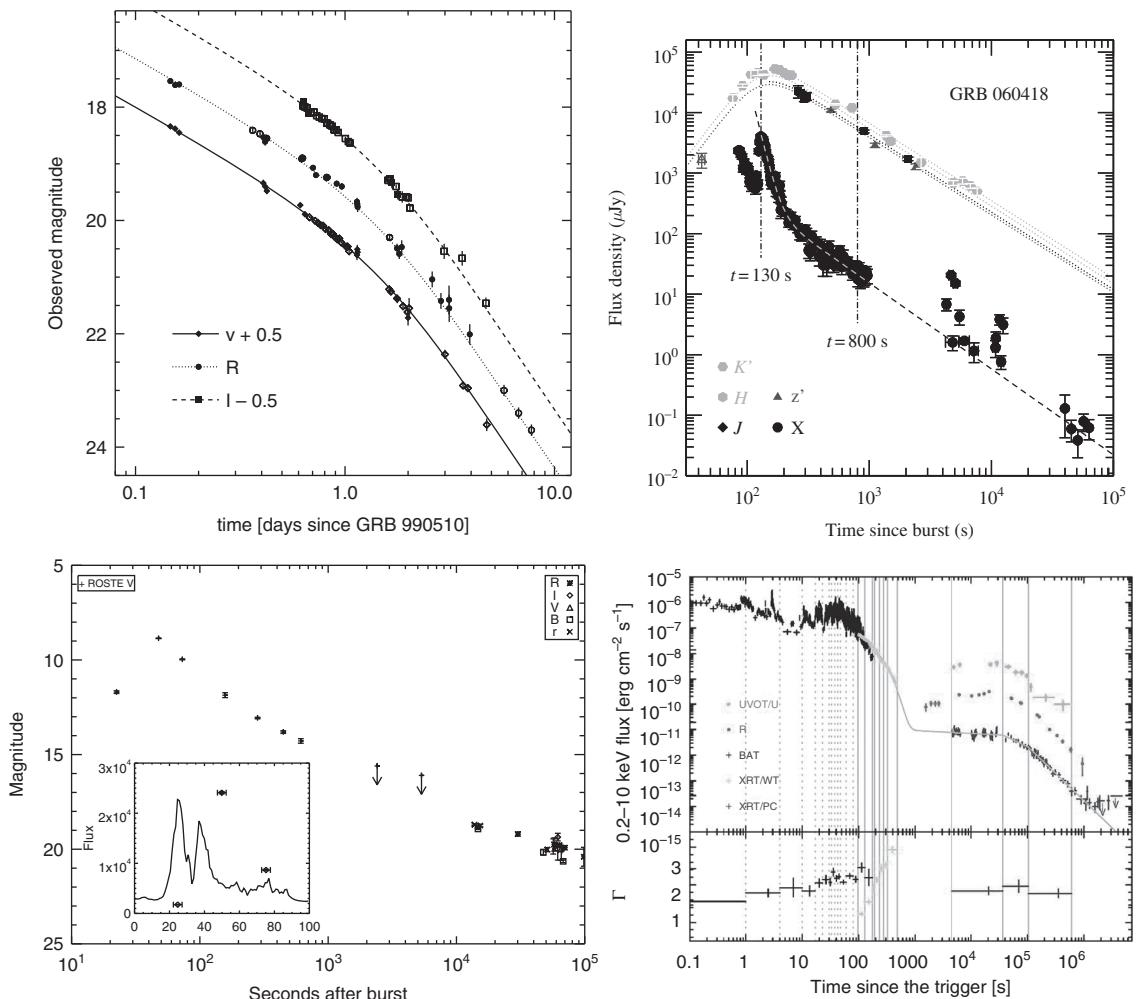


Figure 2.25 Examples of optical afterglow lightcurves that show different patterns. *Upper left:* GRB 990510. Reproduced from Figure 1 in Harrison et al. (1999) with permission. ©AAS. *Upper right:* GRB 060418. From Molinari et al. (2007). *Lower left:* GRB 990123. From Akerlof et al. (1999). *Lower right:* GRB 060614 (the upper two curves in the right part of the figure are optical lightcurves, the lower curve connected to the early prompt emission phase is the X-ray lightcurve). From Mangano et al. (2007).

Early-Time Optical Afterglow

At earlier times (first hours), the optical lightcurves show more complicated behaviors. There are three general patterns. The first pattern, exemplified by GRB 060418 (upper right panel of Fig. 2.25, Molinari et al. 2007), shows a smooth hump at early times, which transitions to a normal decay at late times. This is consistent with the external forward shock model, with the hump explained as the *onset of afterglow* at the blastwave deceleration radius (§8.3). A sample of this type of afterglow was collected by Liang et al. (2013).

The second type, exemplified by GRB 990123 (lower left panel of Fig. 2.25, Akerlof et al. 1999), shows a steeper decay (typically t^{-2} or so) early on, sometimes with a steep rising phase before the steep decay. This type is consistent with a dominant emission from the GRB reverse shock (Mészáros and Rees, 1997a; Sari and Piran, 1999b; Zhang et al., 2003a). A sample of this type of afterglow was collected and modeled by Japelj et al. (2014). The third type, exemplified by GRB 060729 (Grupe et al., 2007) and GRB 060614 (Mangano et al., 2007), shows a shallow decay/plateau (sometimes even with an early slight rise) phase similar to X-ray lightcurves, before breaking to the normal decay phase (lower right panel of Fig. 2.25). A systematic study of early optical afterglow lightcurves with the *Swift* UVOT data was presented by Oates et al. (2009).

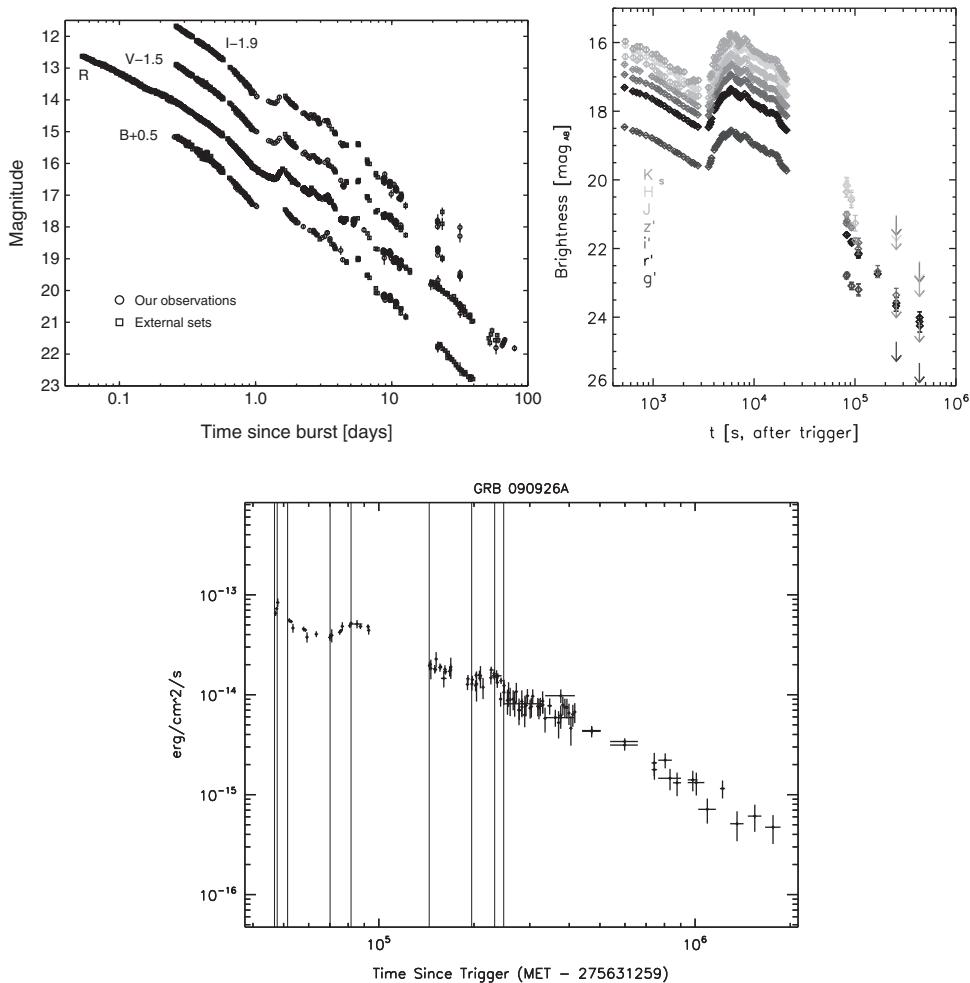
Other Features

For bursts with high-quality data, richer features other than power-law decays have been discovered in the optical lightcurves. Figure 2.26 shows several examples of non-conventional optical lightcurves. Some afterglows show clear *bumps* and *wiggles* that deviate from the simple afterglow model predictions (e.g. GRB 021004 and GRB 030329) (Holland et al., 2003; Lipkin et al., 2004). The upper left panel of Fig. 2.26 displays the wiggling lightcurve of GRB 030329 (Lipkin et al., 2004). The upper right panel of Fig. 2.26 shows a distinct *re-brightening feature* in the multiple optical-band lightcurves of GRB 081029 (Nardini et al., 2011), which is not expected from the simplest afterglow models. The proposed models for interpreting these features include density bumps or voids in the circumburst medium, multiple episodes of energy injection into the blastwave, angular fluctuations in energy per unit solid angle, or the existence of multiple jet components. *Optical flares* were also reported in some GRBs (Swenson et al., 2013), some of which are temporarily correlated with X-ray flares. They are typically less significant than X-ray flares, with less extreme rising and decaying indices. Two examples of optical flares are presented in the lower panel of Fig. 2.26 (from Swenson et al. 2013). Some optical flares may share the same physical origin as X-ray flares due to late central engine activities. Some others are not associated with X-ray flares, which may be of an external shock origin, due to the various non-conventional effects mentioned above. In general, whether an optical flare may be interpreted as of an external shock origin can be determined by comparing the normalized variable time $\Delta t/t$ against the normalized variable flux $|\Delta F_\nu|/F_\nu$ (Ioka et al., 2005).

In the optical afterglow lightcurves of a good fraction of long GRBs, a bump feature, usually with a red color, shows up about a week after the GRB trigger. This is usually interpreted as the signature of an associated supernova (see §2.3 for details).

Lightcurve Gallery and a Synthetic Lightcurve

GRB optical afterglow lightcurves were compiled in the observer's frame and in the rest frame (shifted to a common redshift $z = 1$) by Kann et al. (2010, 2011) (Fig. 2.27). The absolute magnitude of the optical luminosity at 1 day peaks around -23 . Based on some early observations, it was suggested that there might be a possible bimodality of the

**Figure 2.26**

Examples of “non-conventional” optical afterglow lightcurve features. *Upper left:* GRB 030329 shows multiple wiggles. Reproduced from Figure 1 in Lipkin et al. (2004) with permission. ©AAS. *Upper right:* GRB 081029 shows a distinct re-brightening feature in multiple bands (from top to bottom the lightcurves correspond to K, H, J, z', i', r', g' bands, respectively). From Nardini et al. (2011). *Lower panel:* Two “optical flares” detected in GRB 090926A. Reproduced from Figure 1 in Swenson et al. (2013) with permission. ©AAS.

rest-frame optical luminosity at around 1 day (Liang and Zhang, 2006b; Nardini et al., 2006; Kann et al., 2006). The bimodality was weakened and later disappeared as the sample size enlarged (Kann et al., 2010; Zaninoni et al., 2013).

By analogy with the cartoon picture of the canonical X-ray lightcurve (Zhang et al. 2006, Fig. 2.22), Li et al. (2012) attempted to draw a synthetic optical lightcurve, as shown in Fig. 2.28. It turns out that optical lightcurves have richer features and more components. To make a connection to Fig. 2.22, various components in Fig. 2.28 are defined as follows: Ia. prompt optical flares that track the γ -ray emission; Ib. early optical flash that is likely of an external reverse shock origin; III. shallow decay phase (or plateau); III. the standard

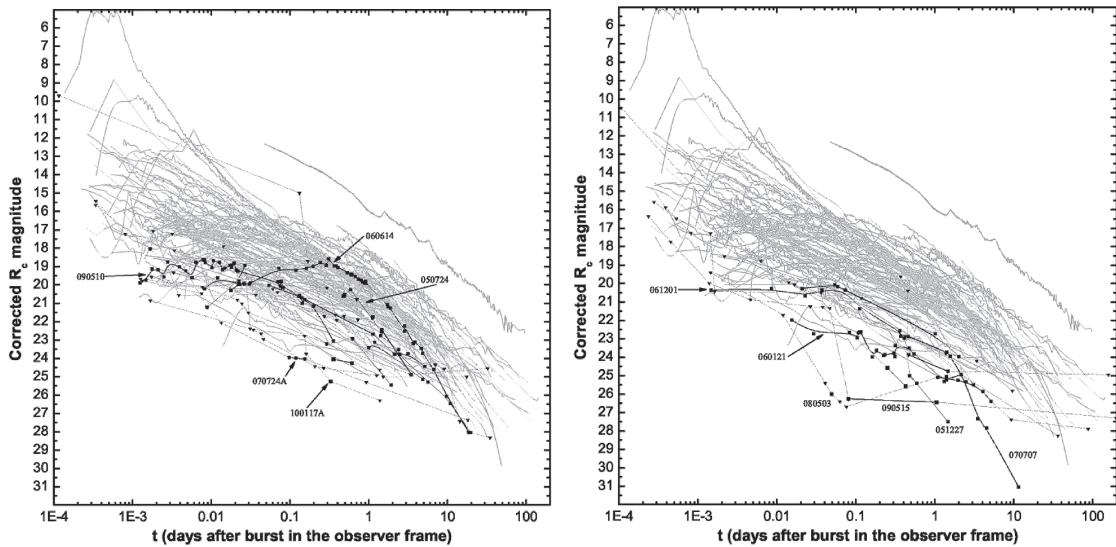


Figure 2.27 *Left:* The observed optical afterglow lightcurves in the observer's frame, magnitudes corrected for extinction. *Right:* The observed optical afterglow lightcurves if all the bursts are placed at $z = 1$. In both figures, gray curves are for the GRBs with a massive star origin (Type II, mostly long GRBs), whereas black curves are for the GRBs with a compact star origin (Type I, mostly short GRBs). From Kann et al. (2011).

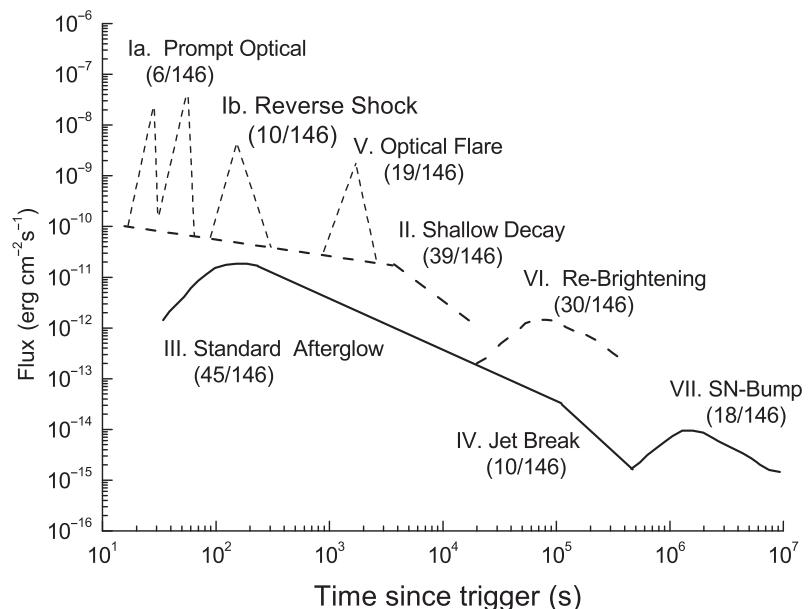


Figure 2.28 A synthetic optical lightcurve. From Li et al. (2012).

afterglow component (normal decay including its early afterglow onset bump); IV. post-jet-break phase; V. optical flares; VI. re-brightening feature occasionally observed in some GRBs; VII. supernova bump.

Polarization

Polarization measurements for late-time optical afterglows typically place an upper limit on, or sometimes give a measurement of, the linear polarization degree of several percent (e.g. Covino et al., 2003). This is consistent with the expectation of the external forward shock model, in which the magnetic fields are believed to be generated through plasma instabilities (e.g. Medvedev and Loeb, 1999; Nishikawa et al., 2005, 2009) or macroscopic turbulence (e.g. Sironi and Goodman, 2007), with a coherent length much smaller than the observable size of the emission region (which is R/Γ in view that only emission within the $1/\Gamma$ cone is bright enough to be observed due to strong relativistic beaming, where R is the emission radius, and Γ is the Lorentz factor of the blastwave) (Gruzinov and Waxman, 1999). Early polarization measurements were made for a handful of bursts in the optical band. Using a ring polarimeter on the robotic Liverpool Telescope, Mundell et al. (2007) placed a 2σ upper limit of $\Pi < 8\%$ for GRB 060418 at 203 s after trigger. The lightcurve has a smooth hump, which suggests a forward-shock-dominated origin. The observational epoch coincides with the peak of the lightcurve. The polarization degree upper limit at this epoch is consistent with that of late-time afterglows, and is also consistent with the theoretical expectation. Observations of another burst GRB 090102 by the same group (Steele et al., 2009) revealed a $\Pi = 10 \pm 1\%$ polarization around 160 s after trigger. The epoch of detection is during a relatively steep decay phase ($F \propto t^{-\alpha}$ with $\alpha = 1.50 \pm 0.06$) before breaking to a more normal decay phase ($\alpha = 0.97 \pm 0.03$) at around 1000 s after trigger. The lightcurve is consistent with a reverse-shock-dominated origin. A relatively high polarization degree may suggest that the emission region carries a significantly ordered magnetic field. This is consistent with the scenario of a magnetized central engine, which gives a bright reverse shock emission in a moderately magnetized ejecta (Fan et al., 2002; Zhang et al., 2003a; Kumar and Panaiteescu, 2003) after significant magnetic dissipation during the prompt emission phase (e.g. Zhang and Yan, 2011). A similar case is GRB 120308A (Mundell et al., 2013). This burst showed $\Pi = 28 \pm 4\%$ in the optical afterglow 4 minutes after the trigger, which decreases to $\Pi = 10^{+5}_{-4}\%$ over the subsequent 10 minutes. The lightcurve, in the meantime, shows a gradual transition from the reverse-shock-dominated phase to the forward-shock-dominated phase (Zhang et al., 2015), consistent with the theoretical expectation. Besides these cases, Uehara et al. (2012) reported a polarization degree of $10.4 \pm 2.5\%$ for GRB 091208B between 149 s and 706 s after the burst trigger. The lightcurve in this case however has a more “normal” decay index $\alpha = 0.75 \pm 0.02$, which is consistent with the prediction of a forward shock. It is puzzling how ordered magnetic fields may be generated in the forward shock, although a proposal was put forward by Uehara et al. (2012). Alternatively, a long-lasting reverse shock can also give rise to the desired afterglow decay index (Uhm et al., 2012). The high polarization degree may then be expected, if emission from a long-lasting reverse shock outshines that from the forward shock.

Optically Dark GRBs

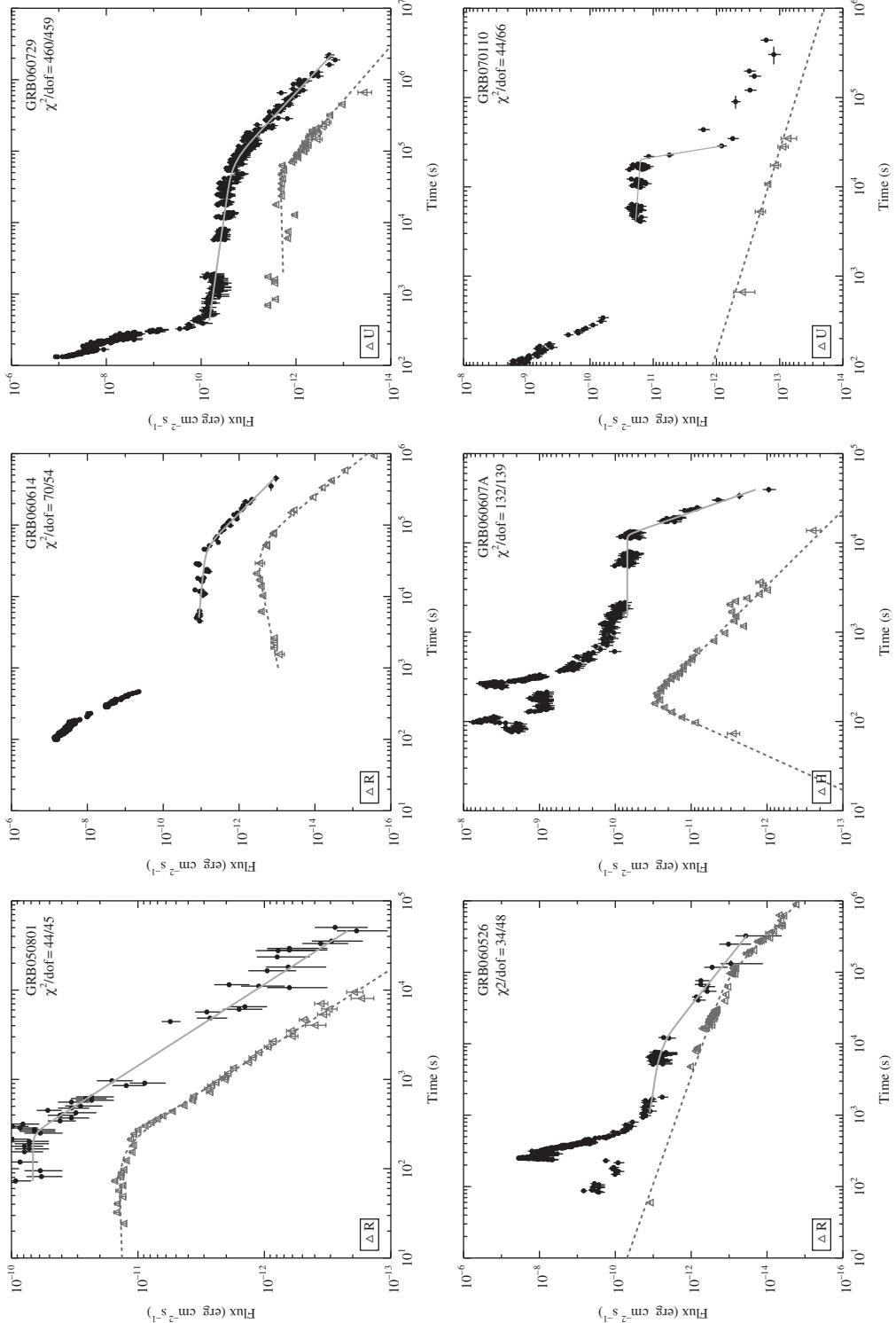
A good fraction of GRBs (30–50%) do not have a detectable optical afterglow. The fraction of these *optically dark GRBs* depends on the epoch of the observations and the sensitivity of the telescopes, so the definition of dark GRBs is subjective. Several attempts have been made to more objectively define optically dark GRBs. Jakobsson et al. (2004) and Rol et al. (2005) defined an optically dark GRB as the GRB whose optical flux or upper limit is “darker” than what is expected from the synchrotron radiation model according to the measured X-ray flux. According to this definition, an afterglow can be “optically dark” even if it is detected, while some others may not be dark even if no detection is made (if the afterglow is faint and the upper limit is not that constraining). Such a definition is relevant if the X-ray and optical afterglows originate from the same emission site. For the GRBs whose X-ray emission is dominated by an internal emission component (X-ray flares or internal plateau), special precaution should be taken in applying this definition.

Leading candidate mechanisms for interpreting optically dark GRBs include dust extinction in the star-forming region where the massive progenitor stars of long GRBs are supposed to reside, and neutral-hydrogen absorption from the intergalactic medium for GRBs at high redshifts. Detailed studies (e.g. Perley et al., 2009) suggest that high- z GRBs account for only a small fraction of the observed dark GRBs. Most dark GRBs are likely caused by heavy dust extinction within the GRB host galaxies.

Chromaticity

Combining X-ray and optical data, one can investigate whether the afterglow behavior from both bands is consistent with being of the same origin. One important test is whether the temporal breaks observed in some GRBs are *achromatic* (i.e. a break occurs at the same time in both optical and X-ray bands) or *chromatic* (i.e. a break occurs in one band but not in the other, or both bands have a break but they occur at different times). According to the standard afterglow model (e.g. Sari et al., 1998), a temporal break can be chromatic as long as the break itself is of a “spectral” origin, i.e. when a spectral break passes across a particular observational band. In this case, one would expect a change in the spectral index across the temporal break in the lightcurve. However, if there is no spectral change across a temporal break (e.g. most X-ray temporal breaks), then the break has to be of a “hydrodynamical” (e.g. energy injection) or “geometrical” (e.g. jet break) origin. Such a break should be achromatic.

A puzzling feature seen in a fraction of GRBs is that the optical and X-ray afterglows are chromatic (Panaitescu et al., 2006a; Fan et al., 2006; Liang et al., 2007b, 2008a; Huang et al., 2007). In some cases there is no temporal break in the optical lightcurve at the epoch when the X-ray lightcurve makes a transition from segment II (plateau phase) to segment III (normal decay phase) or from segment III to IV (jet break phase). In other cases, the optical lightcurve has a break, but at a different epoch from the one in the X-ray band. This is the case particularly when the X-ray lightcurve shows an internal plateau. Figure 2.29 presents a gallery of optical vs. X-ray lightcurves, showing examples of achromatic and chromatic afterglow lightcurves (Liang et al., 2007b).



Examples of GRBs that show achromatic (*upper row*) and chromatic (*lower row*) features in the joint X-ray (*upper*) and optical (*lower*) lightcurves. From Liang et al. (2007b).

Figure 2.29

The fraction of chromatic GRB afterglows depends on the stringency of the achromaticity criteria. If one separately fits the X-ray and optical lightcurves to determine temporal breaks independently, most of the afterglows would be characterized as being “chromatic”, since the obtained break times are usually different (Liang et al., 2007b, 2008a). On the other hand, if one starts with the assumption of “achromaticity” and investigates how the data deviate from this assumption, then the fraction of GRBs that show definite chromatic behavior drops to 10–20%. At least 50% are consistent with being both achromatic and satisfying the *closure relations* of afterglow model predictions (Wang et al., 2015b). The cases with definite chromatic behavior (e.g. GRBs 050730, 060607A, 060526, 070110, lower row of Fig. 2.29) cannot be interpreted within the standard external shock afterglow model. Two different sites need to be introduced to account for the emission from the two bands. Indeed, some of them (GRBs 060607A and 070110) clearly show an “internal X-ray plateau”, which has a very steep decay index at the end of the plateau (§2.2.2), and therefore has a different origin from the external-shock-origin optical afterglow.

2.2.4 Radio Afterglow

The statistical properties of GRB radio afterglows are summarized in Chandra and Frail (2012).

About 30% of GRBs are detected to have radio afterglows. A radio afterglow lightcurve typically shows an early rising phase and reaches a peak around 3–6 days after the trigger at 8.5 GHz, with a median peak luminosity 10^{31} erg s $^{-1}$ Hz $^{-1}$ for long GRBs, and about one order of magnitude or more fainter for short GRBs, X-ray flashes, and nearby SN-associated low-luminosity GRBs. This is consistent with the standard external forward shock model prediction, with the peak corresponding to the crossing of the typical synchrotron frequency ν_m or the self-absorption frequency ν_a .

Some GRBs show evidence of an early *radio flare*. The first instance was GRB 990123 (Kulkarni et al., 1999), which showed a rapid rise and decline of the radio flux, with a peak around 1 day after trigger. Plotting the radio afterglow data of a large sample of GRBs that have radio afterglow detections, Chandra and Frail (2012) discovered an apparent dip around 1 day, suggesting that the earlier emission (the radio flare) may be a distinct emission component. A radio flare was discovered for GRB 130427A with a well-monitored lightcurve (Fig. 2.30, Anderson et al. 2014; van der Horst et al. 2014, see also Laskar et al. 2013; Perley et al. 2014). These early radio flares are usually attributed to the emission from a short-lived external reverse shock (Sari and Piran, 1999b; Kobayashi and Sari, 2000; Kobayashi and Zhang, 2003b), even though a two-component forward shock model was also proposed (van der Horst et al., 2014).

There are theoretical motivations to detect coherent radio emission during or shortly after the prompt emission phase (e.g. Usov and Katz, 2000; Sagiv and Waxman, 2002; Zhang, 2014). Searches for prompt dispersed radio pulse signals from GRBs have been carried out, and only upper limits have been reported so far. Bannister et al. (2012), while reporting an upper limit $1.27 \text{ Jy } w^{-1/2}$ for the putative radio pulses emitted between 200 and 1800 s after 9 GRBs they had observed (with the pulse width w in the range

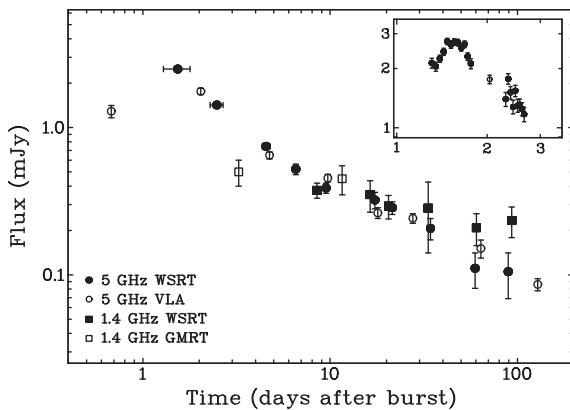


Figure 2.30 Radio afterglow lightcurve of GRB 130427A showing an early radio flare peaking around 1.5 days and a late (forward shock) component peaking around 10 days. From van der Horst et al. (2014).

$6.4 \times 10^{-5} \text{ s}^{-1} < w < 3.2 \times 10^{-3} \text{ s}^{-1}$), did cautiously report possible detections of a dispersed pulse with a duration of several milliseconds following two GRBs: GRB 100704A and GRB 101011A. Both detections are at $\sim 6\sigma$ significance level, which does not meet the criterion for claiming a robust detection. For comparison, the so-called *fast radio bursts* discovered in radio transient surveys (Lorimer et al., 2007; Thornton et al., 2013) typically require greater than 10σ significance in order to claim a detection. Another search for dispersed radio pulses following GRBs also led to negative detections (Palaniswamy et al., 2014).

2.2.5 High-Energy Afterglow

Back in the *CGRO* era, several GRBs detected by BATSE also triggered the high-energy detector EGRET. One famous example was GRB 940217 (Hurley et al., 1994), from which strong GeV emission was still detected 1.5 hours after the trigger when the burst re-emerged from the Earth's limb. The *Fermi* LAT allowed a systematic study of high-energy afterglows of GRBs. Observations suggest that for GRBs jointly detected by *Fermi* GBM and LAT, usually the LAT band (>100 MeV) photons are continuously detected (lasting for $\sim 10^3$ s) after the GBM-band (\sim MeV) emission fades in less than 100 s. By definition, this is the high-energy afterglow of a GRB.

At high energies, the number of photons greatly reduces. As a result, it is not as easy to identify afterglow lightcurve components as in the low-energy bands. Nonetheless, for bright GRBs, a lightcurve can be constructed with reasonable quality. About 9 GRBs per year are jointly detected by *Fermi* LAT and GBM (about 4% of GBM-detected GRBs, Ackermann et al. 2013). The LAT-band afterglow emission typically shows a power-law decay with time (Ghisellini et al. 2010; Zhang et al. 2011; Ackermann et al. 2013, Fig. 2.31). Several investigations suggested that the LAT-band afterglow is generally consistent with emission from the external forward shock (Kumar and Barniol Duran, 2009, 2010; Gao et al., 2009; Ghisellini et al., 2010).

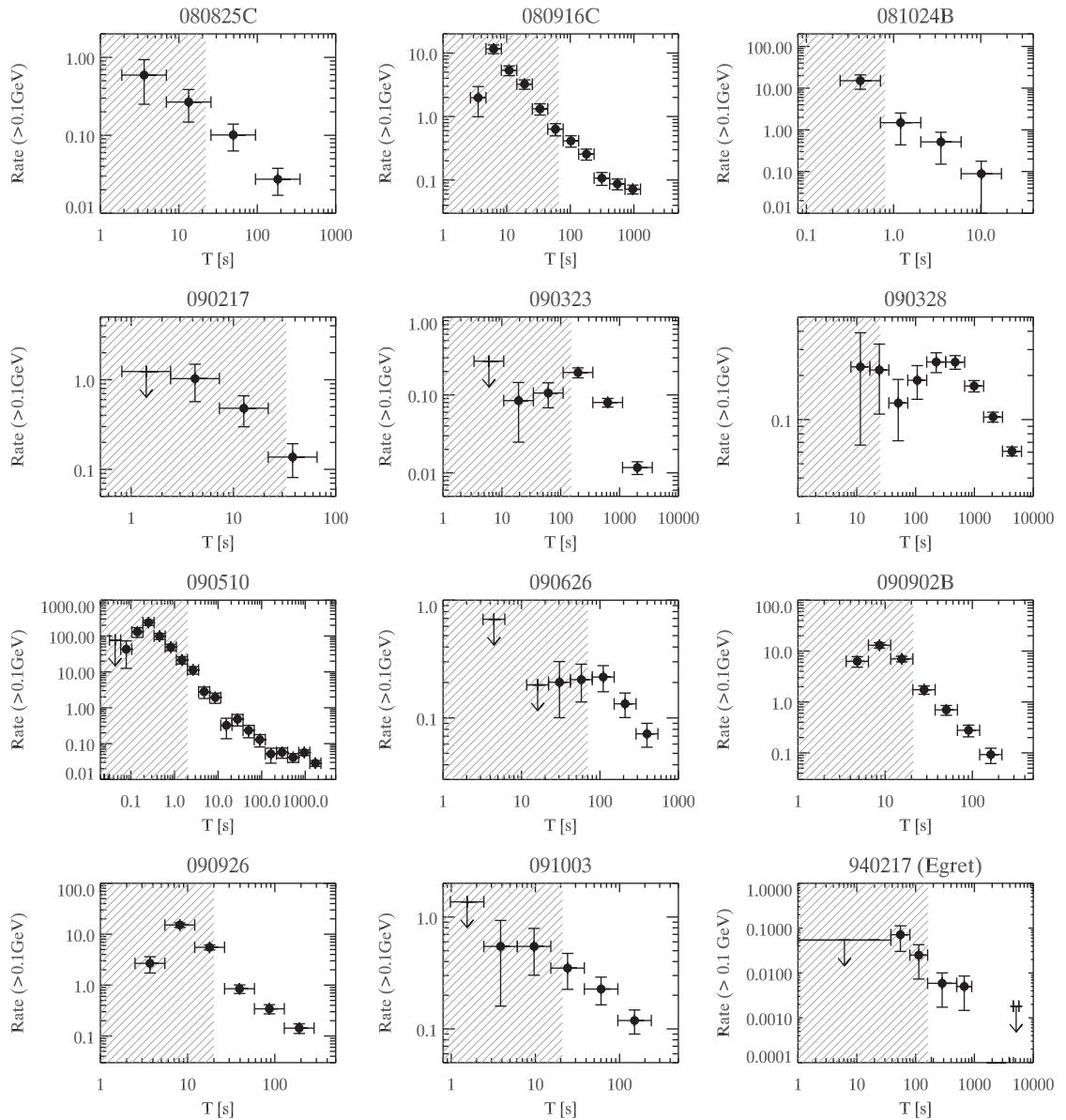


Figure 2.31 High-energy afterglow lightcurves of 16 GRBs as observed by *Fermi* LAT. The shaded regions denote the prompt emission phase. From Ghisellini et al. (2010).

One puzzling fact is that the high-energy afterglow of some GRBs already decays during the prompt emission phase (when emission in the GBM band is still going on). This raises doubts about interpreting the entire GeV emission as emission from the external shock. This is because energy is still being added to the blastwave during the prompt emission phase, so that the predicted lightcurve from the external forward shock should initially

show a shallower decay or even a slight rise (Maxham et al., 2011). Detailed data analyses showed that the spikes in the LAT lightcurves usually track those in the GBM lightcurves (Abdo et al., 2009c; Zhang et al., 2011), suggesting that the LAT-band photons likely have an internal origin during the prompt emission phase. Case-by-case modeling of the LAT-band afterglow lightcurves also cannot account for the early LAT-band emission in the external shock model (He et al., 2011; Liu and Wang, 2011), suggesting that the observed LAT-band emission is a superposition of internal and external emission components. More detailed studies of LAT-band lightcurves revealed that the decay slope at early times is somewhat steeper than that at late times (Ackermann et al., 2013). This is consistent with the superposition picture, with the early steeper decay component attributed to internal emission and the late normal decay component attributed to the standard forward shock emission.

Whereas most LAT-band afterglows are consistent with a synchrotron radiation origin (Kumar and Barniol Duran, 2009, 2010; Gao et al., 2009; Ghisellini et al., 2010), a synchrotron self-Compton (SSC) spectral component has been predicted (Dermer et al., 2000a; Zhang and Mészáros, 2001b), which should give a dominant contribution to the LAT band for a wide parameter space. The highest energy photons from GRB 130427A may demand a SSC emission component (e.g. Fan et al., 2013b; Liu et al., 2013; Tam et al., 2013), even though the case is not crystal clear (e.g. Ackermann et al., 2014; Kouveliotou et al., 2013).

A GeV counterpart of X-ray flares has been predicted (Wang et al., 2006; Fan and Piran, 2006b). A search for *GeV flares* has been carried out in the *Fermi* era. Flaring GeV emission was discovered during an X-ray flare of GRB 100728A (Abdo et al., 2011), which is consistent with being the counterpart of the X-ray flare (He et al., 2012).

2.2.6 Summary

In summary, multi-wavelength observations of GRB afterglows have led to the following picture. In general, the multi-wavelength afterglows have a main component originating from the external forward shock, which is characterized by decaying (including an initial rising phase in low frequencies) broken power-law lightcurves. *However, at least for the initial few hours, the so-called “afterglow” is not simply the external forward shock emission. Instead, it is a superposition of multiple emission components.* Theoretically, there are at least three natural emission sites discussed in the literature: the external forward shock, the external reverse shock, and an internal dissipation site within the late-time outflow launched due to late central engine activities. It is likely that all three sites are contributing to the observed afterglow emission, although emission from different sites may contribute differently in different observational bands, at different epochs, and in different GRBs. Since data robustly suggest erratic late central engine activities (evidenced by, e.g., X-ray flares and internal X-ray plateaus), all three emission sites are relevant, and their emission signatures are entangled in the observational data. All these factors should be properly considered in theoretical modeling.

2.3 Supernova/Kilonova Associations

2.3.1 Long GRBs

Supernova Classification

Supernovae (SNe) are usually classified according to their spectral properties (Filippenko, 1997). Figure 2.32 displays the classification scheme, with the *GRB-associated SNe*, a special type of *broad-line Type Ic*, highlighted. First, based on whether or not hydrogen lines are detected, SNe can be classified as *Type I* (no hydrogen) or *Type II* (yes hydrogen). Next, within Type I, if a singly ionized silicon line (Si II at 615.0 nm) is detected, it is *Type Ia*. If not, one looks for a non-ionized helium line (He I at 587.6 nm). If the He line is detected, it is *Type Ib*. If the He line is missing (or very weak), it is *Type Ic*. Within Type Ic, one may further classify a SN based on the widths of other spectral lines: *narrow-line Type Ic* and *broad-line Type Ic*. It turns out that at least some long GRBs are associated with some (not all) broad-line Type Ic SNe.

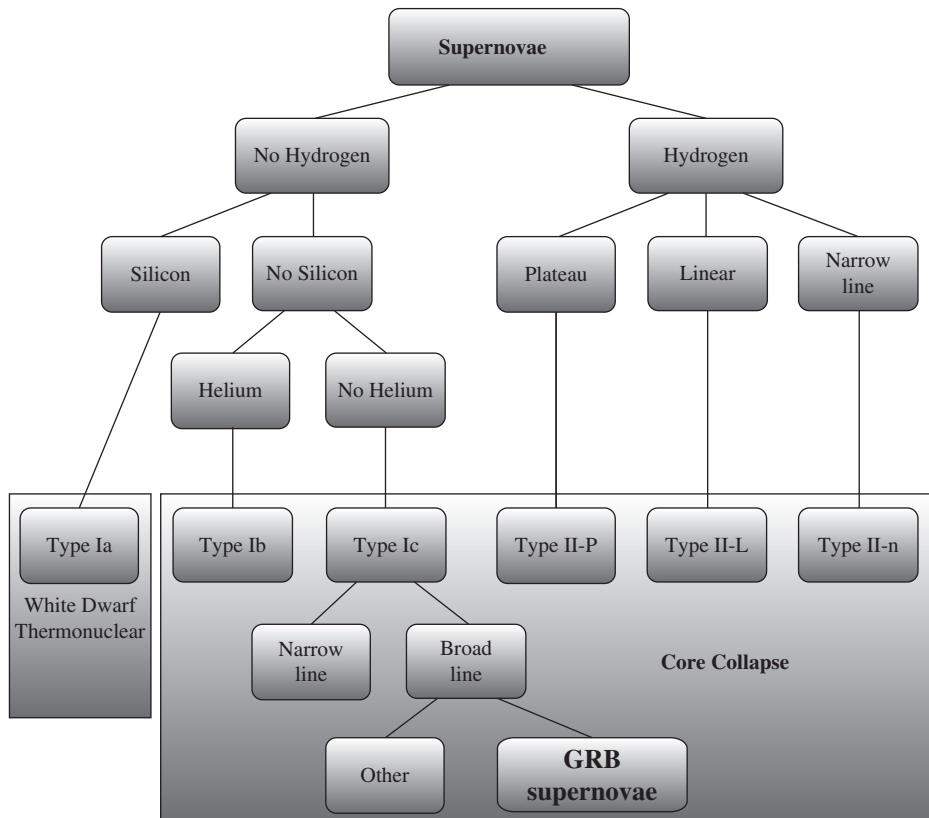


Figure 2.32 Supernova classification scheme and the one associated with long GRBs. Figure courtesy Jared Rice.

Table 2.1 Some spectroscopically identified GRB–SN associations

GRB	SN	z	References
980425	1998bw	0.0085	Galama et al. (1998)
030329	2003dh	0.168	Stanek et al. (2003); Hjorth et al. (2003)
031203	2003lw	0.105	Malesani et al. (2004)
060218	2006aj	0.033	Pian et al. (2006); Campana et al. (2006)
100316D	2010bh	0.059	Starling et al. (2011)
101219B	2010ma	0.55	Sparre et al. (2011)
120422A	2012bz	0.283	Melandri et al. (2012)
130427A	2013cq	0.34	Xu et al. (2013)

Physically, SNe are generated via massive star *core collapse* or runaway *thermonuclear burning* explosions from accreting white dwarfs. Type Ia SNe are white-dwarf-related SNe, since a white dwarf does not have a hydrogen envelope (so no hydrogen lines in the spectrum), but is silicon rich (from the white dwarf itself). All other types of SNe are associated with massive star core collapse. Stars with mass greater than about $8M_{\odot}$ can develop a large enough iron core to undergo catastrophic collapse, resulting in a neutron star or a black hole and a supernova. Near the end of their lives, massive stars usually develop strong winds before going off as a supernova. Depending on the strength of the wind, different layers of the stellar envelope may be stripped. Type II SNe are produced by those stars that are not massive enough and whose winds are not strong enough to blow away the outermost hydrogen envelope, so that hydrogen lines can be observed in the SN spectra. For a more massive star with moderate to high metallicity, a strong wind can develop, which strips off the outer hydrogen envelope(s) of the star. These stars are known as *Wolf–Rayet stars*. For Wolf–Rayet stars that still accommodate helium in their atmospheres, a Type Ib SN is produced after core collapse. For more extreme cases, even the He layer is stripped away before core collapse happens. These Wolf–Rayet stars would give rise to Type Ic SNe.

During a supernova, the ejected material moves in different directions and results in dispersion of the radial velocity. This causes a “Doppler broadening” of the lines. The higher the ejecta velocity, the larger the radial velocity dispersion, and hence the broader the lines. Within Type Ic, a small fraction are broad-line SNe (suggesting energetic explosions). Out of those, only a fraction are observed to be associated with long GRBs. The reason why a GRB progenitor star is so special is unclear, but angular momentum at the core may be a key parameter to determine whether a collimated relativistic jet can be launched.

Spectroscopically Identified GRB–SN Associations

The smoking gun signature of a supernova associated with a GRB is the detection of the characteristic SN spectral features in the optical band. A handful of long GRBs have iron-clad associations with spectroscopically identified SNe. Table 2.1 gives some examples.

Figure 2.33 shows the time-evolving (from 5 to 33 days after the GRB) spectra of the broad-line Type Ic SN 2003dh associated with GRB 030329, as compared with

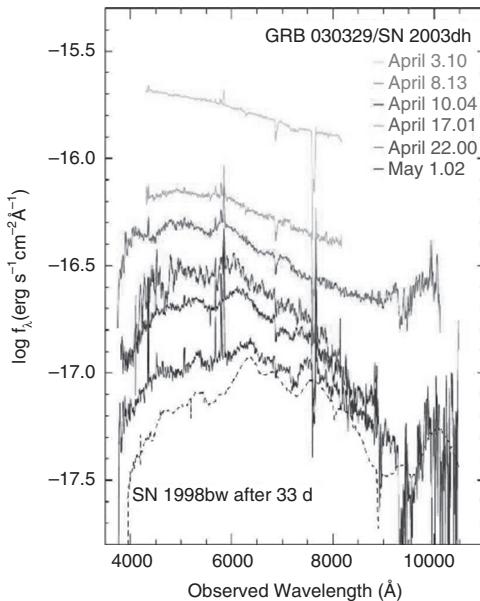


Figure 2.33 Spectra of GRB-associated broad-line Type Ic SN 2003dh in six different epochs (denoted in the upper right corner of the plot following the same order from top to bottom) compared with the spectrum of 1998bw at 33 days. From Hjorth et al. (2003).

the spectrum of SN 1998bw associated with GRB 980425 at 33 days (Hjorth et al., 2003).

A supernova lightcurve usually peaks at \sim 1–2 weeks after the explosion, when the bright optical afterglow (which decays with a power law) has already faded significantly. Therefore a strong SN signature requires two conditions: first, the GRB needs to be nearby (e.g. $z < 0.6$) to allow the SN peak to be bright enough to be detectable; second, the optical afterglow needs to be relatively faint. The latter condition favors GRBs with relatively low luminosities. Indeed, most well-studied SN-associated GRBs are low-luminosity GRBs. These GRBs are typically longer, softer, with a smoother lightcurve. They have a higher local event rate density than the classical high-luminosity GRBs and might have a different physical mechanism from them (e.g. Campana et al., 2006; Soderberg et al., 2006; Liang et al., 2007a; Virgili et al., 2009; Bromberg et al., 2012). Nonetheless, two high-luminosity GRBs have been detected with spectroscopically identified SN associations: GRB 030329 at $z = 0.168$ (Stanek et al., 2003; Hjorth et al., 2003) and GRB 130427A at $z = 0.34$ (Xu et al., 2013), suggesting that most, if not all, high-luminosity GRBs are probably associated with broad-line Type Ic SNe. The low rate of detection is likely due to the relatively high redshifts and bright optical afterglows of most high-luminosity GRBs.

Supernova “Red Bumps”

There are a lot more cases of GRB–SN associations claimed in the literature. Even though the SNe are not spectroscopically identified, a lightcurve bump has been observed in the

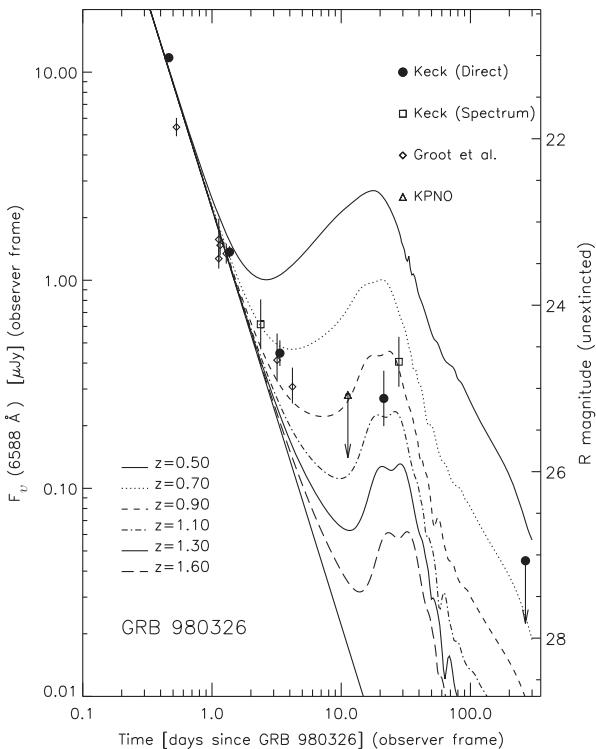


Figure 2.34 A supernova “red bump” in the optical afterglow of GRB 980326. From Bloom et al. (1999).

optical lightcurves of many long GRBs around 1–2 weeks after the GRB trigger. These bumps are relatively “red” (thermal-like spectrum) compared with the afterglow (which has a non-thermal synchrotron spectrum), as is expected for a supernova. Figure 2.34 shows an example of a red bump observed in GRB 980326 (Bloom et al., 1999). Using the multi-color lightcurves of SN 1998bw (the one associated with GRB 980425) as a template, Zeh et al. (2004) found many GRB–SN association candidates, and claimed that essentially every long GRB optical lightcurve may contain light from an underlying supernova. Hjorth and Bloom (2012) reviewed GRB–SN association candidates before 2011, and gave a rank for each association based on the strength of the evidence. Cano et al. (2017) summarized the properties of GRBs and SNe of the confirmed GRB/SN associations before 2016.

Supernova-less Long-Duration GRBs

While data and models before 2006 were consistent with the hypothesis that “ALL long-soft GRBs are accompanied by SNe of Type Ic” (Woosley and Bloom, 2006), two astonishing discoveries in 2006 changed the story.

GRB 060614 was a long-duration burst ($T_{90} \sim 100$ s) at $z = 0.125$ (Gehrels et al., 2006). At such a low redshift a bright SN was expected and should have been detected. After close scrutiny by numerous telescopes worldwide for an extended period of time, no

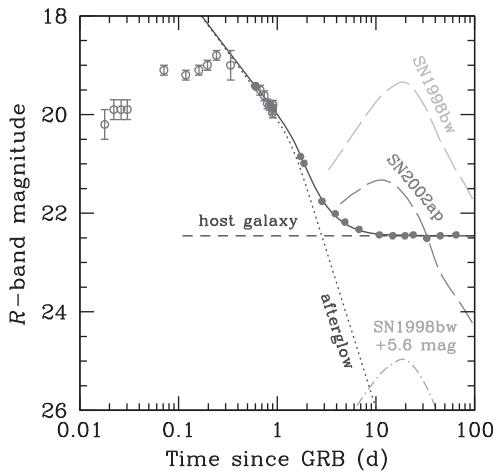


Figure 2.35 Deep upper limit of SN light from long GRB 060614. From Della Valle et al. (2006).

SN light was detected from the source (Gal-Yam et al., 2006; Fynbo et al., 2006a; Della Valle et al., 2006). The upper limit of the luminosity of the SN, if any, is 2 orders of magnitude fainter than the peak luminosity of SN1998bw (Fig. 2.35). A second, nearby, moderately long burst, GRB 060505 ($T_{90} \sim 4$ s), was also found not to be associated with a SN (Fynbo et al., 2006a).

The physical reason for the lack of a SN associated with those two GRBs is subject to debate. Some authors suggested that, under certain conditions, a core-collapse event cannot launch a successful supernova (Tominaga et al. 2007, in connection with the early “failed” SN idea of Woosley 1993). However, several properties of GRB 060614, e.g. a 5-second hard spike followed by softer extended emission with nearly zero spectral lag (Gehrels et al., 2006), and its analogy with the “smoking gun” short GRB 050724 (Zhang et al., 2007b), suggest that it likely belongs to the physical category of most short GRBs, i.e. “Type I” GRBs that are not associated with massive stars. If so, the lack of a SN signature is the natural expectation. A more detailed discussion on the GRB classification schemes is presented in §2.7.

Properties of SNe Associated with GRBs

The spectroscopically identified SNe associated with GRBs all belong to Type Ic. On the other hand, not all Type Ic SNe have GRB associations. A systematic radio survey of Type Ibc SNe suggests that less than 3% of Type Ibc SNe are associated with GRBs (Soderberg, 2007). This suggests that GRBs must invoke a special type of progenitor. The GRB-associated SNe are consistent with being broad-line Type Ic, suggesting a large kinetic energy. As shown in Fig. 2.36 (Pian et al., 2006), they have diverse peak brightness, rise time, lightcurve width, and spectral broadness. Compared with regular Type Ic SNe, the GRB-associated SNe appear to represent the brighter end of the Type Ic population. However, when non-detections and upper limits on SN light are taken into account, the

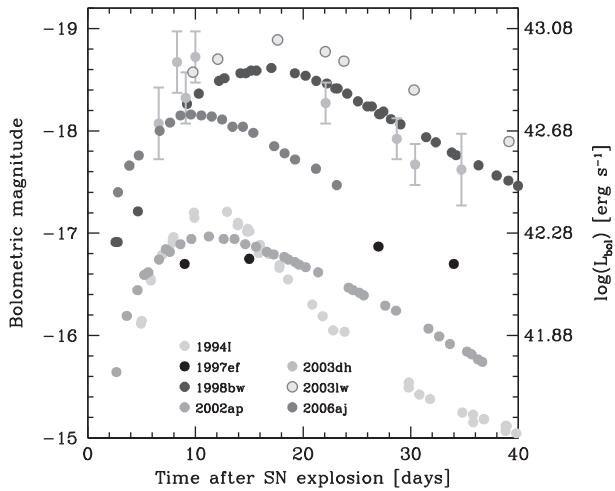


Figure 2.36 Lightcurves of several GRB-associated Type Ic SNe (1998bw, 2003dh, 2003lw, 2006aj) compared with other Type Ic SNe (1994I, 1997ef, 2002ap). From Pian et al. (2006). A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

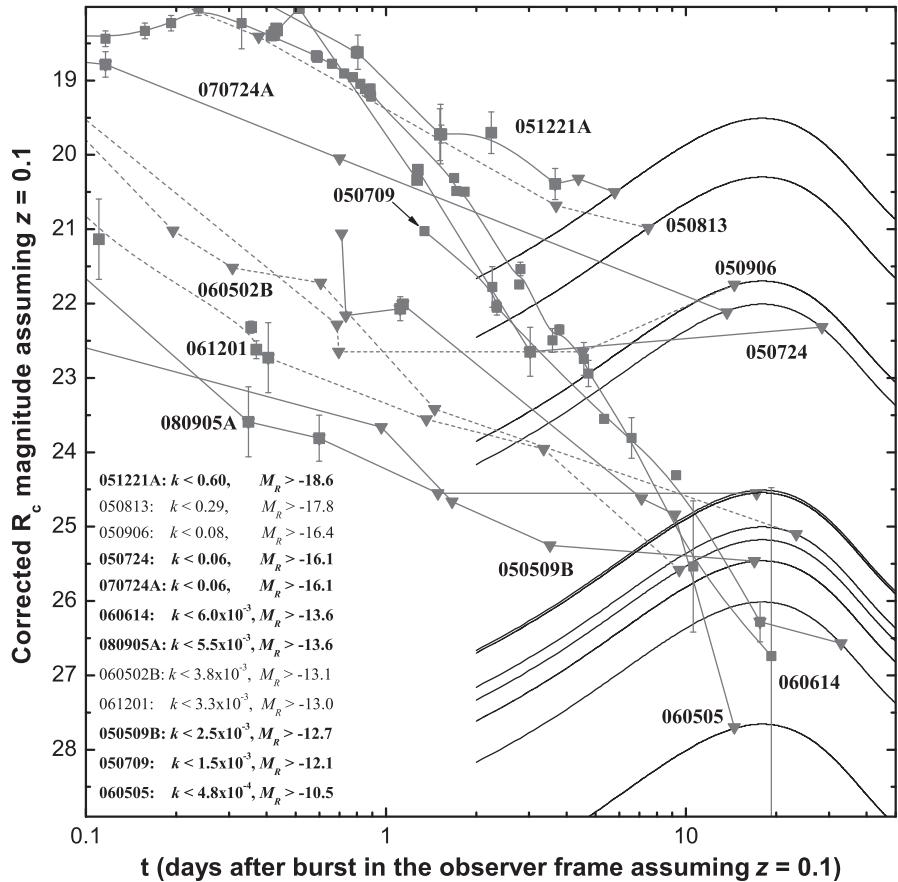
GRB-associated Type Ic SNe may not be special compared with normal Type Ic SNe (Woosley and Bloom, 2006).

2.3.2 Short GRBs

No Association with SNe

Deep searches of possible associated SN light have been carried out for all nearby short GRBs. The upper limits vary from case to case (e.g. Kann et al., 2011; Berger, 2014), but so far no positive detection has been made. Figure 2.37 shows some short GRB afterglow lightcurves superposed with the SN 1998bw-like supernovae with certain stretching factor s and k -correction factor k (Kann et al., 2011). One can see that the observations clearly ruled out the association of SNe with short GRBs. Figure 2 of Berger (2014) shows the upper limits of SN light of short GRBs as compared with the brightness of long-GRB-associated SNe. The non-detection of a bright SN is consistent with a compact star origin (rather than massive star origin) of these GRBs.

It is worth rephrasing that the nearby long GRBs 060614 and 060505 both have deep upper limits of supernova light (Fig. 2.37). However, other observational properties (e.g. a relatively short hard spike, a short spectral lag, a low specific star formation rate at the burst site) of GRB 060614 make it more consistent with belonging to the compact star GRB category (Gehrels et al., 2006; Gal-Yam et al., 2006; Zhang et al., 2007b). Indeed, Zhang et al. (2007b) showed that it would look rather similar to the “smoking gun” compact star GRB 050724 (which lies in an early-type, elliptical galaxy, Barthelmy et al. 2005a; Berger et al. 2005b) if it were somewhat less energetic. The case of GRB 060505 is more controversial, but it is by no means a typical long GRB.

**Figure 2.37**

Deep late detections or upper limits of short GRB afterglows (shifted to $z = 0.1$), compared with the R-band lightcurve of SN 1998bw allowing certain stretching factor s and k -correction factor k . From Kann et al. (2011).

Macronova/Kilonova

A fainter-than-supernova optical/IR transient has been predicted to be associated with NS–NS or NS–BH mergers (Li and Paczyński, 1998; Kulkarni, 2005; Metzger et al., 2010; Barnes and Kasen, 2013; Yu et al., 2013; Metzger and Piro, 2014). The mechanism to power such a transient is rapid neutron capture (r -process) and radioactive decay of the synthesized heavy elements in the neutron-rich ejecta launched during the merger events (e.g. Freiburghaus et al., 1999; Hotokezaka et al., 2013), sometimes with additional energy injection from a post-merger central engine. There is however no standard name for these transients. Kulkarni (2005) suggested using the term *macronova* to reflect that these events are (much) brighter than typical “novae” (produced in accreting white dwarfs) but are (much) dimmer than “supernovae”. Metzger et al. (2010) performed detailed calculations to pin down the previously assumed (Li and Paczyński, 1998) radiative efficiency of these events, and found that the brightness is roughly a thousand times brighter than novae ($L \sim 10^{41} \text{ erg s}^{-1}$). They therefore suggested naming these events *kilonovae*. In

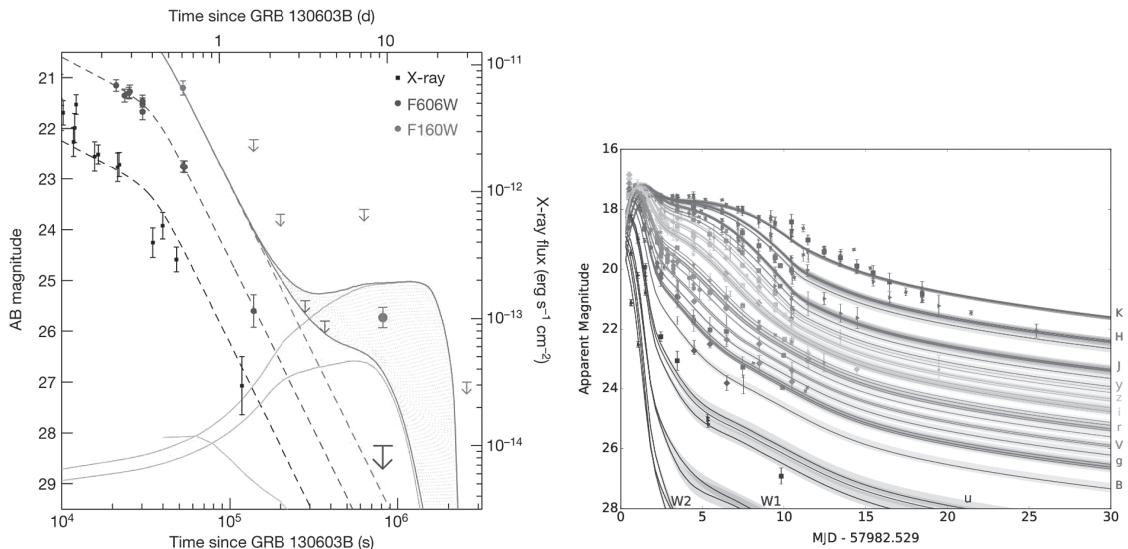
the Japanese community, the term *r-process novae* was sometimes adopted to avoid the “macro-” vs. “kilo-” confusion. Yu et al. (2013) and Metzger and Piro (2014) considered the possibility that a NS–NS merger does not form a black hole promptly, but rather forms a supra-massive, rapidly spinning, highly magnetized neutron star as the merger product. A significant fraction of the spin energy of such a *millisecond magnetar* would be injected to the merger-launched ejecta, making the transients brighter than the r-process-powered one. Depending on the magnetar parameters and ejecta mass, the transients could be brighter than “kilonovae”. The main power of the transients is no longer limited to the r-process, since the spin energy of the magnetar can be comparable to or even larger than the radioactive decay energy. Yu et al. (2013) adopted the name “*merger-nova*” to describe these transients more generally, regardless of whether a magnetar or a black hole forms after the merger. In the literature, the term “kilonova” is the most popular. Here we use these terms interchangeably most of the time, but use “mergernova” when the magnetar-powered optical transients are emphasized.

The original model predicted that the kilonovae (Li and Paczyński, 1998; Metzger et al., 2010) are bright in the optical band. Barnes and Kasen (2013) and Tanaka and Hotokezaka (2013) pointed out that the opacity of the ejecta could be much larger due to the existence of heavier elements, especially the lanthanides. As a result, the photosphere is at a larger radius, where the temperature of the photosphere is lower. The transient is therefore bright in the infrared band rather than the optical band. A bright near-IR emission component was indeed detected from the short GRB 130603B with HST (Tanvir et al. 2013; Berger et al. 2013; Fig. 2.38 left). The lightcurve and spectral behavior of this IR transient seem to be consistent with the prediction of the kilonova model of Barnes and Kasen (2013).⁶

Evidence of a kilonova was also claimed in the SN-less long GRB 060614 (Yang et al., 2015) and the short GRB 050709 (Jin et al., 2016). A systematic search also revealed a mergernova-like bump in GRBs 080503 (Gao et al., 2015c), 050724, 070714B, and 061006 (Gao et al., 2017b). In these latter three cases, a magnetar may be needed to power the mergernova, since the peak luminosity is of the order 10^{42} erg s⁻¹, more than 1 order of magnitude higher than the kilonova luminosity ($\sim 10^{41}$ erg s⁻¹).

The first NS–NS gravitational wave event, GW170817, detected by *Advanced LIGO* and *Advanced Virgo* (Abbott et al., 2017d) was discovered to be associated with a macronova/kilonova event (e.g. Coulter et al., 2017; Pian et al., 2017; Evans et al., 2017; Shappee et al., 2017; Smartt et al., 2017; Nicholl et al., 2017; Chornock et al., 2017). The optical transient seems to have two (blue and red) and possibly even three (an additional “purple”) components (Villar et al. 2017, Fig. 2.38 right). The peak luminosity is of the order 10^{42} erg s⁻¹, at least 10 times greater than the predicted “kilonova” luminosity. There is indirect evidence of lanthanides in the ejecta (e.g. Pian et al., 2017; Tanvir et al., 2017; Smartt et al., 2017). The spectrum and lightcurve of the event are generally consistent with the macronova/kilonova model.

⁶ The broad-band data of GRB 130603B may also be interpreted by having a magnetar as the post-merger product (Fan et al., 2013a).

**Figure 2.38**

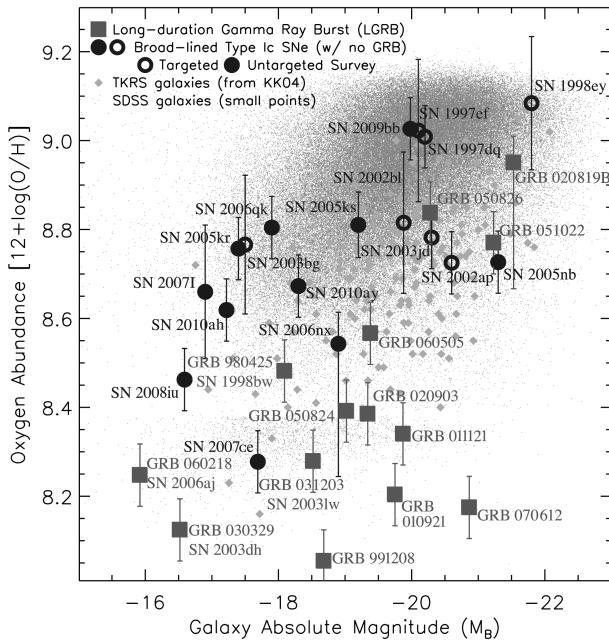
Left: Multi-wavelength lightcurves of GRB 130603B showing an IR excess that is consistent with a “kilonova”. From Tanvir et al. (2013). *Right:* Broad-band lightcurves of GW170817 showing a clear signature of a macronova/kilonova. Reproduced from Figure 1 in Villar et al. (2017) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

2.4 Host Galaxies

2.4.1 Long GRBs

Host Galaxy Properties

The majority of *long GRB host galaxies* are *irregular, star-forming galaxies*, with a few being spiral galaxies with active star formation (Fruchter et al., 2006). One important question is how GRB hosts compare with field galaxies at comparable redshifts, in particular, whether long GRBs prefer a *low-metallicity* environment, as favored by the collapsar progenitor model (Woosley and Bloom, 2006). Studies have shown that long GRB hosts are relatively metal poor (e.g. Fynbo et al., 2003; Prochaska et al., 2004; Fruchter et al., 2006) compared with field galaxies. They are also systematically more metal poor than broad-line Type Ic SNe without GRB associations (Modjaz et al., 2008). Counter-arguments suggest that this apparent metal-poor property of long GRB hosts may not be intrinsic, but is rather a consequence of anti-correlation between star formation and metallicity seen in galaxies in general (Savaglio et al., 2009). Graham and Fruchter (2013) compared the metallicity distributions among the host galaxies of long GRBs, broad-line Type Ic SNe, and Type II SNe, and also against the metallicity distribution of local star-forming galaxies in the SDSS sample. They concluded that such an anti-correlation between star formation rate and metallicity is not adequate in interpreting the data, and long GRBs indeed favor a low-metallicity environment (Fig. 2.39). Such a conclusion is consistent with the expectation of

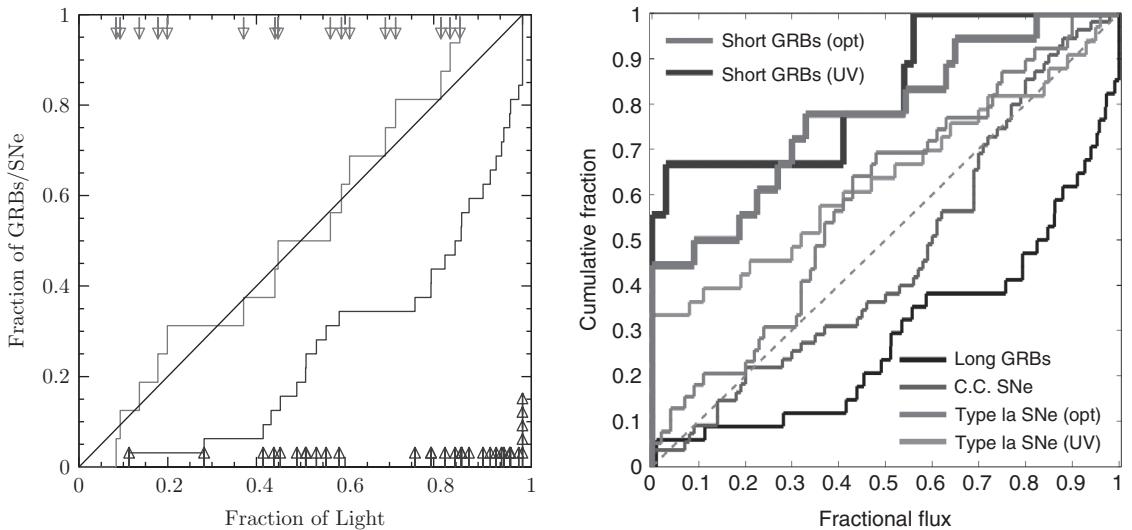
**Figure 2.39**

A comparison of metallicity of long GRB host galaxies with the host galaxies of other broad-line Type Ic SNe not associated with GRBs, Type II SNe, as well as the Sloan Digital Sky Survey galaxy sample. Long GRB hosts on average tend to be more metal poor than other samples. Reproduced from Figure 3 in Graham and Fruchter (2013) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

the *collapsar* model of GRBs (MacFadyen and Woosley, 1999), as well as numerical simulations of the GRB host galaxy luminosity function (Niino et al., 2011). Nonetheless, some GRBs, especially some dark GRBs, are found in relatively metal-rich host galaxies (e.g. Holland et al., 2010; Perley et al., 2013), suggesting that long GRBs can also be produced in relatively metal-rich galaxies. However, at the redshifts of these GRBs, the spatial resolution (\sim kpc) is too poor to pin down the metallicity in the immediate environment of GRBs. Considering small-scale metallicity variations, the data do not rule out the possibility that long GRBs are born exclusively in a low-metallicity environment (Niino et al., 2015).

GRB Location within the Host Galaxy

A systematic study by Fruchter et al. (2006) showed that long GRBs track the brightest light in the host galaxies, suggesting a very high *specific star formation rate* (star formation rate per unit mass) at the burst site. This can be seen in a diagram comparing the cumulative fraction of optical light in the host galaxies vs. the cumulative fraction of light in pixels fainter than or equal to the one at the location of the transient (SN: upper histogram; long GRB: lower histogram). Fruchter et al. (2006) showed that whereas the two fractions

**Figure 2.40**

Left: A comparison between the locations of long GRBs and core-collapse SNe in their host galaxies. From Fruchter et al. (2006). *Right:* A more extended study also including short GRBs and Type Ia SNe. Reproduced from Figure 7 in Fong and Berger (2013) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

roughly track each other for core-collapse SNe (mostly Type II), those of long GRBs significantly deviate from the tracking line, with a much larger fraction of GRBs existing in a relatively small fraction of galaxy light on average (Fig. 2.40 left). In other words, most long GRBs reside in the brightest core regions of the host galaxies, where specific star formation rate is the highest. This is quite consistent with the massive star origin of long GRBs. The conclusion was confirmed by Blanchard et al. (2016) with a much larger long GRB sample.

Some long GRBs have been discovered in regions with a relatively low local specific star formation rate (e.g. Levesque et al., 2012). Long GRB 071025 was found to be located in a halo environment, far away from the bulk of massive star formation (Cenko et al., 2008). The bright afterglow of the GRB, on the other hand, suggests a relatively high density (compared with what is generally expected in a halo environment) of the circumburst medium. Cenko et al. (2008) argued that the GRB may still be associated with a collapsed massive star in a compact star-forming cluster during the merger of two galaxies.

2.4.2 Short GRBs

Host Galaxy Properties

One important difference between the host galaxies of short and long GRBs is that some short GRBs are located in *elliptical* or *early-type galaxies* (Gehrels et al., 2005; Bloom et al., 2006; Barthelmy et al., 2005a; Berger et al., 2005b). A mix of elliptical and spiral host galaxies for short GRBs suggests that they are likely not directly associated with deaths of massive stars, but are more consistent with the *compact-star-merger* models. Fong et al. (2010, 2013) and Fong and Berger (2013) systematically analyzed the host

galaxy properties of a sample of short GRBs and compared them with the hosts of long GRBs and Type II SNe. They found that about 20% (Fong et al., 2013) of short GRBs have an early-type host galaxy. Most short GRB host galaxies are late type, with a moderate star formation rate. As a whole and compared with long GRB hosts, the host galaxies of short GRBs are somewhat larger, the stellar population is relatively older (Leibler and Berger, 2010), and metallicity is relatively higher (Berger, 2014).

Offset

Another important aspect of short GRB host galaxy phenomenology is the relative location of the short GRB with respect to its host. Performing the same exercise as for long GRBs (Fruchter et al., 2006), one can study the cumulative fraction of light in the pixels fainter than or equal to the one at the location of the short GRB against the cumulative fraction of light in the host galaxy (Fig. 2.40 right; Fong and Berger 2013). It is found that the curve of short GRBs is very different from those of long GRBs, core-collapse GRBs, and even Type Ia supernovae. Most short GRBs are found to be far from the bright light of the host galaxies. Another way to look at this is to plot the projected *offset* (both physical and normalized) of the location of the GRB with respect to the center of the host. It is found that short GRBs on average have much larger offsets than long GRBs (Fig. 2.41, Fong and Berger 2013). All these are consistent with the compact-star-merger model for short GRBs, since when two compact stars are born in a binary system, they have undergone two supernovae, each giving a “kick” to the binary system, so that the system drifts away from the star-forming regions. When a merger happens after the binary loses orbital angular momentum through gravitational wave radiation, the system is already far from the star-forming regions (Bloom et al., 2002).

There is a population of short GRBs that are *hostless*. They may be “kicked” away from their host, or reside in distant faint host galaxies (Berger, 2010). A statistical study

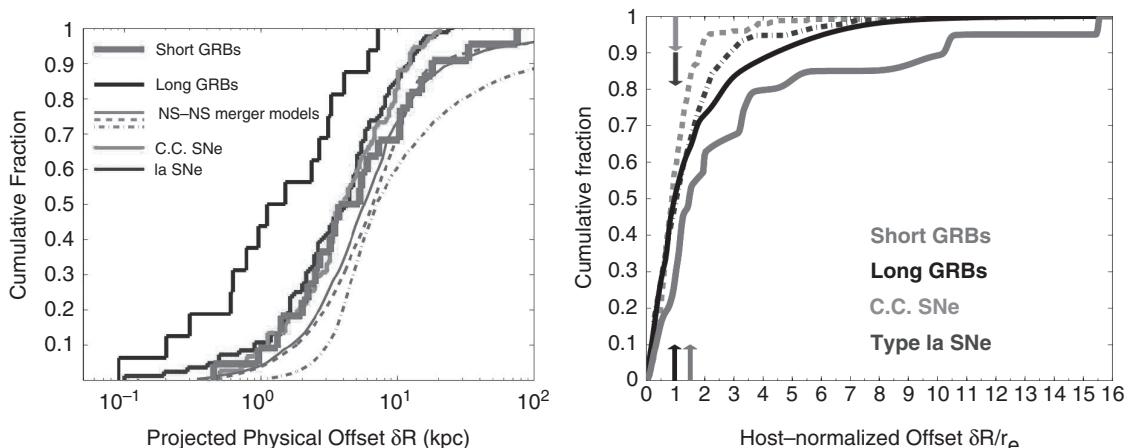


Figure 2.41 Projected physical and normalized offsets of short GRBs with respect to the center of their host galaxies, as compared with the offsets of other transients. Reproduced from Figures 5 and 6 in Fong and Berger (2013) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

suggests that some hostless GRBs are consistent with being kicked out from a nearby galaxy, which is again consistent with the expectation of the compact-star-merger models (Fong and Berger, 2013).

Directly comparing the host galaxy properties of long GRBs and short GRBs, Li et al. (2016b) found that the two populations are not as clearly separated as originally thought. Rather, they have significant overlaps in most properties (Fig. 2.42). In other words, based on host galaxy information only (star formation rate, specific star formation rate, afterglow offset, etc.), one cannot always confidently determine whether a galaxy is the host of a long or short GRB.

The properties of the host galaxy of the first NS–NS gravitational wave event GW170817/GRB 170817A, i.e. NGC4993, fall into the distributions of short GRBs in terms of its size, luminosity, and offset (Fong et al., 2017; Zhang et al., 2018a). Nonetheless, relative to typical short GRB hosts, NGC4993 is superlative in terms of its large optical luminosity, old stellar population age, and low star formation rate (Fong et al., 2017).

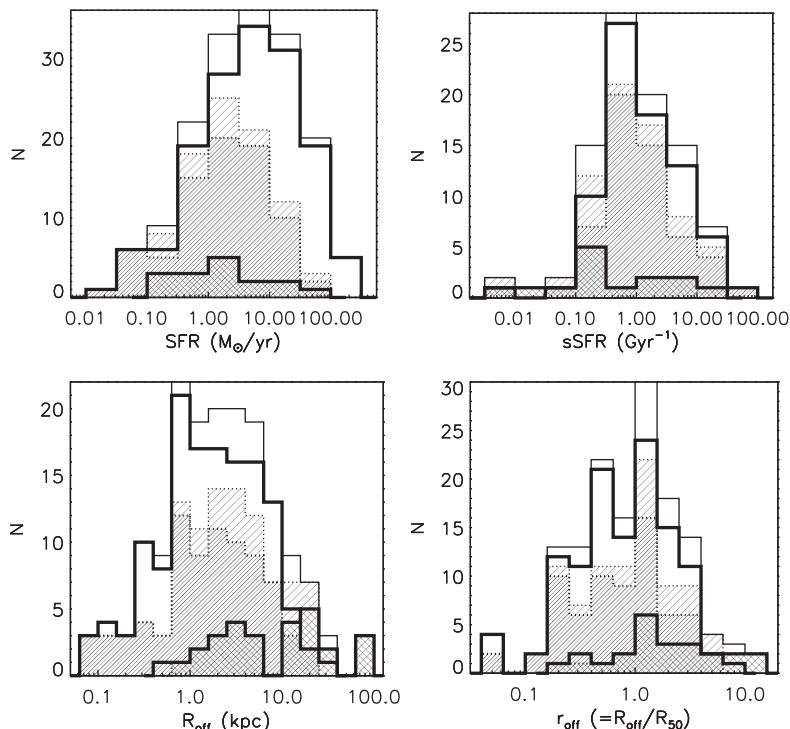


Figure 2.42

A comparison of the distribution of long GRBs and short GRBs (both marked as thick solid histograms with the long population having more events). The thin solid histogram is the total distribution. The thick dotted histogram is for $z < 1.4$ long GRBs only (which covers the same redshift range as short GRBs), and the thin dotted histogram is the total distribution ($z < 1.4$). From upper left to lower right, the distributions are for star formation rate, specific star formation rate, physical offset, and normalized offset, respectively. From Li et al. (2016b).

2.5 Global Properties

2.5.1 Directional Distribution

Both long and short GRBs have an *isotropic* directional distribution. This robust observational fact was established in the BATSE era (Meegan et al. 1992, Fig. 2.43) and is confirmed by mounting data collected by later missions such as *Swift* and *Fermi*. Such a distribution is fully consistent with the cosmological origin of GRBs.

2.5.2 Peak Flux/Fluence Distribution

Before GRB redshifts were measured in the afterglow era in 1997, we did not have distance information for GRBs. No luminosity and energetics information could be retrieved. Two important statistical properties are the distributions of GRB peak flux ($\log N - \log P$) and fluence ($\log N - \log S$). The peak flux/fluence distributions are a convolution of the intrinsic peak luminosity/energy distributions and the distance (or redshift) distribution of GRBs. In the era when no redshift (distance) information was available, these distributions already carried clues about the spatial distribution of GRBs, most importantly whether GRBs were nearby (Galactic) or at cosmological distances.

There are two important *Euclidean* criteria that have been used to test whether GRBs are homogeneously (uniformly) distributed in space.

The first criterion states: *For a certain type of astrophysical object uniformly distributed in a Euclidean space, the number of objects observed above a fluence (or peak flux) S (or P) satisfies*

$$N(> S) \propto S^{-3/2} \quad (2.17)$$

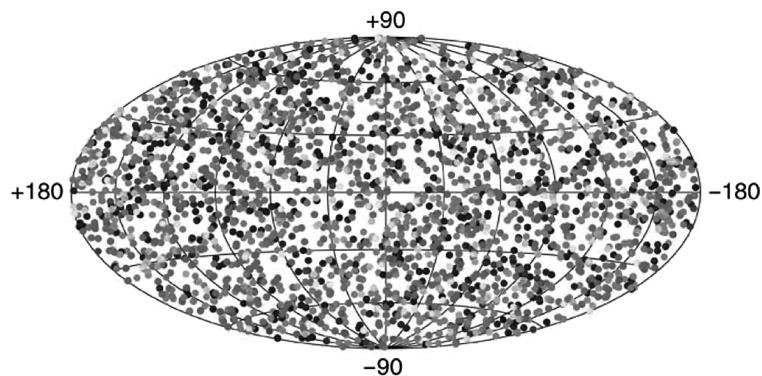


Figure 2.43 The directional distribution of 2704 BATSE GRBs that show an isotropic distribution in the sky (Paciesas et al., 1999). Reproduced from the BATSE GRB 4B Catalogs.

and

$$N(> P) \propto P^{-3/2}, \quad (2.18)$$

regardless of the energy (luminosity) function of the objects.

This can be proven as follows. We first consider a constant total energy E emitted by all the objects. The fluence of an object at a distance r is therefore $S = E/(4\pi r^2) \propto r^{-2}$, or $r \propto S^{-1/2}$. For a uniform distribution, the number density of the objects, n_0 , is constant. The total number of objects above S is $N(> S) = (4\pi/3)r^3 n_0 \propto r^3 \propto S^{-3/2}$.

Next, we consider objects with a distribution of emitted energy dN/dE . Since $N(> S) \propto S^{-3/2}$ is satisfied regardless of energy, the superposed $N(> S)$ should keep the same dependence ($\propto S^{-3/2}$) regardless of the shape of the energy distribution function dN/dE .

Replacing the total emitted energy E by the peak luminosity L , one can prove $N(> P) \propto P^{-3/2}$ regardless of the form of luminosity function $N(L)$.

The second Euclidean criterion states: *For a certain type of astrophysical object uniformly distributed in a Euclidean space, regardless of the energy (luminosity) function, one has the average V-to-V_{max} ratio*

$$\left\langle \frac{V}{V_{\max}} \right\rangle = \frac{1}{2}, \quad (2.19)$$

where $V = (4\pi/3)r^3$ is the volume enclosed in the sphere defined by the source distance r , and $V_{\max} = (4\pi/3)r_{\max}^3$ is the maximum volume inside which the source can be detected, where r_{\max} is the distance at which the fluence (flux) reaches the sensitivity threshold S_{\lim} (F_{\lim}) of the detector.

This can be proven as follows. For a uniform distribution, the number of objects in the distance range r to $r + dr$ is $N(r) = n_0 4\pi r^2 dr$. For a certain fluence S (or peak flux P), the average volume within the radius r_{\max} corresponding to S_{\lim} (or F_{\lim}) is therefore

$$\langle V \rangle = \frac{\int_0^{r_{\max}} (4\pi r^3/3) n_0 (4\pi r^2 dr)}{\int_0^{r_{\max}} n_0 (4\pi r^2 dr)} = \frac{4\pi}{3} \frac{r_{\max}^6/6}{r_{\max}^3/3} = \frac{4\pi}{3} \frac{r_{\max}^3}{2} = \frac{V_{\max}}{2}. \quad (2.20)$$

This gives Eq. (2.19).

In practice, one does not need to measure r and r_{\max} in order to perform the $\langle V/V_{\max} \rangle$ test. For a given source with luminosity L , one has $F = L/4\pi r^2$ and $F_{\lim} = L/4\pi r_{\max}^2$. One therefore has $V/V_{\max} = (r/r_{\max})^3 = (F_{\lim}/F)^{3/2}$. For a detector with known threshold flux F_{\lim} , the V/V_{\max} value for any event with flux F can be readily derived (Schmidt, 1968).

In the BATSE era, great effort was made to study $\log N - \log S$ ($\log N - \log P$) and $\langle V/V_{\max} \rangle$ in order to investigate whether GRBs are uniformly distributed in the nearby (Euclidean) space. If Eqs. (2.17)–(2.19) were observed, it would have given a strong indication that GRBs were local events even without distance information. These investigations however always gave negative results, suggesting an *inhomogeneous* or *non-uniform* distribution. The BATSE $\log N - \log P$ curves (Fig. 2.44) showed a slope shallower than $-3/2$ at low P values, suggesting a deficit of low P GRBs with respect to the Euclidean prediction (e.g. Pendleton et al., 1996). The $\langle V/V_{\max} \rangle$ value was measured to be between 0.3 and 0.4, which is smaller than the expected value 0.5 (e.g. Fishman et al., 1994).

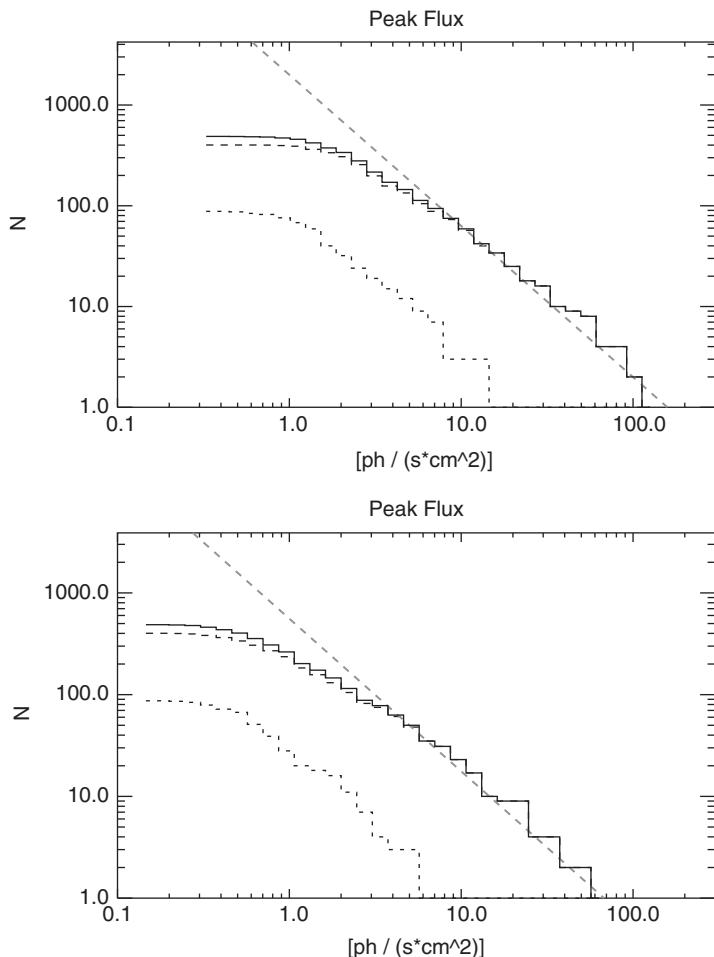


Figure 2.44 The log N –log P distributions of long (dashed), short (dotted), and total (solid) GBM GRBs in the energy ranges of 10–1000 keV (*upper*) and 50–300 keV (*lower*) from the first *Fermi* GBM GRB catalog. The peak flux is defined on the 1.024 s time scale. Reproduced from Figure 4 in Paciesas et al. (2012) with permission. ©AAS.

We now know that these deviations are due to a combination of the following three effects. First, GRBs are cosmological events. The volume enclosed by a certain luminosity distance is no longer a simple $V \propto r^3$ function at relatively large redshifts. The flux is proportional to D_L^{-2} , where D_L is the *luminosity distance*, which becomes very different from the Euclidean distance r as z approaches unity. Second, the GRB event rate density is not a constant. Long GRBs follow the star formation history of the universe, and the star formation rate $\dot{\rho}_*(z)$ is a function of redshift z . Short GRBs may originate from mergers of two compact objects, whose *redshift distribution* invokes more complicated factors but is by no means a uniform distribution. Finally, near the sensitivity threshold of a detector, the number of identified GRBs depends on complicated trigger criteria, which affect the shape of log N –log P . Indeed, by including “untriggered events”, the shape of log N –log P usually changes in the low- P end. Putting all these factors together, the GRB number at low P is reduced, and a smaller $\langle V/V_{\max} \rangle$ value than 0.5 is obtained.

2.5.3 Event Rate and Event Rate Density

Detected Event Rate

The *detected GRB event rate* (number per year) is sensitivity and energy dependent. The current generation of GRB detectors (e.g. BATSE, *Swift*/BAT, *Fermi*/GBM) have a sensitivity limit of the order of 10^{-8} erg cm $^{-2}$ s $^{-1}$. With such a threshold, an ideal imaginary 4π all-sky detector on average would detect \sim 600–1000 GRBs/yr, or \sim 2–3 GRBs/day. For example, *Swift*/BAT has a field of view of \sim 1/7 all sky, and detects 2–3 GRBs per week.

Since short GRBs are typically harder than long GRBs, the long-to-short GRB ratio depends on the bandpass of the detectors. For example, the ratio is \sim 3:1 for BATSE (Paciesas et al., 1999), but is \sim 9:1 for *Swift*/BAT (Sakamoto et al., 2008b, 2011), and \sim 5:1 for *Fermi*/GBM (von Kienlin et al., 2014).

Nearby low-luminosity, long-duration GRBs may form a distinct population. Their detected event rate is about 0.2–0.5 per year by *Swift*/BAT.

Physical Event Rate Density

A physically relevant quantity is the *event rate density* of GRBs (in units of # · Gpc $^{-3}$ · yr $^{-1}$), also called the *volumetric event rate*. It describes how often these events happen in time (rate) and in space (density) in the universe. The *observed event rate density* is redshift dependent (the event rate density varies with cosmic time) and energy/luminosity dependent (more common at lower energies and luminosities). Since GRBs are beamed, the *intrinsic event rate density* also depends on the beaming factor.

In the literature, the event rate density is sometimes denoted as $R_{\text{GRB}}(z)$ (with emphasis on “rate”) or $\rho(z)$ (with emphasis on “density”). To reflect both effects, here we use $\dot{\rho}(z)$ to denote the event rate density at z , which can be written as

$$\dot{\rho}(z) = \dot{\rho}_0 f(z), \quad (2.21)$$

with $\dot{\rho}_0$ denoting the local GRB event rate density. The redshift evolution factor is absorbed in the function $f(z)$, which is discussed in detail below in §2.5.4. Notice that $\dot{\rho}_0$ is still a function of minimum luminosity considered. Strictly, it should be denoted as $\dot{\rho}_0(> L_m)$.

The local event rate density of various species of GRBs can be derived by counting the observed number of GRBs in different redshift bins, supplemented by modeling their *luminosity function* and *redshift distribution* (e.g. Guetta et al., 2005; Liang et al., 2007a; Virgili et al., 2009, 2011a,b; Wanderman and Piran, 2010, 2015; Sun et al., 2015). The currently constrained values (subject to uncertainties) are:

- High-luminosity (HL) long GRBs:

The HL long GRBs are the classical long GRBs with an isotropic luminosity \sim 10 49 –10 54 erg s $^{-1}$. Most observed long-duration GRBs are HL-GRBs. Their observed local event rate density above 10 50 erg s $^{-1}$, which we denote as $\dot{\rho}_{0,50}^{\text{HL}}$ (same convention for other $\dot{\rho}_0$ values also), is

$$\dot{\rho}_{0,50}^{\text{HL}} \sim (0.5 - 1) \text{ Gpc}^{-3} \cdot \text{yr}^{-1}. \quad (2.22)$$

The average local galaxy density is $n_g \sim 0.02 \text{ Mpc}^{-3} \sim 2 \times 10^7 \text{ Gpc}^{-3}$, so one may also write

$$\dot{\rho}_{0,50}^{\text{HL}} \sim (0.025 - 0.05) \text{ gal}^{-1} \cdot \text{Myr}^{-1}. \quad (2.23)$$

Long GRBs roughly trace the star formation history of the universe.⁷ The star formation rate at $z \sim 1$ is roughly 1 order of magnitude higher than at $z = 0$. So one may estimate the event rate density at a typical cosmological distance, i.e.

$$\dot{\rho}_{50}^{\text{HL}}(z \sim 1) \sim 10\dot{\rho}_{0,50}^{\text{HL}} \sim (0.25 - 0.5) \text{ gal}^{-1} \cdot \text{Myr}^{-1}. \quad (2.24)$$

This is roughly once every 2–4 million years per galaxy.

The beaming factor of a GRB is defined by⁸

$$f_b \equiv \frac{\Delta\Omega}{4\pi}, \quad (2.25)$$

where $\Delta\Omega$ is the solid angle of the jet. Considering a bipolar, conical jet with a half-opening angle of θ_j , one has

$$\Delta\Omega = 2 \int_0^{2\pi} \int_0^{\theta_j} \sin\theta d\theta d\phi = 4\pi \int_0^{\theta_j} d(-\cos\theta_j) = 4\pi(1 - \cos\theta_j), \quad (2.26)$$

so that

$$f_b \equiv 1 - \cos\theta_j \simeq \frac{\theta_j^2}{2}, \quad (2.27)$$

where the second approximation applies when $\theta_j \ll 1$.

For HL long GRBs, data suggest the jet correction factor is $(f_b^{\text{HL}})^{-1} \sim 500$, which corresponds to a mean *jet opening angle* $\theta_j^{\text{HL}} \sim 3.6^\circ$ (Frail et al., 2001). One can derive the total intrinsic event rate density of long GRBs at the local universe ($z \sim 0$):

$$\dot{\rho}_{0,\text{tot},50}^{\text{HL}} \sim 500\dot{\rho}_{0,50}^{\text{HL}} \sim (250 - 500) \text{ Gpc}^{-3} \cdot \text{yr}^{-1} \sim (12.5 - 25) \text{ gal}^{-1} \cdot \text{Myr}^{-1}, \quad (2.28)$$

and that at $z \sim 1$:

$$\dot{\rho}_{\text{tot},50}^{\text{HL}}(z \sim 1) \sim (2500 - 5000) \text{ Gpc}^{-3} \cdot \text{yr}^{-1} \sim (125 - 250) \text{ gal}^{-1} \cdot \text{Myr}^{-1}. \quad (2.29)$$

- Low-luminosity (LL) long GRBs:

A small fraction of the observed long GRBs have low isotropic luminosity ($\sim 5 \times 10^{46} - 10^{49} \text{ erg s}^{-1}$). They are only observable at relatively low redshifts. Their observed event rate density above $L_{\min} = 5 \times 10^{46} \text{ erg s}^{-1}$ ($\log L_{\min} = 46.7$) is much higher than that of HL-GRBs:

$$\dot{\rho}_{0,46.7}^{\text{LL}} \sim (100 - 200) \text{ Gpc}^{-3} \cdot \text{yr}^{-1} \sim (5 - 10) \text{ gal}^{-1} \cdot \text{Myr}^{-1} \gg \dot{\rho}_{0,50}^{\text{HL}}. \quad (2.30)$$

The relatively low detection rate of LL-GRBs can be attributed to their low luminosity, since most of them are below the detection sensitivity limit if the redshift is slightly

⁷ This is a very good approximation at low redshifts. At high redshifts (e.g. $z > 5$), it is found that the GRB rate is in excess of what is predicted from the star formation rate. Some additional factors (e.g. low-metallicity preference, evolution of luminosity function) may play a role in defining the GRB redshift distribution.

⁸ In the literature, sometimes f_b is defined as $4\pi/\Delta\Omega$. Throughout the book, we follow the original convention of Frail et al. (2001).

higher. Observations do not show strong evidence of collimation for these events, suggesting a much wider jet opening angle, or that the emission is essentially isotropic. With a beaming factor of $(f_b^{\text{LL}})^{-1} \geq 1$, the total local event rate is

$$\dot{\rho}_{0,\text{tot},46.7}^{\text{LL}} \sim (100 - 200)(f_b^{\text{LL}})^{-1} \text{ Gpc}^{-3} \cdot \text{yr}^{-1} \sim (5 - 10)(f_b^{\text{LL}})^{-1} \text{ gal}^{-1} \cdot \text{Myr}^{-1} \quad (2.31)$$

One can see that the total intrinsic event rate densities of both HL-GRBs (Eq. (2.28)) and LL-GRBs (Eq. (2.31)) are comparable.

- Short GRBs:

Most short GRBs are believed to be of a compact-star-merger origin. Their observed local event rate density has a large uncertainty. At above 10^{50} erg s $^{-1}$, it is estimated to be about (Wanderman and Piran, 2015; Sun et al., 2015)

$$\dot{\rho}_{0,50}^{\text{SGRB}} \sim (0.5 - 3) \text{ Gpc}^{-3} \cdot \text{yr}^{-1} \sim (0.025 - 0.15) \text{ gal}^{-1} \cdot \text{Myr}^{-1} \sim \dot{\rho}_{0,50}^{\text{HL}}. \quad (2.32)$$

It is known that at least some short GRBs are collimated (Burrows et al., 2006; Soderberg et al., 2006; De Pasquale et al., 2010). The beaming factor is however not well constrained and is spread in a wide range from case to case, with a mean value $f_b^{\text{SGRB}} \sim 0.04$ (Fong et al., 2015), or $(f_b^{\text{SGRB}})^{-1} \sim 25$. The total local event rate may be then estimated as

$$\dot{\rho}_{0,\text{tot},50}^{\text{SGRB}} \sim (13 - 75) \text{ Gpc}^{-3} \cdot \text{yr}^{-1} \sim (0.6 - 4) \text{ gal}^{-1} \cdot \text{Myr}^{-1}. \quad (2.33)$$

Short GRBs are often detected at even lower luminosities (e.g. 10^{49} erg s $^{-1}$), at which the local event rate density is even higher.

The detection of GRB 170817A associated with GW170817 at a distance ~ 40 Mpc suggests that the short GRBs can have an isotropic luminosity as low as $\sim 10^{47}$ erg s $^{-1}$. The local event rate density above this luminosity derived from this single event is much higher, i.e. at least (Zhang et al., 2018a)

$$\dot{\rho}_{0,47}^{\text{SGRB}} \sim (190_{-160}^{+440}) \text{ Gpc}^{-3} \cdot \text{yr}^{-1} \sim (9.5_{-8}^{+22}) \text{ gal}^{-1} \cdot \text{Myr}^{-1}. \quad (2.34)$$

This is comparable to or slightly smaller (by a factor of a few) than the estimated NS-NS merger event rate density $\dot{\rho}_0^{\text{NS-NS}} = 1540_{-1220}^{+3200} \text{ Gpc}^{-3} \cdot \text{yr}^{-1}$ inferred from the detection of GW170817 (Abbott et al., 2017d).

2.5.4 Redshift Distribution

The number of GRBs detected per unit (observed) time dt per unit redshift bin dz can be written as (noting $\dot{\rho}^{\text{GRB}}(z) = dN/dt_z dV(z)$ and $dt_z = dt/(1+z)$)

$$\frac{dN}{dt dz} = \frac{\dot{\rho}^{\text{GRB}}(z)}{1+z} \frac{dV(z)}{dz}, \quad (2.35)$$

where

$$\frac{dV(z)}{dz} = \frac{c}{H_0} \frac{4\pi D_L^2}{(1+z)^2 [\Omega_m(1+z)^3 + \Omega_\Lambda]^{1/2}} \quad (2.36)$$

for a flat Λ CDM universe. Here the Hubble constant is

$$H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1} \quad (2.37)$$

with $0.6 < h < 0.8$ and a favored value $h \sim 0.67$,

$$\Omega_m \equiv \frac{8\pi G\rho_0}{3H_0^2} \quad (2.38)$$

and

$$\Omega_\Lambda \equiv \frac{\Lambda c^2}{3H_0^2} \quad (2.39)$$

are the dimensionless matter density and dark energy density parameters of the universe, respectively, ρ_0 is the matter density at the current epoch,

$$D_L(z) = (1+z)D_c(z) \quad (2.40)$$

is the *luminosity distance* of the source at redshift z , and

$$D_c(z) \equiv \frac{c}{H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_m(1+z')^3 + \Omega_\Lambda}} \quad (2.41)$$

is the *comoving distance*⁹ of the source at redshift z .

Given a set of measured *cosmological parameters* ($H_0, \Omega_m, \Omega_\Lambda$), the GRB redshift distribution therefore depends on the functional form of $\dot{\rho}^{\text{GRB}}(z)$, which is different for long and short GRBs.

Long GRBs

Most long-duration GRBs are consistent with having a massive star core-collapse progenitor. To first order, their redshift distribution traces the star formation rate (SFR), $\dot{\rho}_*(z)$, of the universe, which can be mapped with various SFR indicators (e.g. Madau et al., 1998; Hopkins and Beacom, 2006). Whether long GRBs are unbiased tracers of star formation is subject to debate. Most researchers believe that there is an additional weighting factor at play, so that

$$\dot{\rho}^{\text{LGRB}}(z) \propto \dot{\rho}_*(z)\xi(z), \quad (2.42)$$

where $\xi(z)$ is the weighting factor. One important factor is metallicity. Various studies suggest that long GRBs preferentially reside in a low-metallicity environment (e.g. Modjaz et al., 2008; Graham and Fruchter, 2013). If so, $\xi(z)$ would reflect a z -dependent metallicity weighting factor, which becomes more significant at higher redshifts. Alternatively, the GRB luminosity function may evolve with redshift, and $\xi(z)$ may be characterized as a certain functional form, e.g. $(1+z)^\delta$, to mimic such an evolution effect. It has been noticed that the high- z long GRB event rate exceeds the expected rate based on a simple extrapolation of the known star-forming history to higher redshifts, and various effects accounting

⁹ In general, the line-of-sight comoving distance and the transverse comoving distance can be different (Hogg, 1999). The equations presented here are valid for a flat universe (the curvature density term $\Omega_k = 0$), in which the two comoving distances are the same. Such a flat universe is predicted by inflation theory, and is supported by cosmological data. See §13.4.1 for more discussion.

for this high- z excess of GRB rate have been vigorously discussed in the literature (e.g. Kistler et al., 2008; Li, 2008; Salvaterra et al., 2009; Qin et al., 2010; Virgili et al., 2011b; Robertson and Ellis, 2012).

Short GRBs

Most short GRBs are believed to be associated with mergers of two compact objects (two neutron stars or one neutron star and one black hole). Within such a scenario, the redshift distributions of these events have a more complicated functional form. The epoch when a short GRB occurs is jointly defined by the epoch of star formation (when the two massive stars in the binary system were born) and the merger delay time scale (which is the time scale for the two compact objects, NS–NS or NS–BH, to merge due to gravitational wave radiation). In a Λ CDM universe, the *look back time* to a redshift z is defined as

$$t(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{(1+z')[\Omega_m(1+z')^3 + \Omega_\Lambda]^{1/2}}. \quad (2.43)$$

The short GRB event rate density $R_{\text{GRB}}(z)$ may be related to $\dot{\rho}_*$ through the following relation (e.g. Virgili et al., 2011a):

$$\dot{\rho}^{\text{SGRB}}(z) \propto \dot{\rho}_*(z_1), \quad (2.44)$$

where

$$t(z) + \tau = t(z_1), \quad (2.45)$$

z_1 and z are the redshifts for star formation and the short GRB, respectively, and τ is the merger delay time scale for the compact star binary system. The distribution of the delay time scale τ is not well constrained. Analytical models invoke simple function forms, such as a Gaussian or a log-normal distribution with a characteristic delay time scale τ_c , or a power-law distribution, $f(\tau) \propto \tau^\eta$ (e.g. Nakar et al., 2006; Virgili et al., 2011a). More advanced models invoke population synthesis models that closely track the history of star formation, binary evolution, and gravitational wave loss of the binary systems (e.g. Belczynski et al., 2010). One special population synthesis model, dubbed the “twin” model (e.g. Belczynski et al., 2007), has two components in the τ distribution: one “prompt” component and another delayed component. The prompt component is related to those binaries that were born with very tight orbits. The delay time scale τ in this case is not much longer than the massive star lifetimes, so that the mergers happen “promptly” after star formation.

Various authors have applied observed short GRB data to constrain the merger delay time distribution and luminosity function of short GRBs. Using the data of six short GRBs with redshift measurements, Nakar et al. (2006) concluded that long delay time distributions, either a log-normal distribution centered around ~ 4 Gyr or a power-law model with index ~ 0.6 , are consistent with the data. Using BATSE short GRB $\log N - \log P$ data, Guetta and Piran (2006) reached a similar conclusion. However, if one simultaneously considers the $\log N - \log P$ data of BATSE and *Swift* samples, as well as the $L-z$ distribution of the z -known *Swift* sample GRBs, most models cannot satisfy all the data constraints (Virgili et al., 2011a). It is likely that there is a non-negligible contamination of

massive-star-origin GRBs in the observed short GRB samples (Virgili et al., 2011a; Wanderman and Piran, 2015). Another possibility is to have a relatively narrow typical delay time scale peaking around 2–3 Gyr (Virgili et al., 2011a; Wanderman and Piran, 2015), which is not consistent with the expected delay time distribution derived from the Galactic NS–NS binary population. More data are needed to make a tighter constraint on the intrinsic redshift distribution of compact-star-origin short GRBs.

2.5.5 Luminosity Function

Since GRBs are highly variable objects, one needs to specify a time interval to define a luminosity. Usually average luminosity (total isotropic energy divided by cosmological rest-frame duration) and “peak” luminosity are measured and used to study the luminosity function. The peak luminosity is the luminosity at the peak time of a burst. In principle, one needs to specify the same unit time interval in the rest frame for all GRBs to give an accurate measurement. In practice, since some GRBs do not have redshift measurements, one usually derives the peak flux using the detector’s energy band for a particular time bin in the observer frame (e.g. 64 ms or 1 s).

For both average and peak luminosities, a physically meaningful way to define them is to extrapolate what is seen in a relatively narrow energy band to a common cosmological rest-frame energy band, e.g. 1– 10^4 keV. In order to derive such a *bolometric luminosity*, one needs to perform a fit to the observed spectrum, and use the spectral parameters to conduct a *k-correction*. For example, suppose the GRB photon number spectrum can be delineated by a functional form $N(E)$, one then has the *k* factor defined by

$$k \equiv \frac{\int_{1/(1+z)}^{10^4/(1+z)} EN(E)dE}{\int_{e_1}^{e_2} EN(E)dE}, \quad (2.46)$$

where E is in units of keV, and (e_1, e_2) brackets the energy band of the detector. Assuming isotropic emission, the “bolometric” (i.e. “wide-band”) luminosity (denoted as $L_{\text{iso,bol}}$ or L_{iso} , or simply L) can be derived from the observed γ -ray flux F_γ according to

$$L = 4\pi D_L^2 k F_\gamma. \quad (2.47)$$

The GRB *luminosity function* is the distribution of such an isotropic, bolometric luminosity, either the average value (\bar{L}) or the peak value (L_p). The distribution spans a wide range, from several $\times 10^{46}$ to $\sim 10^{55}$ erg s $^{-1}$.

For HL long GRBs, the luminosity function can be characterized as a broken power law with the form (e.g. Liang et al., 2007a)¹⁰

$$\Phi(L)dL = \Phi_0 \left[\left(\frac{L}{L_b} \right)^{\alpha_1^{\text{HL}}} + \left(\frac{L}{L_b^{\text{HL}}} \right)^{\alpha_2^{\text{HL}}} \right]^{-1} dL. \quad (2.48)$$

¹⁰ In some papers (e.g. Wanderman and Piran, 2010), the luminosity function is defined as $\Phi(L)d\log L$ rather than $\Phi(L)dL$. With that definition, the two indices are both smaller by 1.

The parameters α_1^{HL} , α_2^{HL} , and L_b^{HL} are constrained to different values by different authors (e.g. Liang et al., 2007a; Virgili et al., 2009; Wanderman and Piran, 2010; Sun et al., 2015) dependent on the sample size and whether the LL-GRBs are taken into consideration as a separate luminosity function (LF) component (see more below). The most updated values as of 2015 (Sun et al., 2015) are $\alpha_1^{\text{HL}} = 1.0$, $\alpha_2^{\text{HL}} = 2.0$, and $L_b^{\text{HL}} = 7.8 \times 10^{52} \text{ erg s}^{-1}$ for an average isotropic luminosity function.

The luminosity function of LL long GRBs is not well constrained with a small sample size due to the detectors' sensitivity limit. In any case, they cannot be accounted for by simply extrapolating the luminosity function of HL-GRBs to low luminosities. Instead, they likely form a new component with a higher event rate density than the HL-GRB extrapolation (Liang et al., 2007a; Virgili et al., 2009; Sun et al., 2015). Assuming a similar form as Eq. (2.48), the current data cannot give a constraint on α_1^{LL} and L_b^{LL} . The upper part of the LF index, α_2^{LL} , is steeper than the lower part of the LF index, α_1^{LL} , for the HL-GRB component. Fitting LL- and HL-GRBs together the best constraint as of 2015 gives $\alpha_2^{\text{LL}} = 1.7$ (Sun et al. 2015, Fig. 2.45).

The luminosity function of short GRBs is less well constrained than that of long GRBs. The assumption that all short GRBs are due to compact star mergers does not give self-consistent results to account for the $L-z$ and $\log N-\log P$ distributions (Virgili et al., 2011a). However, the exact fraction of the massive-star-GRB contamination in the observed short GRB sample is not well constrained, and multi-wavelength data are needed to judge the physical origin of a short GRB (Zhang et al., 2009a). Removing some short GRBs from the sample, Wanderman and Piran (2015) obtained a broken power-law LF for short

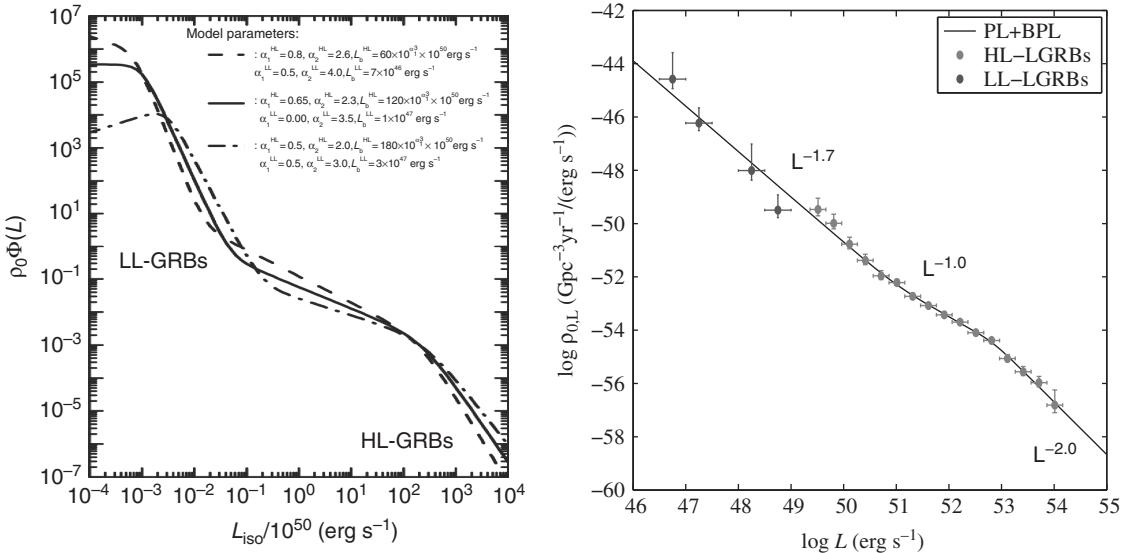


Figure 2.45 Constrained long GRB luminosity function that includes two components: a HL-GRB component with a broken power-law function of the luminosity function, and a distinct LL-GRB component dominating in low luminosities, whose luminosity function form is not well constrained. *Left:* From Liang et al. (2007a); *Right:* From Sun et al. (2015).

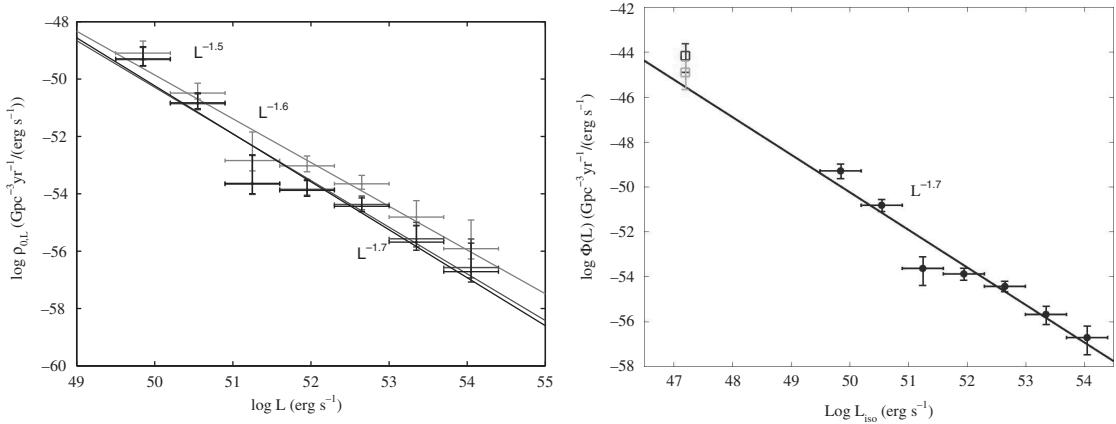


Figure 2.46 *Left:* Constrained short GRB luminosity function assuming that all GRBs are of the compact-star-merger origin. Three merger delay time distribution models are adopted: Gaussian (lower), log-normal (middle), and power law (upper). From Sun et al. (2015). *Right:* The updated short GRB luminosity function with the inclusion of GRB 170817A associated with GW180717. The upper and lower data points at $\sim 10^{47}$ erg s $^{-1}$ are derived from the detections of GW170817 (GW detection) and GRB 170817A (GRB detection), respectively. From Zhang et al. (2018a).

GRBs. Assuming that all short GRBs are due to the compact-star-merger origin, Sun et al. (2015) found that the LF of short GRBs can be fit roughly with a simple power law, with index $\sim 1.5\text{--}1.7$, depending on the assumed merger delay time distribution model (Fig. 2.46 left). It is intriguing that GRB 170817a lies right on the extrapolation of this LF to $\sim 10^{47}$ erg s $^{-1}$ (Zhang et al. 2018a, Fig. 2.46 right).

2.5.6 Isotropic Energy

Integrating the bolometric luminosity over the intrinsic duration of the burst, one can get the *isotropic bolometric emission energy* of a GRB. More conveniently, one can calculate the isotropic bolometric emission energy with the observed γ -ray fluence S_γ via a k -correction, i.e.

$$E_{\gamma,\text{iso}} = \frac{4\pi D_L^2 k S_\gamma}{1+z}. \quad (2.49)$$

Here the subscript γ stands more for “photons” rather than γ -rays, since the range 1–10⁴ keV covers a wide frequency range including both γ -rays and X-rays.

The distribution of $E_{\gamma,\text{iso}}$ is wide, from 10^{49} erg to 10^{55} erg for long GRBs, and from $\sim 3.3 \times 10^{46}$ erg (GRB 170817A at ~ 40 Mpc) to $\sim 10^{53}$ erg (10 keV – 30 GeV for GRB 090510 at $z = 0.903$) for short GRBs.

The *isotropic kinetic energy* of the blastwave after prompt emission, $E_{\text{K,iso}}$, can be derived from the afterglow data through afterglow modeling (e.g. Panaitescu and Kumar, 2002; Yost et al., 2003; Zhang et al., 2007a). See §8.10.2 for the detailed method. In general, the derived $E_{\text{K,iso}}$ scale with $E_{\gamma,\text{iso}}$ among GRBs.

2.5.7 Beaming-Corrected Energy and Luminosity

With the beaming factor f_b (Eqs. (2.25) and (2.27)) inferred from the afterglow “jet break” observations, one can derive the true energetics of GRBs.

The *true (beaming-corrected) bolometric emission energy* is

$$E_\gamma = \frac{\Delta\Omega D_L^2 k S_\gamma}{1+z} = f_b E_{\gamma,\text{iso}}. \quad (2.50)$$

Similarly, the *true (beaming-corrected) afterglow kinetic energy* is

$$E_K = f_b E_{K,\text{iso}}. \quad (2.51)$$

The *radiative efficiency* of a GRB can be defined as (Lloyd-Ronning and Zhang, 2004)

$$\eta_\gamma = \frac{E_\gamma}{E_\gamma + E_K} = \frac{E_{\gamma,\text{iso}}}{E_{\gamma,\text{iso}} + E_{K,\text{iso}}}, \quad (2.52)$$

since the *beaming correction factor* f_b cancels out.¹¹ Observationally, the GRB radiative efficiency is found to vary from less than 1% to over 90% in GRBs (e.g. Zhang et al., 2007a; Wang et al., 2015b).

One interesting observation was that the beaming-corrected energies seem to have a narrower distribution than the isotropic ones. The pre-*Swift* data suggest a very narrow distribution of E_γ , which is clustered around $5 \times 10^{50} - 10^{51}$ erg (Frail et al. 2001; Bloom et al. 2003; upper left panel of Fig. 2.47). The kinetic energy E_K also has a similar narrow distribution (Berger et al. 2003b, upper right panel of Fig. 2.47). This led to the suggestion that long GRBs might have a “standard” energy reservoir (Frail et al., 2001), with an almost constant energy distributed in a range of collimated angles. Alternatively, it was suggested that GRBs may have a quasi-universal “structured” jet with a varying energy per unit solid angle, and different GRBs correspond to different viewing angles from the axes of these quasi-universal jets (Zhang and Mészáros, 2002b; Rossi et al., 2002; Zhang et al., 2004a). With a larger sample of GRBs collected in the *Swift* era, it was found that the distributions of E_γ and E_K are not as narrow as before, even though they are still narrower than the distributions of $E_{\gamma,\text{iso}}$ and $E_{K,\text{iso}}$ (Liang et al. 2008a; Racusin et al. 2009; lower panels of Fig. 2.47).

Similarly, the beaming-corrected average bolometric luminosity, \mathcal{L} , and the peak bolometric luminosity, \mathcal{L}_p , can be corrected from the isotropic values with f_b , i.e.

$$\mathcal{L} = f_b L_{\text{iso}}, \quad (2.53)$$

$$\mathcal{L}_p = f_b L_{p,\text{iso}}. \quad (2.54)$$

Notice that the symbol L has been adopted for isotropic luminosities in the literature (and §2.5.5) in luminosity function studies. For clarity, here we use \mathcal{L} to denote the beaming-corrected luminosities.

¹¹ Here it is assumed that the jet opening angle θ_j remains the same for both prompt emission and afterglow. This is a good assumption for a hydrodynamical conical jet.

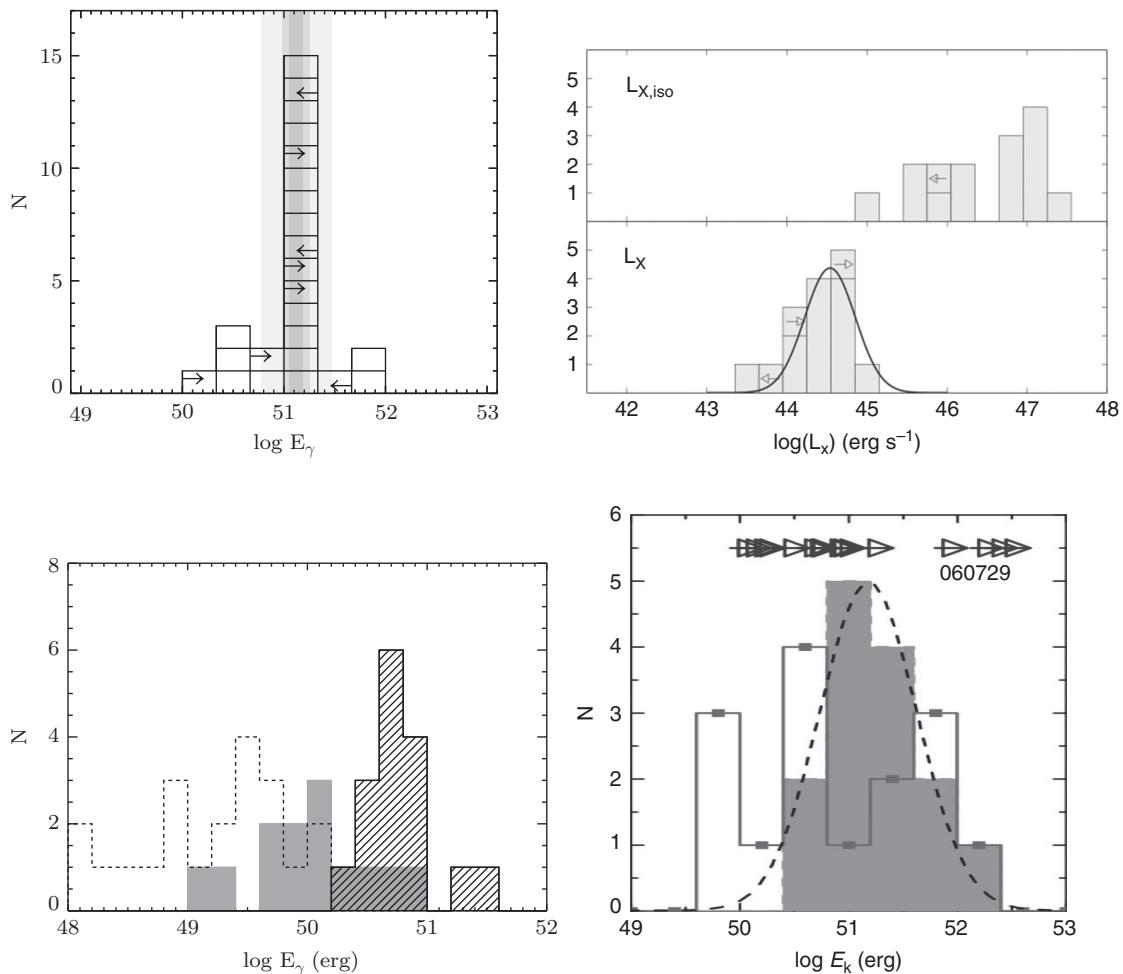


Figure 2.47 *Upper left:* The pre-Swift E_γ histogram. Reproduced from Figure 1 in Bloom et al. (2003) with permission. ©AAS.
Upper right: The pre-Swift 10-hr beaming-corrected X-ray luminosity distribution (lower panel, which is a good proxy for E_K) compared with the isotropic X-ray luminosity distribution (upper panel, which is a good proxy for $E_{K,\text{iso}}$). Reproduced from Figures 1b and 1c in Berger et al. (2003b) with permission. ©AAS. *Lower left:* The post-Swift E_γ distribution histogram (grey solid for the prominent jet break sample and dashed lines for the hidden and possible jet break samples) compared with the pre-Swift sample (filled hatched histogram) by Bloom et al. (2003). From Racusin et al. (2009). *Lower right:* The post-Swift E_K histogram (solid open histogram) compared with the pre-Swift sample (shaded histogram). From Liang et al. (2008a). One can see that the beaming-corrected energies in the post-Swift era have wider distributions than those in the pre-Swift era.

2.6 Empirical Correlations

Some empirical correlations among observational parameters have been reported in the literature.

2.6.1 $E_{p,z}$ – $E_{\gamma,\text{iso}}$ (Amati) and $E_{p,z}$ – $L_{\gamma,p,\text{iso}}$ (Yonetoku) Relations

Amati et al. (2002) discovered a correlation between the GRB isotropic bolometric emission energy, $E_{\gamma,\text{iso}}$, and the rest-frame peak energy, $E_{p,z} = (1+z)E_p$. The correlation has a rough positive dependence, $E_{p,z} \propto E_{\gamma,\text{iso}}^{1/2}$. More precisely, this relation may be written as

$$\frac{E_{p,z}}{100 \text{ keV}} = C \left(\frac{E_{\gamma,\text{iso}}}{10^{52} \text{ erg}} \right)^m, \quad (2.55)$$

with $C \sim 0.8\text{--}1$ and $m \sim 0.4\text{--}0.6$ (Amati, 2006). This relation is found for long GRBs with known redshifts (Amati et al., 2002; Amati, 2006; Frontera et al., 2012), and it covers a wide range of $E_{\gamma,\text{iso}}$ and $E_{p,z}$ values, from bright hard GRBs to low-luminosity X-ray flashes (Sakamoto et al., 2006). The correlation states the fact that more energetic GRBs tend to be harder. Significant outliers have also been observed. For example, GRB 980425 is a low-luminosity nearby GRB with a supernova association (SN 1998bw). Its spectrum is actually much harder than what the Amati relation predicts, with E_p comparable to that of a typical high-luminosity GRB. The upper left panel of Fig. 2.48 shows the relation in 2008 (with GRB 980425 excluded, Amati et al., 2008).

A related correlation is between the isotropic, bolometric peak luminosity $L_{\gamma,p,\text{iso}}$ and $E_{p,z}$ (Wei and Gao, 2003; Yonetoku et al., 2004). Adapted from Yonetoku et al. (2004), this relation reads

$$\frac{E_{p,z}}{100 \text{ keV}} \simeq 1.8 \left(\frac{L_{\gamma,p,\text{iso}}}{10^{52} \text{ erg s}^{-1}} \right)^{0.52}. \quad (2.56)$$

The original plot in Yonetoku et al. (2004) is presented in the upper right panel of Fig. 2.48. This is also a correlation with broad scatter.

Several groups have argued that the Amati relation (and the Yonetoku relation) could be an artifact due to an observational selection effect (e.g. Nakar and Piran, 2005; Band and Preece, 2005; Butler et al., 2007; Kocevski, 2012). Counter-arguments suggested that the selection effect may not completely destroy the correlation (Ghirlanda et al., 2008). Possible supporting evidence for the Amati and Yonetoku correlations is that, within the same burst, an “internal” L – E_p correlation also exists (Liang et al., 2004; Frontera et al., 2012; Lu et al., 2012; Guiriec et al., 2013). This might be, however, dominated by the data points during the decay phase of the bright pulses (Lu et al., 2012; Preece et al., 2014).

Short GRBs do not fall on the long GRB Amati relation. On the other hand, they seem to form a parallel track above the long GRB Amati relation. In other words, given the same $E_{p,z}$, short GRBs are systematically less energetic. This can be attributed to their short durations, which hints that luminosity may be more intrinsically related to $E_{p,z}$. Indeed, in the $E_{p,z}$ – $L_{\gamma,p,\text{iso}}$ space, short and long GRBs are no longer well separated, suggesting that their radiation processes may be similar (Zhang et al., 2009a; Ghirlanda et al., 2009; Guiriec et al., 2013).

2.6.2 $E_{p,z}$ – E_{γ} (Ghirlanda) Relation

Ghirlanda et al. (2004b) claimed that there exists an even tighter correlation between $E_{p,z}$ and the beaming-corrected bolometric emission energy E_{γ} , i.e.

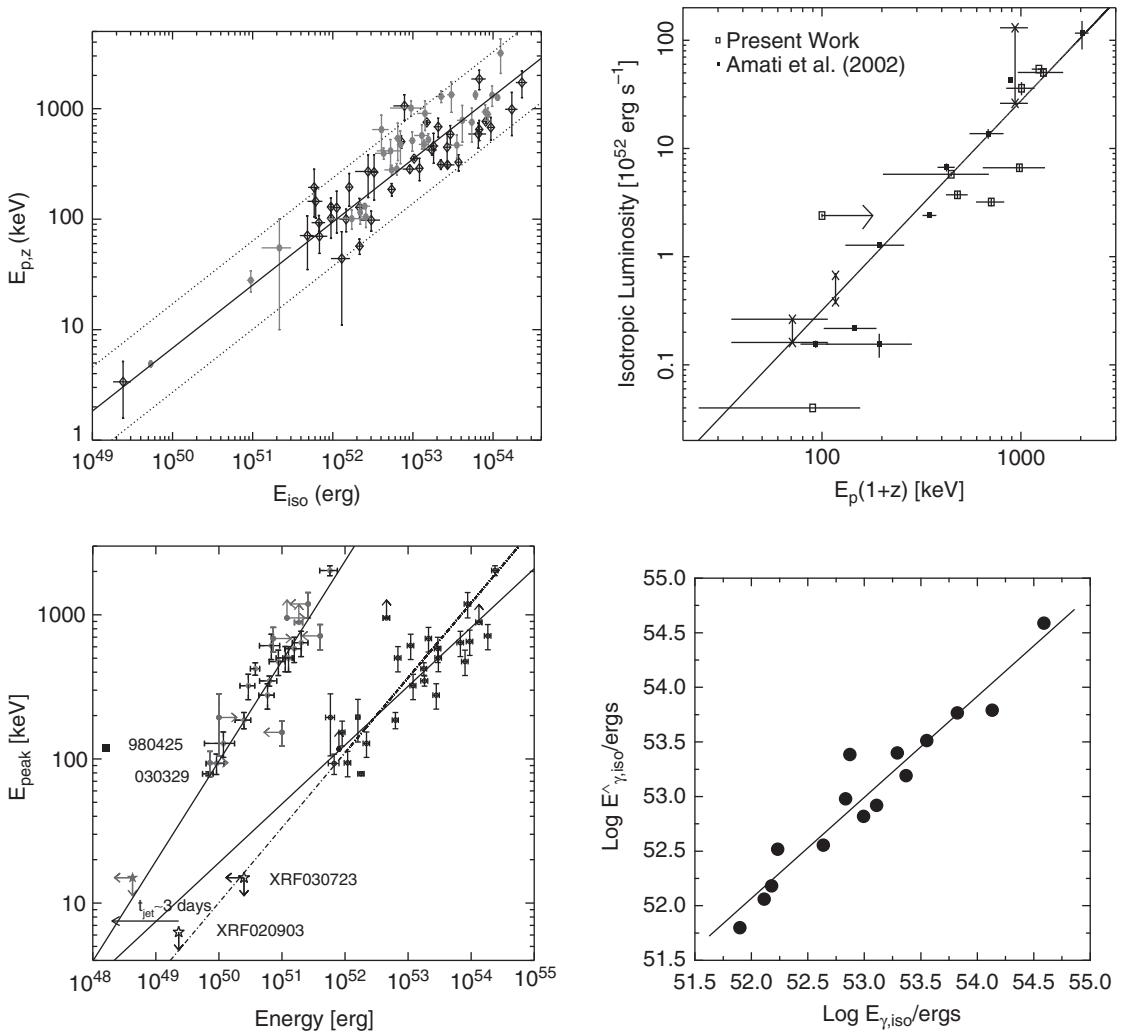


Figure 2.48 *Upper left:* The $E_{p,z}$ - $E_{\gamma,\text{iso}}$ Amati relation. From Amati et al. (2008). *Upper right:* The $E_{p,z}$ - $L_{\gamma,\text{iso}}$ Yonetoku relation. From Yonetoku et al. (2004). *Lower left:* The $E_{p,z}$ - E_{γ} Ghirlanda relation as compared with the Amati relation. From Ghirlanda et al. (2004b). *Lower right:* The $E_{p,z}$ - $E_{\gamma,\text{iso}}-t_{b,z}$ Liang-Zhang relation, where $\hat{E}_{\gamma,\text{iso}}$ is the predicted $E_{\gamma,\text{iso}}$ based on the relation. From Liang and Zhang (2005).

$$\frac{E_{p,z}}{100 \text{ keV}} \simeq 4.8 \left(\frac{E_{\gamma}}{10^{51} \text{ erg}} \right)^{0.7}. \quad (2.57)$$

The lower left panel of Fig. 2.48 shows the Ghirlanda relation as compared with the Amati relation (Ghirlanda et al., 2004b).

In the *Swift* era, multi-wavelength observations allowed properly selecting jet breaks based on the “achromatic” criterion. A re-analysis in 2018 suggested that the Ghirlanda relation becomes less tight, especially when the early jet breaks are included in the analysis (Wang et al., 2018).

2.6.3 $E_{p,z}$ - $E_{\gamma,\text{iso}}$ - $t_{b,z}$ (Liang–Zhang) Relation

Regardless of the interpretation of the afterglow temporal breaks, Liang and Zhang (2005) discovered a fundamental-plane correlation among $E_{p,z}$, $E_{\gamma,\text{iso}}$, and $t_{b,z}$, where $t_{b,z} = t_b/(1+z)$ is the break time in the rest frame of the burst as measured in the *optical* band. In its original form, this relation reads

$$\frac{E_{p,z}}{100 \text{ keV}} \simeq 1.09 \left(\frac{E_{\gamma,\text{iso}}}{10^{52} \text{ erg}} \right)^{0.52} \left(\frac{t_{b,z}}{\text{day}} \right)^{0.64}, \quad (2.58)$$

and it is presented in the lower right panel of Fig. 2.48. Such an empirical correlation is not attached to the jet theoretical framework, and leaves room for theoretical interpretation. The Ghirlanda relation, when expanded to explicitly include the jet break time t_j , has a similar form as the Liang–Zhang relation if one replaces t_j by t_b . This relation also becomes less tight in the *Swift* era when early jet breaks are included (Wang et al., 2018)

2.6.4 $E_{p,z}$ - $L_{\gamma,p,\text{iso}}$ - $T_{0.45}$ (Firmani) Relation

With prompt emission parameters only, Firmani et al. (2006) claimed another three-parameter correlation:

$$\frac{E_{p,z}}{100 \text{ keV}} \simeq 1.37 \left(\frac{L_{\gamma,p,\text{iso}}}{10^{52} \text{ erg s}^{-1}} \right)^{0.62} \left(\frac{T_{0.45,z}}{10 \text{ s}} \right)^{-0.30}. \quad (2.59)$$

Here $T_{0.45,z} = T_{0.45}/(1+z)$, and $T_{0.45}$ is the time span during which the brightest 45% of the total counts are detected above the background. The main difference between $T_{0.45}$ and the traditional T_{90} (or T_{50}) is that the former deducts any quiescent period that may exist during the burst, and therefore better represents the duration of the emission episodes of a burst. The 45% percentage has no physical significance, which was adopted to achieve the most significant correlation.

2.6.5 $E_{\gamma,\text{iso}}$ - θ_j (Frail) Relation

As already mentioned earlier, Frail et al. (2001) found that the measured jet opening angle θ_j of pre-*Swift* GRBs seem to be anti-correlated with $E_{\gamma,\text{iso}}$ through $E_{\gamma,\text{iso}} \propto \theta_j^{-2}$, so that the jet-corrected γ -ray energy $E_{\gamma} \simeq (\theta_j^2/2)E_{\gamma,\text{iso}}$ is roughly constant, $\sim 10^{51}$ erg for long-duration GRBs. The correlation was confirmed by Bloom et al. (2003) and was extended to kinetic energy by Berger et al. (2003b). The implication is that long GRBs have a standard energy reservoir. Wider jets tend to have a lower energy concentration, while narrow jets have a higher energy concentration. Alternatively, this may be understood as a universal (Zhang and Mészáros, 2002b; Rossi et al., 2002) or quasi-universal (Zhang et al., 2004a) jet for GRBs, with the inferred jet opening angle replaced by the observer’s viewing angle.

In the *Swift* era, the Frail relation was found to be no longer tight. Both E_{γ} and E_K are found to have a wider distribution than the pre-*Swift* sample (Liang et al., 2008a; Kocevski

and Butler, 2008; Racusin et al., 2009). The Ghirlanda relation discussed above is in conflict with the Frail relation: instead of having a constant E_γ as the Frail relation suggests, the Ghirlanda relation suggests a correlation between E_γ and $E_{p,z}$.

The pre-*Swift* and post-*Swift* histograms of E_γ and E_K (or the X-ray luminosity, which is a proxy of E_K) (Bloom et al., 2003; Berger et al., 2003b; Liang et al., 2008a; Racusin et al., 2009) are presented in Fig. 2.47.

2.6.6 $L-\tau$ (Norris) Relation

Norris et al. (2000) discovered an anti-correlation between GRB peak luminosity and the delay time (spectral lag) τ for the arrival of low-energy photons (25–50 keV) with respect to high-energy photons (100–300 keV and >300 keV) for a sample of BATSE GRBs. In its original form, it is written as

$$\frac{L_{\gamma,p,\text{iso}}}{10^{53} \text{ erg s}^{-1}} \simeq 1.3 \left(\frac{\tau}{0.01 \text{ s}} \right)^{-1.14}, \quad (2.60)$$

where τ is measured in the observer frame. Several groups later investigated this correlation by considering the spectral lags in the burst rest frame. One way is to correlate $L_{\gamma,p,\text{iso}}$ with $\tau/(1+z) \times (1+z)^{0.33} = \tau/(1+z)^{0.67}$ (Gehrels et al. 2006; Zhang et al. 2009a; left panel of Fig. 2.49). By doing so, one has assumed that the spectral lag is proportional to the pulse width w (which has an energy dependence of ~ 0.33 power). This is valid for individual pulses. For complex bursts with overlapping pulses, Ukwatta et al. (2012) argued that it is more appropriate to investigate a correlation between $L_{\gamma,p,\text{iso}}$ and $\tau_z = \tau/(1+z)$, and gave

$$\log \left(\frac{L_{\gamma,p,\text{iso}}}{\text{erg s}^{-1}} \right) = (54.7 \pm 0.4) - (1.2 \pm 0.2) \log \frac{\tau_z}{\text{ms}} \quad (2.61)$$

for the lag defined between the 100–150 keV and the 200–250 keV energy bands in the rest frame of the GRB sources (right panel of Fig. 2.49).

There are significant outliers in the $L-\tau$ correlation. It seems that even though the low-luminosity GRB 060218 may be moderately accommodated within the correlation (Liang et al., 2006a), several other low-luminosity GRBs (e.g. GRB 980425, GRB 031203) and the supernova-less long GRBs 060614 and 060505 all lie well below the correlation (Gehrels et al., 2006; McBreen et al., 2008). Short GRBs all have negligible lags (Yi et al., 2006), and do not follow the correlation.

2.6.7 $L-V$ (Fenimore–Reichart) Relation

Fenimore and Ramirez-Ruiz (2000) and Reichart et al. (2001) proposed a correlation between the GRB luminosity and the complexity of GRB lightcurves, a parameter defined as “variability” V . The definition of variability depends on how the smoothed background lightcurve is defined, and can be technically very different among authors. In any case, a positive correlation $L_{\gamma,p,\text{iso}} \propto V^m$ with large scatter was found, although the index m ranges from 3.3 (Reichart et al., 2001) to 1.1 (Guidorzi et al., 2005).

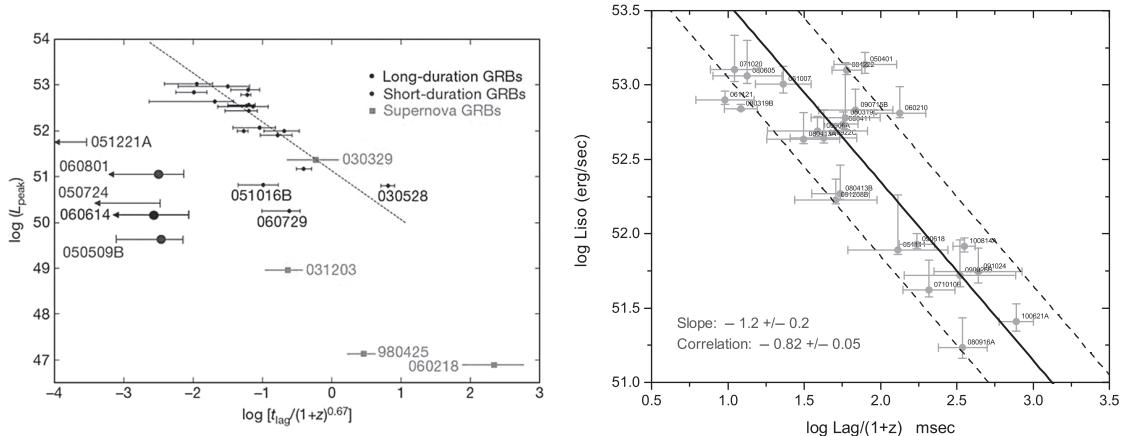


Figure 2.49 The luminosity–spectral lag correlations. *Left:* The $L_{\gamma,p,\text{iso}} - \tau/(1+z)^{0.67}$ correlation and noticeable outliers. From Gehrels et al. (2006). *Right:* The $L_{\gamma,p,\text{iso}} - \tau/(1+z)$ correlation. From Ukwatta et al. (2012).

2.6.8 X-ray Plateau $L_X - T_{a,z}$ (Dainotti) Relation and $L_X - T_{a,z} - E_{\gamma,\text{iso}}$ (Xu–Huang) Relation

Dainotti et al. (2008) discovered that there exists a rough anti-correlation between the rest-frame X-ray plateau ending time ($T_{a,z} = T_a/(1+z)$, where T_a is the plateau ending time as defined by Willingale et al. 2007) and the X-ray luminosity L_X at T_a . The slope is roughly -1 (Dainotti et al., 2013). This suggests that the total plateau energy has a relatively small scatter: a longer plateau tends to have a lower luminosity and vice versa.

Xu and Huang (2012) introduced a third parameter $E_{\gamma,\text{iso}}$ and claimed that a three-parameter correlation is tighter. In its original form, this relation is expressed as

$$L_X \propto T_a^{-0.87} E_{\gamma,\text{iso}}^{0.88}. \quad (2.62)$$

2.6.9 $E_{\gamma,\text{iso}} - \Gamma_0$ (Liang–Ghirlanda), $L_{\gamma,p,\text{iso}} - \Gamma_0$ (Lü), and $L_{\gamma,p,\text{iso}} - E_{p,z} - \Gamma_0$ (Liang) Relations

A sample of GRBs have high-quality early optical afterglow data. A good fraction of them show an early hump in the lightcurve, which is consistent with being due to deceleration of the blastwave. Within the framework of such an interpretation, the initial Lorentz factor of the outflow, Γ_0 , of a moderate sample of GRBs can be measured. Liang et al. (2010) discovered a positive correlation between Γ_0 and the isotropic γ -ray energy $\Gamma_0 \propto E_{\gamma,\text{iso}}^a$, with $a \sim 1/4$. The positive correlation was verified by Ghirlanda et al. (2011) and Lü et al. (2012). Lü et al. (2012) further discovered a similar correlation between Γ_0 and the average isotropic γ -ray luminosity $L_{\gamma,\text{iso}}$, i.e. $\Gamma_0 \propto L_{\gamma,\text{iso}}^b$, with b also close to $1/4$. When considering beaming correction, a correlation between Γ_0 and L_{γ} still exists (although with larger dispersion), with a similar index (Yi et al., 2015).

Liang et al. (2015) investigated a list of three-parameter correlations among $L_{\gamma,p,\text{iso}}$ (or $E_{\gamma,\text{iso}}$), $E_{p,z}$, and Γ_0 , and found that the tightest one is $L_{\gamma,p,\text{iso}} - E_{p,z} - \Gamma_0$. This relation is

tighter than the $L_{\gamma,p,\text{iso}}-E_{p,z}$ and the $L_{\gamma,p,\text{iso}}-\Gamma_0$ relations. In its original form, the two ways of expressing this three-parameter correlation are

$$L_{\gamma,p,\text{iso},52} = 10^{-6.38 \pm 0.35} \left(\frac{E_{p,z}}{\text{keV}} \right)^{1.34 \pm 0.14} \Gamma_0^{1.32 \pm 0.19},$$

$$E_{p,z} = 10^{3.71 \pm 0.38} \text{ keV } L_{\gamma,p,\text{iso},52}^{0.55 \pm 0.06} \Gamma_0^{-0.50 \pm 0.17}. \quad (2.63)$$

Figure 2.50 shows the $E_{\gamma,\text{iso}}-\Gamma_0$, $L_{\gamma,\text{iso}}-\Gamma_0$, and $L_{\gamma,p,\text{iso}}-E_{p,z}-\Gamma_0$ relations (Liang et al., 2010; Lü et al., 2012; Liang et al., 2015).

2.6.10 GRB-Associated SNe as Standard Candles?

Type Ia supernovae have been regarded as standard candles thanks to a clear correlation between the peak luminosity and the decline rate of the SN lightcurve (Phillips, 1993).

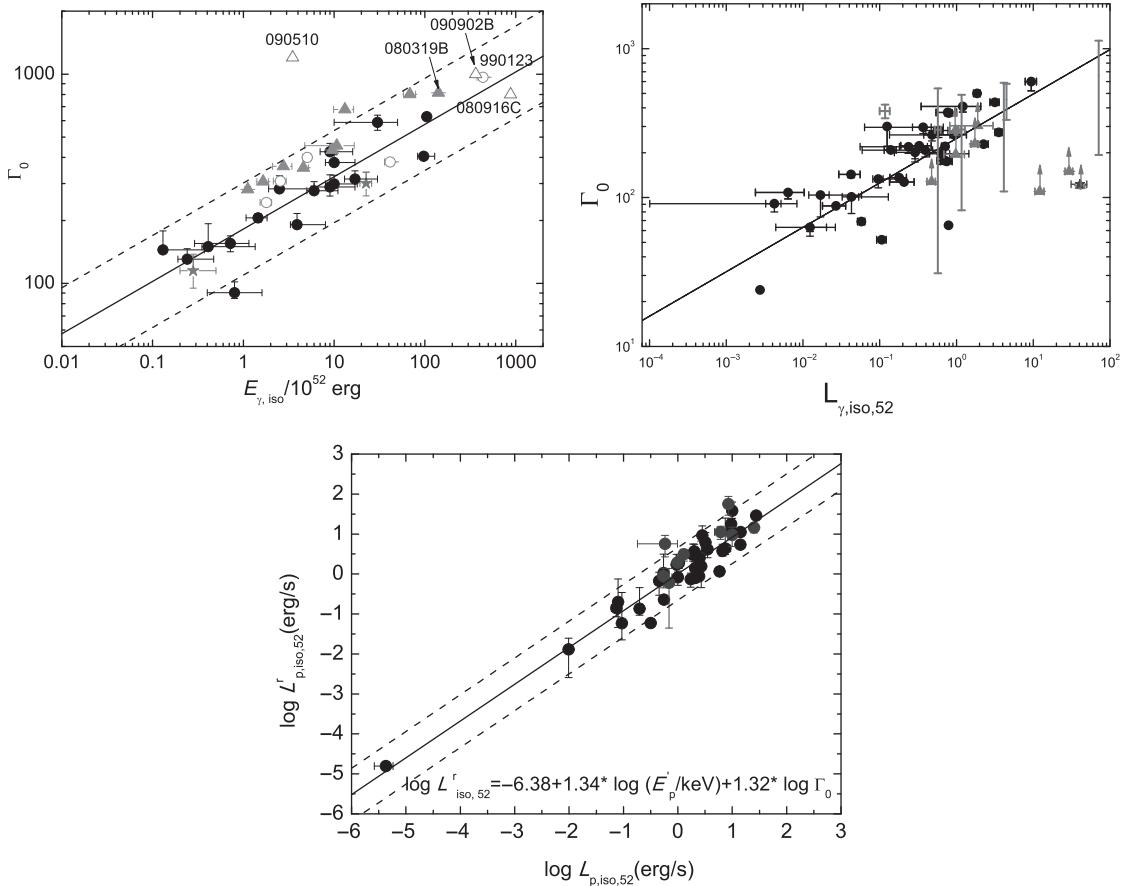


Figure 2.50 *Upper left:* The $E_{\gamma,\text{iso}}-\Gamma_0$ relation. From Liang et al. (2010). *Upper right:* The $L_{\gamma,\text{iso}}-\Gamma_0$ relation. From Lü et al. (2012). *Lower:* The $L_{\gamma,p,\text{iso}}-E_{p,z}-\Gamma_0$ relation. $L'_{\gamma,p,\text{iso}}$ is the derived luminosity based on the three-parameter correlation. From Liang et al. (2015).

Since the discovery of GRB-associated SNe, efforts have been made to look for similar correlations in order to establish GRB-associated SNe as standard candles. Despite the failure of earlier attempts, Li and Hjorth (2014) and Cano (2014) argued that the Type Ic SNe associated with GRBs also seem to have a luminosity–lightcurve decline rate correlation similar to Type Ia. In particular, Li and Hjorth (2014) claimed a correlation,

$$M_{V,\text{peak}} = 1.59^{+0.28}_{-0.24} \Delta m_{V,15} - 20.61^{+0.19}_{-0.22}, \quad (2.64)$$

where $M_{V,\text{peak}}$ is the peak V-band absolute magnitude and $\Delta m_{V,15}$ is the change of apparent magnitude in 15 days (positive value), which denotes the decline rate of the SN lightcurve. Cano (2014), on the other hand, claimed a correlation between the luminosity (k in his notation) and the lightcurve “stretching factor” (s in his notation). Since the sample of GRB-associated SNe is still small, both correlations are subject to confirmation with larger samples in the future.

2.7 Classification

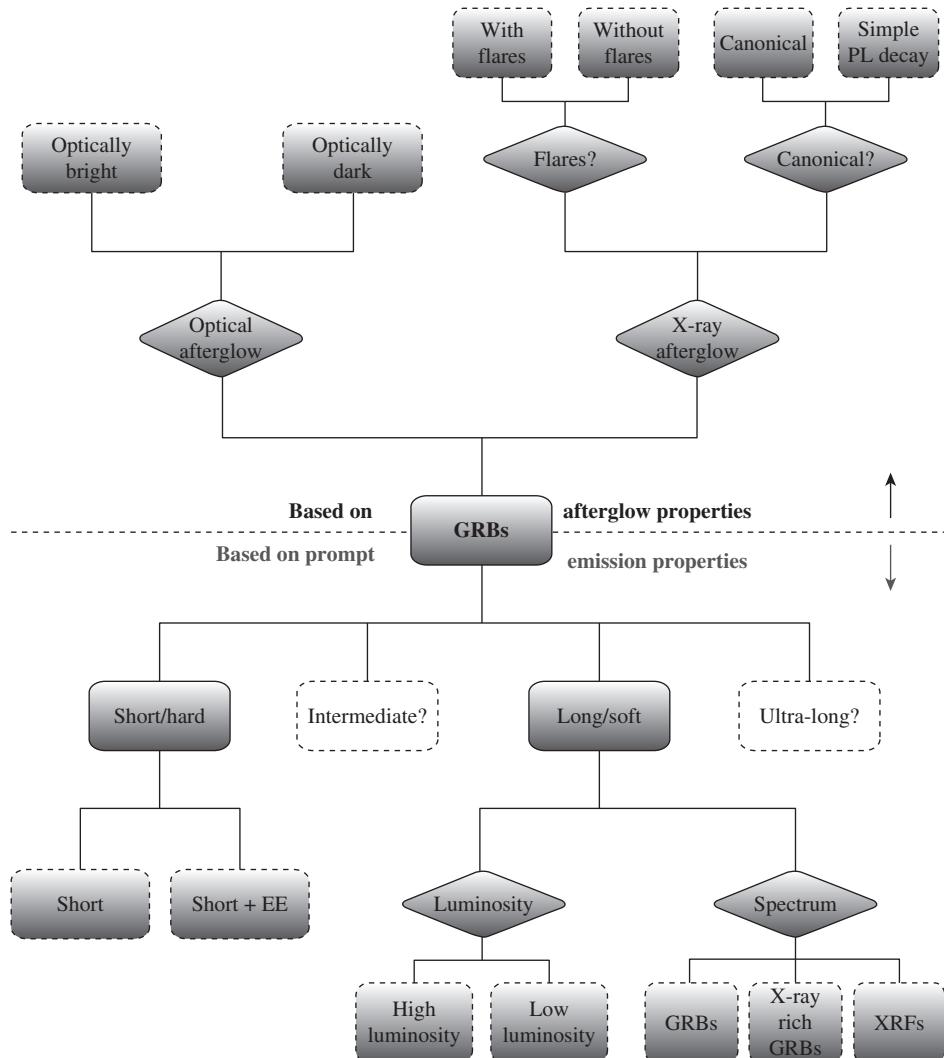
Classification is an important ingredient in understanding the nature of astronomical objects. It usually starts with phenomenological classification schemes according to some apparent divisions in one or more observational quantities based on some well-defined observational criteria. As observational data gradually accumulate, one can gain physical insights into the origins of the observed phenomena, and then try to classify the observed objects “physically”. One example is the supernova classification schemes (Fig. 2.32). Based on some well-defined observational criteria (whether or not H, He, and Si lines are observed in the spectra), one classifies SNe into Types I and II, and further into subtypes Ia, Ib, and Ic. With physical understanding, one could re-classify SNe into two types: massive star core-collapse SNe, which includes Type II, Ib, and Ic; and white dwarf thermonuclear SNe triggered by accretion in binary systems, which are Type Ia.

The GRB classification schemes also follow a similar path. Figures 2.51 and 2.54 summarize various phenomenological and physical classification schemes of GRBs and their possible connections.

2.7.1 Phenomenological Classification Schemes

Duration–Hardness Classification Scheme

The main classification scheme is the *long/soft* vs. *short/hard* dichotomy in the duration domain supplemented by the hardness information (Kouveliotou et al. 1993, see Figs. 2.2 and 2.3). The boundary between the two classes is vague. The duration separation line is around 2 seconds in the BATSE band (30 keV – 2 MeV). Long and short GRBs roughly make up 3/4 and 1/4 of the total population in the BATSE sample, but the short GRB fraction is smaller for other detectors (Sakamoto et al., 2008b, 2011; Paciesas et al., 2012; Zhang et al., 2012d; Qin et al., 2013). This is because the duration T_{90} of a GRB is energy

**Figure 2.51**

GRB phenomenological classification schemes. The bottom portion is the scheme based on prompt emission properties, whereas the upper portion is the scheme based on the afterglow properties. Solid shaded boxes denote the most robust classification schemes supported by observational data; dashed shaded boxes denote secondary, reasonable classification schemes indirectly supported by observations; dashed open boxes denote classification schemes speculated but not fully confirmed by the data; rhombus-shaped boxes denote classification criteria. Figure courtesy Jared Rice.

dependent and detector-sensitivity dependent (Qin et al., 2013). It is possible that a short GRB detected by BATSE would appear as “long” to a detector with a softer bandpass (e.g. *Swift*). Indeed, in the *Swift* era, about 2% of GRBs have a short/hard spike typically shorter than 2 s, but have an extended emission (EE) tail lasting tens to \sim 100 seconds (Norris and Bonnell, 2006). So the unfortunate consequence of the T_{90} classification is that

membership to a certain category of the *same* GRB is detector dependent. Nonetheless, the confusion in T_{90} classification mostly arises in the “grey” zone between the two classes. Practically, one usually defines long and short GRBs based on a rough duration separation line around 2 seconds even for *Swift* and *Fermi* GRBs, even though the separation line may be detector dependent (Qin et al., 2013; Bromberg et al., 2013).

Based on the duration (T_{90}) information, various authors have suggested the existence of a third intermediate class (e.g. Mukherjee et al., 1998; Horváth, 1998; Hakkila et al., 2003; Horváth et al., 2006, 2010; Veres et al., 2010).

Several ultra-long GRBs (with T_{90} longer than thousands of seconds) have been detected (e.g. Thöne et al., 2011; Campana et al., 2011; Gendre et al., 2013; Levan et al., 2014b; Virgili et al., 2013; Stratta et al., 2013). Some authors argued that they form another distinct group (e.g. Gendre et al., 2013; Levan et al., 2014b). Applying a more generally defined duration with both γ -ray and X-ray data taken into account, Zhang et al. (2014) found that these ultra-long GRBs may be the long-duration tail of normal long GRBs. Virgili et al. (2013) reached a similar conclusion based on afterglow observations and modeling. More data are needed in order to address whether a distinct population is needed to account for ultra-long GRBs.

Supplementary Criteria

Other observational information is helpful to refine the duration–hardness classification scheme. For example, the spectral lag τ has been applied as a supplementary parameter. While long GRBs typically have lags, short GRBs have zero or even negative lags (Norris and Bonnell 2006; Yi et al. 2006, see Fig. 2.49 left). As a result, a not-too-short burst (e.g. short GRBs with extended emission) may be regarded “short” if the spectral lag is essentially zero. This criterion is, however, not definitive, since GRBs with very high luminosity and rapid variability also tend to have negligible lags. Theoretically the duration of the spectral lag is related to the duration of an emission unit (e.g. pulse) in the lightcurve (e.g. Zhang et al., 2009a). As a result, the negligible lag in a short GRB is naturally expected because of its short duration, and the negligible lag of a bright long GRB is related to its rapid variability.

Lü et al. (2014) suggested that the *amplitude* of an observed lightcurve may be taken into account as a third dimension in classifying GRBs (Fig. 2.52). First, one can define an *f parameter* as the ratio between the peak flux and the background flux of a GRB. This parameter reflects the “apparent brightness” of a GRB. A high-flux GRB would stick out from the background significantly and have a high *f* value. This parameter alone does not help much, since the *f* distributions for long and short GRBs are similar. Next, for each long GRB, one can simulate a “pseudo GRB” by scaling down the flux globally, until the “signal” above the background has a duration shorter than 2 seconds. One then makes a “pseudo short GRB” from a long GRB. The amplitude parameter of the pseudo GRB is defined as f_{eff} of the original long GRB. Its physical meaning is the amplitude of a “disguised short” GRB due to the “tip-of-iceberg” effect (i.e. a long GRB which is confused as a short GRB because most of its emission is buried beneath the background). Comparing the f_{eff} distribution of long GRBs and the *f* distribution of short GRBs, Lü et al. (2014)

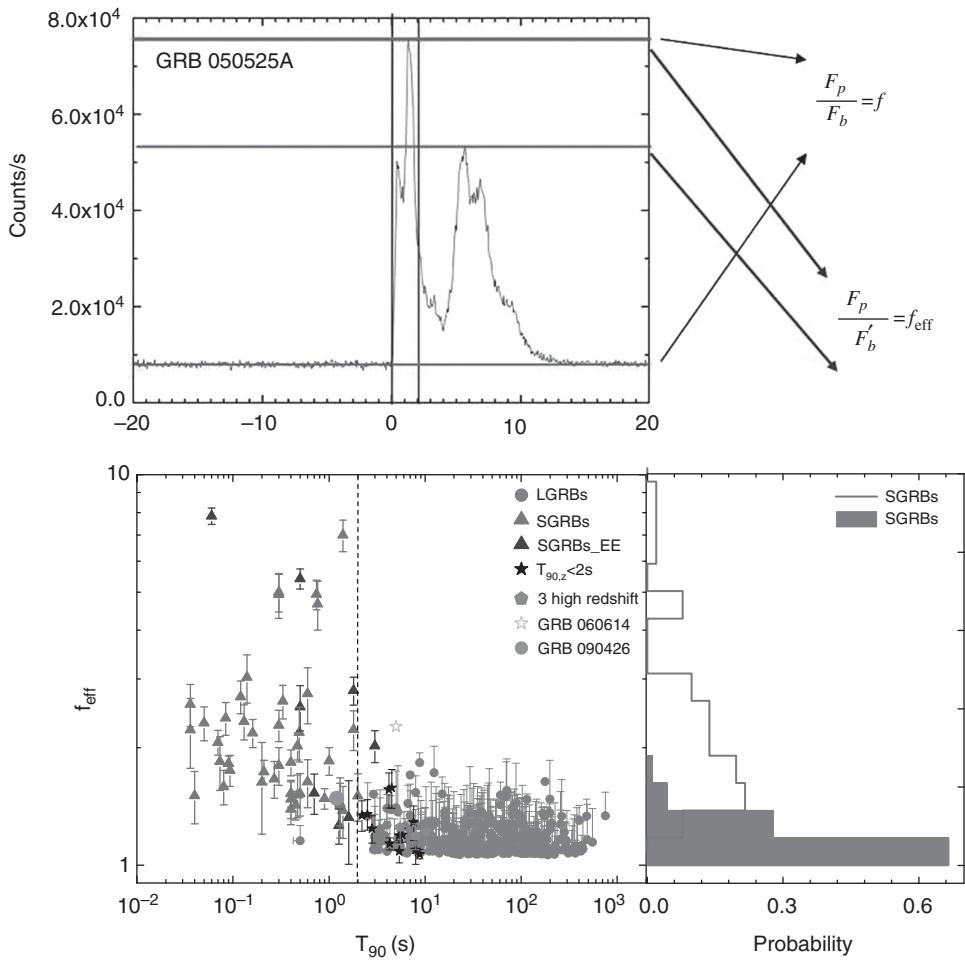


Figure 2.52 Application of the amplitude f and f_{eff} parameters in the phenomenological GRB classification scheme. *Upper:* The definitions of f and f_{eff} parameters. Figure courtesy Hou-Jun Lü. *Lower:* The distribution of f_{eff} for both long and short GRBs. For short GRBs, $f_{\text{eff}} = f$. One can see that most short GRBs are not disguised since their f values are much larger than the f_{eff} values of long GRBs. From Lü et al. (2014).

found that most short GRBs have f values large enough so that they are not disguised (Fig. 2.52 lower panel). Nonetheless, contamination from long GRBs indeed happens when the observed f value of a short GRB is small, and the contamination probability rapidly increases with decreasing f as $P(< f) \sim 0.78^{+0.71}_{-0.4} f^{-4.33 \pm 1.84}$.

A good fraction of GRBs with the highest redshifts have a *rest-frame duration*, $T_{90}/(1 + z)$, shorter than 2 seconds (Zhang et al., 2009a). Multi-wavelength data suggest that these GRBs are related to deaths of massive stars. Lü et al. (2014) showed that by “moving” a normal long GRB to progressively higher redshifts, the rest-frame duration progressively drops due to the tip-of-iceberg effect, so that a short $T_{90}/(1 + z)$ should naturally be expected for high- z GRBs detected near the threshold (with low amplitude).

Similar conclusions were also obtained by Kocevski and Petrosian (2013) and Littlejohns et al. (2013).

HL and LL Long GRBs

Within the long GRB category, based on their luminosities, bursts can be classified into *high-luminosity* (HL) and *low-luminosity* (LL) sub-categories. HL-GRBs typically have a luminosity above $\sim 10^{49}$ erg s $^{-1}$ with significant variability, and are discovered in a wide redshift range. LL-GRBs, on the other hand, typically have a luminosity below $\sim 10^{49}$ erg s $^{-1}$, a long duration, usually a smooth lightcurve, and are discovered at low redshifts (a selection effect due to their low luminosities). Figure 2.53 shows the lightcurves of the LL-GRB 060218 detected with *Swift* BAT, XRT, and UVOT (Campana et al., 2006), which are all very smooth. Table 2.2 lists several well-known LL-GRBs with measured parameters. The separation line between HL- and LL-GRBs is not clearly defined. An indication that they may have separate origins is that there might be a break in the luminosity function of long GRBs: the LL-GRBs are more abundant, and have a steeper luminosity function slope than the low-luminosity portion of the HL-GRBs (Liang et al. 2007a; Sun et al. 2015, see Fig. 2.45). In terms of whether rapid variability is observed, Zhang et al. (2012b) showed that 10 48 erg s $^{-1}$ may be the separation line, below which the lightcurves are smooth without significant variability.

GRBs, X-ray-Rich GRBs, and X-ray Flashes

Based on spectral properties, long GRBs are sometimes further grouped into three sub-categories: GRBs, *X-ray-rich GRBs*, and *X-ray flashes*. There are no distinct peaks in the E_p or hardness ratio distributions to define these sub-categories. The classification into these three sub-categories is therefore subjective. For example, Sakamoto et al. (2005) defined X-ray-rich GRBs and X-ray flashes as those events for which

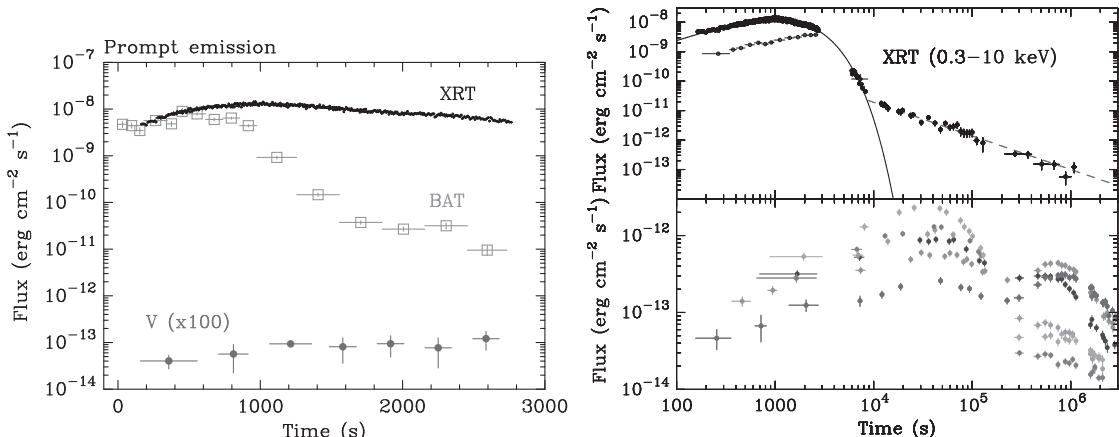


Figure 2.53 Lightcurves of GRB 060218 as detected with *Swift* BAT, XRT, and UVOT. *Left:* Prompt emission lightcurves as detected with BAT and XRT. *Right:* Prompt and afterglow emission lightcurves as detected with XRT and UVOT. From Campana et al. (2006).

Table 2.2 Properties of some low-luminosity GRBs

GRB	z	T_{90} (s)	E_p (keV)	$E_{\gamma,\text{iso},50}$	$L_{\gamma,\text{iso},48}$	References ^a
980425	0.0085	34.9 ± 3.8	~ 120	0.009	~ 0.03	1,2,3
020903	0.251	~ 20	~ 2	0.11	~ 0.7	1,2,4,5
031203	0.105	37.0 ± 1.3	~ 70	1.7	~ 5	1,2,6
060218	0.033	2100 ± 100	~ 5	0.4	~ 0.02	1,2,7
100316D	0.059	~ 1300	~ 18	0.6	~ 0.05	1,2,8
120422A	0.283	5.35 ± 1.4	~ 50	0.45	~ 10	2

^a References: 1. Hjorth and Bloom (2012); 2. Zhang et al. (2012b); 3. Galama et al. (1998); 4. Sakamoto et al. (2004); 5. Soderberg et al. (2004a); 6. Soderberg et al. (2004b); 7. Campana et al. (2006); 8. Starling et al. (2011).

$\log[S_X(2-30 \text{ keV})/\log S_{\gamma}(30-400 \text{ keV})] > -0.5$ and > 0.0 , respectively, to be differentiated from GRBs (which are harder, having $\log[S_X(2-30 \text{ keV})/\log S_{\gamma}(30-400 \text{ keV})] \leq -0.5$). More casually, one may use $E_p = 50 \text{ keV}$ and 30 keV as the separation lines for the three sub-classes. Studies showed that these events form a continuum in their properties, and therefore likely share a similar physical origin (Sakamoto et al., 2005). Most LL-GRBs are X-ray flashes, which might have a somewhat different physical origin (see §2.7.2) within the same progenitor model framework.

Short GRBs and Short GRBs with Extended Emission

Swift BAT has a relatively softer bandpass than BATSE. As a result, some “short” GRBs are found to have soft *extended emission* (or “EE”) following the short, hard spike (Norris and Bonnell, 2006). Such extended emission is temporarily separated from the initial short/hard spike, and lasts for tens to ~ 100 seconds. Short GRBs can therefore be further classified as those with EE and those without. Whether or not these two sub-groups have a distinct physical origin is subject to debate (e.g. Troja et al., 2008; Fong et al., 2010). It is possible that there is a continuous distribution of the flux level of the EE, and the fraction of short GRBs with EE would increase with softer, more sensitive detectors. Indeed, a good fraction of short GRBs have *internal plateaus* typically lasting for ~ 100 seconds (Rowlinson et al., 2010, 2013). The so-called “extended emission” detected in the BAT band could be simply the internal plateau emission when the emission is bright and hard enough (Lü et al., 2015).

Classification Schemes Based on Afterglow Data

Based on optical afterglow data, GRBs can be classified into *optically bright* and *optically dark* GRBs. As discussed in §2.2.3, about 30–50% of GRBs are optically dark. They are defined to be “darker” than the lowest predicted flux level based on the observed X-ray flux and spectral index (Jakobsson et al., 2004; Rol et al., 2005). The prompt emission properties of dark GRBs are usually not distinctly different from those of optically bright GRBs. The physical reason for the optical darkness of most dark GRBs is likely dust

extinction, even though a small fraction of dark GRBs may be high-redshift ones whose optical light is absorbed by neutral hydrogen in the intergalactic medium at $z > 6$ (Perley et al., 2009).

One may also classify GRBs into *SN-GRBs* (those associated with SNe) and *SN-less GRBs* (those not associated with SNe). The definition of the latter is subjective. For HL-GRBs at relatively high redshifts, usually an associated SN (even if one is there) cannot be firmly detected because of the faintness of the SN. On the other hand, similar GRBs at relatively low redshifts, e.g. GRB 030329 (Stanek et al., 2003; Hjorth et al., 2003) and GRB 130427A (Xu et al., 2013), were found to be associated with Type Ic SNe. As a result, it is generally believed that most long GRBs are not intrinsically SN-less even if no associated SN is detected. A small fraction of long GRBs, e.g. GRB 060614 and GRB 060505, are intrinsically SN-less (Gal-Yam et al., 2006; Fynbo et al., 2006a; Della Valle et al., 2006). However, they likely have a different origin from the typical long GRBs (Gehrels et al., 2006; Zhang et al., 2007b).

Based on the X-ray afterglow data, one can classify GRBs into those with X-ray flares and those without. There are no distinct differences in the prompt emission properties for these two sub-classes.

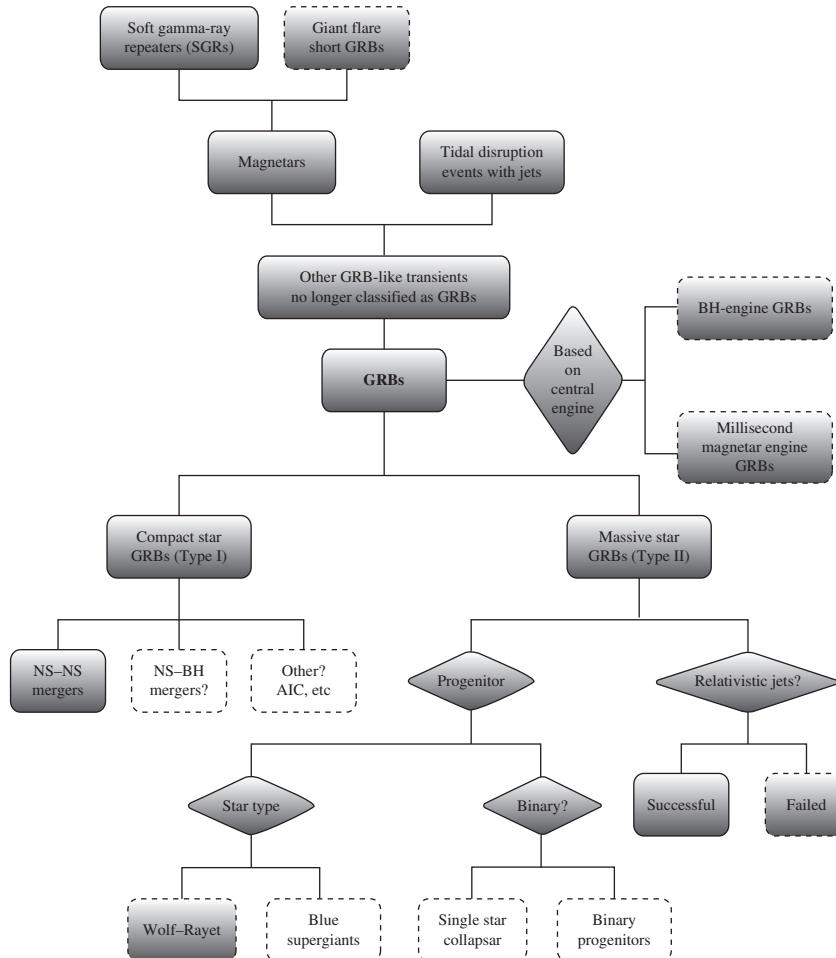
Again based on X-ray afterglow data, one may also classify GRBs into those having a “canonical” multi-segment lightcurve (Zhang et al., 2006; Nousek et al., 2006) and those having a simple power-law decay (Evans et al., 2009; Liang et al., 2009). The latter seems to be somewhat more energetic than the former (e.g. Lü and Zhang, 2014), even though there exists a significant overlap in the properties between the two groups (Liang et al., 2009).

2.7.2 Physical Classification Schemes

Massive Star (Type II) GRBs vs. Compact Star (Type I) GRBs

Even though the long/soft vs. short/hard dichotomy has been known since the BATSE era, it was not until the discoveries of the afterglow emission of both classes of GRBs that the physical origins of these events were unveiled. The majority of GRBs can be included into two broad physical classes: one related to the deaths of massive stars and the other not associated with massive stars (Fig. 2.54).

Initially there was a cozy picture: long-duration GRBs are related to the deaths of massive stars, whereas short-duration GRBs are not related to massive stars. The supporting evidence can be summarized as the following: observations led by *BeppoSAX*, *HETE-2*, and *Swift* suggested that at least some long GRBs are associated with Type Ic SNe (e.g. Galama et al., 1998; Hjorth et al., 2003; Stanek et al., 2003; Campana et al., 2006; Pian et al., 2006). Most long GRB host galaxies are dwarf star-forming galaxies, and long GRBs typically reside in the brightest regions (which have the highest specific star formation rate) in the host galaxies (Fruchter et al., 2006). These facts establish the connection between long GRBs and deaths of massive stars (Woosley, 1993). The breakthrough led by *Swift* unveiled that some nearby short GRBs (or short GRBs with EE) have host galaxies that are elliptical or early type, with little star formation (Gehrels et al., 2005; Barthelmy et al.,

**Figure 2.54**

GRB physical classification schemes. Solid shaded boxes denote the most robust classification schemes supported by observational data; dashed shaded boxes denote secondary, reasonable classification schemes indirectly supported by observations; dashed open boxes denote classification schemes speculated but not fully confirmed by the data; rhombus-shaped boxes denote classification criteria. Figure courtesy Jared Rice.

2005a; Berger et al., 2005a). Some short GRBs have star-forming host galaxies, but the local specific star formation rate is not high (Fox et al., 2005). The GRB site usually has a large offset from the center of the galaxy (Fong et al., 2010). All these point towards another type of progenitor that does not involve massive stars. Rather, these GRBs are likely related to compact stars (neutron stars or black holes), with the leading scenarios being NS-NS or NS-BH mergers (e.g. Eichler et al., 1989; Paczyński, 1991; Narayan et al., 1992).

Such a neat picture was soon destroyed by several observations. GRB 060614 and GRB 060505 are both nearby, long-duration GRBs, but deep searches showed no association of a supernova accompanying either GRB (Gehrels et al., 2006; Gal-Yam et al., 2006; Fynbo

et al., 2006a; Della Valle et al., 2006), unlike other nearby long GRBs. Moreover, the γ -ray properties of GRB 060614 share many features with short GRBs (Gehrels et al., 2006), and it would resemble GRB 050724 (a smoking gun “short” GRB that has a definite non-massive star origin) if it were somewhat less luminous (Zhang et al., 2007b). Although theoretically some massive star core-collapse events can have faint supernova signals (e.g. Nomoto et al., 2006), the available data for GRB 060614 seem not to demand such a scenario since, except for the long duration, all the other properties are similar to those of other nearby short GRBs. Rather, it suggests that some GRBs that are not related to massive stars can have a long duration. Later, it was noticed that the three GRBs with the highest redshifts, i.e. GRB 080913 at $z = 6.7$ (Greiner et al., 2009), GRB 090423 at $z = 8.2$ (Tanvir et al., 2009; Salvaterra et al., 2009), and GRB 090429B at $z = 9.4$ (Cucchiara et al., 2011a) all have a *rest-frame duration* $T_{90}/(1+z)$ shorter than 2 seconds. Yet various arguments suggest that they still originate from deaths of massive stars (Zhang et al., 2009a). Later, an observer-frame short GRB 090426 at $z = 2.609$ was discovered, which shared many properties of long GRBs with a massive star origin (Levesque et al., 2010; Antonelli et al., 2009; Xin et al., 2011; Thöne et al., 2011). Independent arguments suggest that at least some short GRBs, especially those at high redshifts with high luminosities, are probably not related to compact star mergers (Zhang et al., 2009a; Virgili et al., 2011a; Cui et al., 2012; Bromberg et al., 2012).

These observations force the physical classification scheme to be somewhat different from the phenomenological duration classification regime. Zhang (2006) and Zhang et al. (2007b) proposed classifying GRBs physically into *massive star type (Type II)* and *compact star type (Type I)*. The challenge is how to identify the physical class based on the data. Zhang et al. (2009a) summarized a list of multi-wavelength observational criteria that could be connected to the physical nature of a GRB, and suggested applying them to identify the physical class of a GRB. Table 2.3 (from Zhang et al. 2009a) summarizes various observational criteria that would be helpful in judging the physical category of a GRB. In particular, the observational criteria that are most related to the physical nature of a GRB include supernova association, host galaxy properties, as well as the location within the host galaxy. Figure 2.55 shows a flowchart of applying multi-wavelength criteria to diagnose the physical category of a GRB (Zhang et al., 2009a). The chart was applicable for inferring the physical type of a large sample of long and short GRBs before 2011 (Kann et al., 2010, 2011).

The multi-wavelength data cannot be obtained immediately when a GRB is detected. So looking for extra information based on prompt γ -ray data to infer the physical class of GRBs is important. Several attempts have been made. For example, Lü et al. (2010) showed that for GRBs with z measurements, the parameter $\varepsilon \equiv E_{\gamma,\text{iso},52}/E_{p,z,2}^{5/3}$ has a clearer bimodal distribution. The high- ε vs. low- ε categories are found to be more closely related to massive star GRBs vs. compact star GRBs, respectively. Similarly, Qin and Chen (2013) proposed using the Amati relation to classify GRBs.

Bromberg et al. (2012) found that there exists a plateau in the dN/dT_{90} duration distribution of GRBs (for all the samples with different detectors including BATSE, *Swift*, and *Fermi*). They argued that this is direct evidence of a massive-star-GRB jet propagating inside the progenitor star. The idea is that it takes about 10 seconds for a (hydrodynamic)

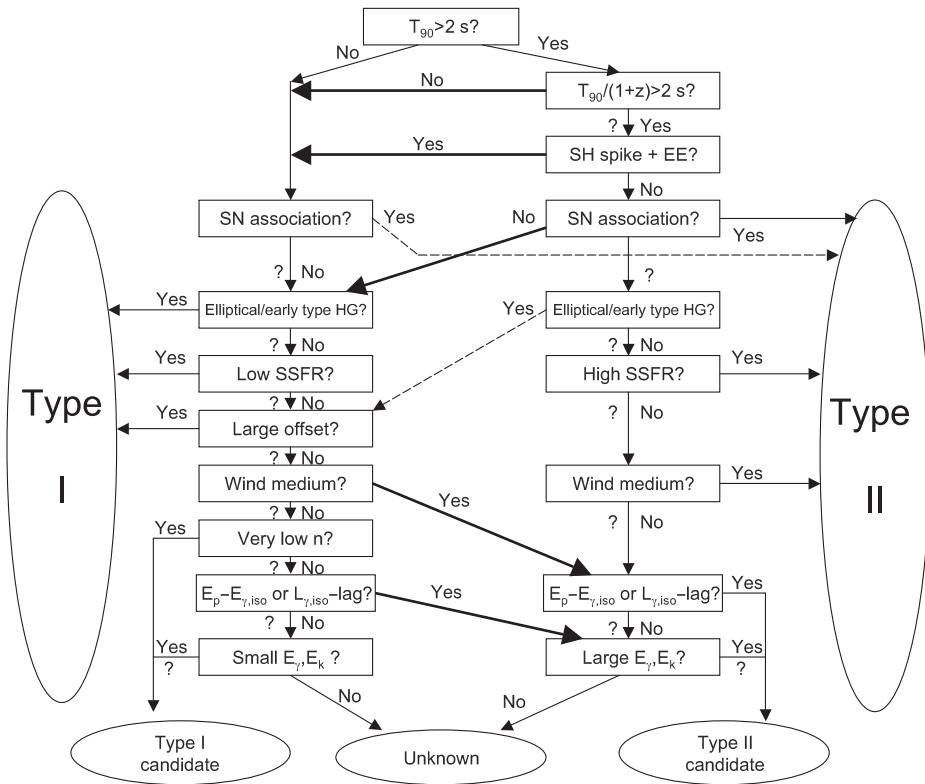
Table 2.3 Observational criteria for physically classifying GRBs. Adapted from Zhang et al. (2009a)

Criterion	Type I	Type II
Duration	Usually short, but can have extended emission	Usually long, can be short
Spectrum	On average hard (soft tail)	On average soft
Spectral lag	Usually short	Usually long, can be short
$E_{\gamma,\text{iso}}$	On average low	On average high
$E_p - E_{\gamma,\text{iso}}$	Usually off Amati relation	Usually on the relation
$L_{\gamma,\text{iso}}^p - \text{lag}$	Usually off Norris relation	Usually on the relation
SN association	No	Yes
Medium type	Usually low- n ISM	Wind or usually high- n ISM
$E_{K,\text{iso}}$	On average low	On average high
Jet angle	On average wide	On average narrow
E_{γ} and E_K	On average low	On average high
Host galaxy type	Elliptical, early, or late	Usually late
SSFR	From low to high	Usually high
Offset	Outskirts or “hostless”	Well inside
z -distribution	On average low z	On average high z

jet to penetrate through a typical Wolf–Rayet star. If the central engine duration has a uniform distribution spanning a wide range, then a plateau in the dN/dT_{90} can emerge below the jet propagation time scale (Bromberg et al., 2012). Applying this formalism to define the massive star population, they found that there should be a noticeable contamination of massive star GRBs in the observed short GRBs. This conclusion is generally consistent with the suggestions of Zhang et al. (2009a), Virgili et al. (2011a), and Cui et al. (2012), which are based on very different arguments.

Sub-categories in Massive Star GRBs: Successful vs. Choked Jets

For GRBs associated with massive stars, growing evidence suggests that the HL-GRBs and LL-GRBs may have somewhat different physical origins. The first piece of evidence is that LL-GRBs have a much higher event rate density in the local universe and may form a distinct population (Liang et al., 2007a; Virgili et al., 2009; Sun et al., 2015). Next, the smooth lightcurves (e.g. Fig. 2.53) of LL-GRBs are very different from those of HL-GRBs, which are usually much more erratic, with rapid variability (e.g. Fig. 2.4). Third, a thermal component was discovered in the X-ray spectrum of GRB 060218 during the prompt emission phase, which is consistent with having a shock breakout origin (Campana et al., 2006). Several authors proposed that not only GRB 060218, but also all LL-GRBs are of a shock breakout origin (e.g. Nakar and Sari, 2012; Bromberg et al., 2012). A jet is probably also launched from the central engine in these GRBs. However, the engine lasts for a time that is shorter than the time for the jet head to emerge from the envelope. After the engine stops, the jet loses power, and starts to spread out within the envelope. As the

**Figure 2.55**

A flowchart for applying multiple observational criteria to diagnose the physical category of a GRB. From Zhang et al. (2009a).

shock breaks out from the star, the shock is at most mildly relativistic, leading to a broad, smooth pulse. In contrast, HL-GRBs are successful jets launched from massive stars whose central engine lasts for a longer time scale than the jet propagation time inside the star. The rapid variability is the imprint of erratic behavior at the central engine, which is manifested in the emission properties of the jet.

Within this picture, HL- and LL-GRBs do not need to have distinct types of progenitors. They can be the manifestations of a continuous family of jets with a distribution of central engine activity times. Indeed, the associated SNe for LL-GRBs and the two nearby HL-GRBs (GRB 030329 and GRB 130427A) are all Type Ic, with similar properties, even though SN 2006aj (the counterpart of GRB 060218) had a smaller energy and ejecta mass than other GRBs (including other LL-GRBs such as GRB 980425), which may have a neutron star rather than a black hole central engine (Mazzali et al., 2006).

Other Speculated Sub-categories in Massive Star GRBs

The leading candidate progenitor of massive star GRBs is a Wolf–Rayet star. No direct observational evidence is available to prove such a progenitor, but some indirect evidence seems to support it. The key evidence is that all GRBs with SN associations have a

broad-line Type Ic SN companion, suggesting that the massive-star-GRB progenitor has likely lost its extended H and He envelopes before core collapse happened. Another indirect supporting observation is the existence of a plateau in dN/dT_{90} distribution, as pointed out by Bromberg et al. (2012). The onset time of the plateau in the duration distribution is consistent with the jet propagation time inside a Wolf–Rayet star.

The discovery of *ultra-long GRBs* with durations in excess of 1000 seconds led some researchers to propose that they might have a different type of progenitor star which has a much larger size. One often discussed possibility is a *blue supergiant* (e.g. Gendre et al., 2013; Levan et al., 2014b; Piro et al., 2014). While launching a relativistic jet from such a large star is not impossible (Mészáros and Rees, 2001), more data are needed to make the case for a blue supergiant progenitor star. Zhang et al. (2014) suggested that, when X-ray flares are considered, many *Swift* GRBs actually have long central engine activity times (defined as t_{burst}). The ultra-long GRBs may be simply those flare-dominated GRBs whose flares are bright enough to be detected in γ -rays. Indeed, afterglow observations also suggest that ultra-long GRBs are not special compared to other long GRBs (Virgili et al., 2013). Also, the long duration of a burst may not necessarily point towards a large star. A long-lasting central engine (e.g. a spinning-down magnetar) within a small star (e.g. Wolf–Rayet) could also power an ultra-long GRB.

Possible Sub-categories in Compact Star GRBs

Within the compact star GRB category, two leading progenitor models are NS–NS and NS–BH mergers. The gravitational wave observations already confirmed the NS–NS merger progenitor (GW170817/GRB 170817A association) and will tell whether NS–BH mergers can also make short GRBs. The electromagnetic signals do not show the clear dichotomy that demands further classification of compact star GRBs into these two sub-categories. Some indirect evidence, e.g. extended emission (e.g. Norris and Bonnell, 2006), internal plateaus (e.g. Rowlinson et al., 2010, 2013; Lü et al., 2015), as well as X-ray flares (e.g. Barthelmy et al., 2005a) following short GRBs, suggests that the central engine of (at least some) short GRBs may be a stable or supra-massive magnetar (e.g. Dai et al., 2006; Gao and Fan, 2006; Metzger et al., 2008; Zhang, 2013; Gao et al., 2013b; Yu et al., 2013; Metzger and Piro, 2014). Within such a scenario, the progenitor has to be a NS–NS merger, since a NS–BH merger would not leave behind a magnetar.

In principle, the compact star GRBs can have a non-merger origin. For example, the scenario invoking *accretion-induced collapse (AIC)* of a neutron star to a black hole (e.g. Qin et al., 1998; MacFadyen et al., 2005; Dermer and Atoyan, 2006) seems to satisfy most of the observational constraints.

GRBs with Different Central Engines: Hyper-accreting Black Holes vs. Millisecond Magnetars

The central engine(s) of GRBs have not been identified. Hyper-accreting black holes and millisecond magnetars have been speculated as possible candidates. Observationally, X-ray afterglow data of some GRBs show the features (e.g. internal plateaus or external plateaus

with the correct slopes, Troja et al. 2007; Lyons et al. 2010; Rowlinson et al. 2010; Gompertz et al. 2014; Lü and Zhang 2014; Lü et al. 2015) that are consistent with the prediction of a magnetar central engine. Some other GRBs do not have these features. Therefore it is tempting to classify GRBs into those powered by magnetars and those powered by black holes. Lü and Zhang (2014) made an effort to characterize long GRBs into Gold, Silver, and Aluminum samples of magnetar candidates, along with some non-magnetar GRBs. They found that magnetar GRBs are statistically less energetic than black hole GRBs. It is interesting to note that both long and short GRBs can have both types of central engine, so that they may both be further separated into two sub-categories based on the central engine type. However, in practice it is very difficult to prove a central engine type for a particular GRB based on its observational properties.

Other GRB-like Transients

Besides the two well-known physical categories (massive star origin vs. compact star origin), the observed GRBs also have contaminations from other physical types. In history, these contaminating bursts, whenever identified, were given other names, so that they are no longer classified as GRBs in modern language.

Back in the pre-BATSE era, one source from the Large Magellanic Cloud was found to emit repeating bursts of γ -rays. Later, more sources with similar properties were identified. These sources, named *soft gamma-ray repeaters* (SGRs), are now commonly interpreted as slowly rotating, strongly magnetized neutron stars dubbed “magnetars” (Duncan and Thompson, 1992; Thompson and Duncan, 1995, 1996; Kouveliotou et al., 1998). These magnetars occasionally give rise to “giant flares”, which display a short hard spike followed by an extended soft tail (e.g. Palmer et al., 2005). These magnetar giant flares, if observed from nearby galaxies, would appear as short GRBs (Hurley et al., 2005). Searches for the associations of short GRBs with nearby galaxies have been carried out. It is found that the fraction of these magnetar-giant-flare-origin events to cosmological short GRBs is less than 10–25% (Tanvir et al., 2005). The short GRB 051103 triangulated by the Inter-Planetary Network (IPN) was found to fall in the direction of the nearby M81/M82 galaxy group, which is a good candidate for a short GRB with a SGR giant flare origin (Frederiks et al., 2007).

“GRB 110328A” triggered the *Swift* BAT multiple times for 2–3 days. Soon it was realized that it is not a regular GRB. Broad-band data suggested that it originated from a jetted *tidal disruption event*, i.e. a jet is launched from a super-massive black hole that tidally disrupted a star (Bloom et al., 2011; Burrows et al., 2011; Levan et al., 2011; Zauderer et al., 2011). It was then renamed *Swift* J16449.3+573451, or Sw J1644+57 for short. Another candidate event of this type, *Swift* J2058.4+0516 or Sw J2058+05, was also reported (Cenko et al., 2012).

An oddball burst, GRB 101225, occurred on Christmas Day in 2010, so it is also known as the “Christmas burst”. It had a series of peculiar properties that were different from traditional long GRBs. In particular, it had an ultra-long duration, smooth lightcurve, no host galaxy and no redshift, so the debate about its distance scale (similar to the distance debate of GRBs in the pre-*Beppe-SAX* era) came alive once more. More interestingly,

its afterglow shows a peculiar thermal-dominated feature (Thöne et al., 2011). The suggested progenitors vary from a helium star–neutron star merger at a moderate redshift (Thöne et al., 2011) to a comet falling onto a Galactic neutron star (Campana et al., 2011). Levan et al. (2014a) included it as a member of the ultra-long GRBs, which they proposed might have a large-size massive star (e.g. blue supergiant) progenitor. In any case, this burst presents the case of a unique, peculiar burst, whose nature will take years to unveil.