11 Central Engine

After the catastrophic destruction of the progenitor system, a *central engine* must form to power a GRB relativistic jet. The leading candidate is a hyper-accreting black hole, but recent observations seem to call for a non-black hole engine, at least for some GRBs, i.e. likely a rapidly spinning, strongly magnetized neutron star known as a millisecond magnetar. It is possible that both types of central engine may operate, possibly even in both types (massive star core collapse and compact star merger) of progenitor systems. In §11.1, the general observational requirements for a GRB central engine are listed. The hyper-accreting black hole engine is introduced in §11.2, and the millisecond magnetar central engine is introduced in §11.3. In both sections, the cases for both the massive star GRBs and compact star GRBs are discussed. In §11.4, a more exotic scenario, i.e. the quark star central engine, is discussed. In §11.5, some ideas that account for the late central engine activities of GRBs (as manifested by X-ray flares and the internal X-ray plateau) are discussed.

11.1 General Observational Constraints

A successful GRB central engine model should satisfy the following observational constraints.

- The engine should be able to power an energetic ($E_{\gamma,\rm iso} \sim 10^{49}$–$10^{55}$ erg) and luminous ($L_{\gamma,\rm iso} \sim 10^{46}$–$10^{54}$ erg s^{-1}) event;
- The engine should be able to launch a clean outflow: the jet can reach a Lorentz factor $\Gamma > 100$;
- The engine should be able to produce GRB emission with diverse temporal behavior, ranging from smooth, single-pulse events to erratic, highly variable events. In most cases, the engine should be intermittent and produce rapid variability;
- The engine should be able to restart itself at late times to power X-ray flares occurring after the prompt emission ends. The reactivation time ranges from hundreds to over ten thousand seconds in some cases, as observed by *Swift*;
- In some events, the engine should be able to power an extended *internal plateau* followed by an extremely steep drop at the end, as observed by *Swift*. During the plateau the emission is steady, although with small-scale temporal variability;

- The engine should be able to launch jets with a variety of compositions, ranging from matter-dominated fireballs to Poynting-flux-dominated outflows, as indicated by prompt emission observations by *Fermi*.

In the following sections, we will discuss three possible engines of GRBs proposed in the literature:

- Hyper-accreting black holes;
- Rapidly spinning magnetars;
- Quark stars (more exotic).

11.2 Hyper-Accreting Black Holes

11.2.1 General Consideration

If a GRB is powered by accretion onto a stellar-mass black hole (BH), a very high accretion rate is required. The ultimate jet power should come from either the gravitational potential energy of the accreted material or the spin energy of the BH. For the latter scenario, the rate of tapping the BH spin energy also depends on the accretion rate. In general, one may write

$$L_{\rm GRB} = \zeta \dot{M} c^2 = 1.8 \times 10^{51}\ {\rm erg\ s^{-1}}\ \zeta_{-3} \left(\frac{\dot{M}}{1 M_\odot\ {\rm s^{-1}}} \right). \tag{11.1}$$

Given a reasonable efficiency (say $\zeta = 10^{-3}\zeta_{-3}$), the required accretion rate for a typical GRB is 0.1–1$M_\odot$ s^{-1}. In general, a broader range of the accretion rate, e.g. (10^{-3}–several)$M_\odot$ s^{-1} may be possible.

With such a high accretion rate, the accretion flow is extremely hot. At a high enough accretion rate, or at a radius sufficiently close to the BH, the temperature is so high that e^-/e^+ capture processes,

$$e^- + p \rightarrow n + \nu_e, \quad e^+ + n \rightarrow p + \bar{\nu}_e, \tag{11.2}$$

become dominant. Abundant neutrinos are generated within the disk, which escape and therefore cool the disk. The accretion flow in this regime is called a neutrino-dominated accretion flow (NDAF). At a lower accretion rate, or at a larger radius from the BH, neutrino cooling is not ignited. Heat and photons are trapped and advected inside the disk, making a thick torus. The accretion flow in this regime is called an advection-dominated accretion flow (ADAF). The left panel of Fig. 11.1 shows a cartoon picture of the disk structure near the BH central engine of a GRB (Chen and Beloborodov, 2007).

The accreting BH may carry a large angular momentum. Such a rapidly spinning BH may be formed in the rapidly rotating core of the progenitor star. Subsequent accretion onto the BH would further spin up the BH. Because of the large accretion rate involved, the BH spin would increase rapidly and achieve a large value. If a strong magnetic field threads the spinning BH and is connected to an external astrophysical load, the BH spin

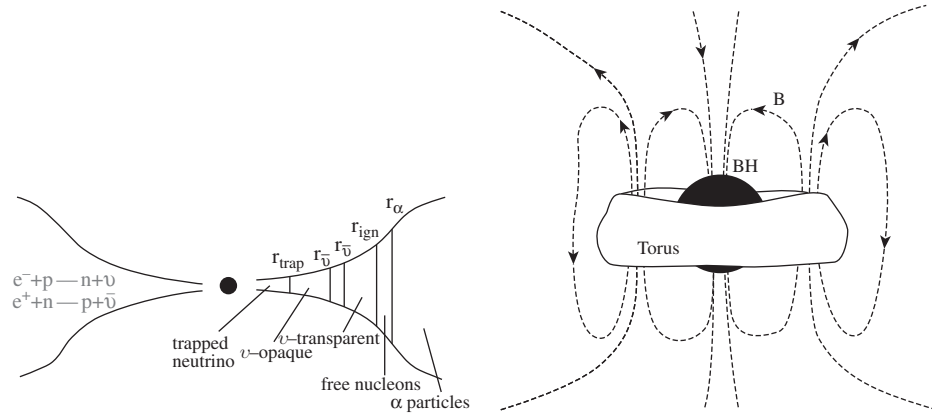

Figure 11.1 *Left:* The structure of a hyper-accreting BH disk. The radius $r_{\rm ign}$ separates the NDAF (inside) and ADAF (outside) regimes. Reproduced from Figure 10 in Chen and Beloborodov (2007) with permission. ©AAS. *Right:* A possible magnetic field configuration of a BH-torus system. From van Putten (2001).

energy may be tapped via the *Blandford–Znajek (BZ) mechanism* (Blandford and McKee, 1976). In such a case, the spin energy of the BH would be the ultimate power source of the jet.

In general, a GRB jet may be launched from a hyper-accreting BH via two well-known mechanisms:

- Neutrinos (ν) and anti-neutrinos ($\bar\nu$) generated in a NDAF would annihilate above the disk and produce photons and electron–positron pairs. Neutrinos can also strip baryons from the disk (e.g. Qian and Woosley, 1996; Popham et al., 1999; Lei et al., 2013). Therefore, a hot "fireball" with a small baryon contamination would form above the disk. Neutrino annihilations have relatively large optical depths near the spin axis, so that a broad, relativistic jet may be launched. For a BH central engine in a massive star GRB, the stellar envelope would further collimate the outflow, making a narrow jet of a few degrees as observed in long GRBs.
- For a highly magnetized accretion disk and a rapidly spinning BH with magnetic field lines threading the BH horizon and connecting to a remote astrophysical load, the spin energy of the BH may be tapped through the BZ mechanism. A Poynting-flux-dominated jet would be launched, which may be self-collimated by a strong toroidal field produced by the rapidly rotating central engine.

11.2.2 $\nu\bar\nu$ Annihilation in a Neutrino-Dominated Accretion Flow (NDAF)

Rich physics (relativistic hydrodynamics, neutrino physics, nucleosynthesis, and thermodynamics) is required to describe the radial and vertical structures of a NDAF (Liu et al., 2017b, for a review). For a one-dimensional treatment, the following solutions for a GRB accretion disk may be obtained (e.g. Narayan et al., 2001; Beloborodov, 2003b):

$$\rho = (1.2 \times 10^{14}\ \mathrm{g\ cm^{-3}})\, \alpha_{-2}^{-1.3} \dot{M}_{-1} \left(\frac{M}{3M_\odot}\right)^{-1.7} \left(\frac{r}{r_s}\right)^{-2.55}, \tag{11.3}$$

$$T_c = (3 \times 10^{10}\ \mathrm{K})\, \alpha_{-2}^{0.2} \left(\frac{M}{3M_\odot}\right)^{-0.2} \left(\frac{r}{r_s}\right)^{-0.3}, \tag{11.4}$$

$$v_r = (2 \times 10^{6}\ \mathrm{cm\ s^{-1}})\, \alpha_{-2}^{1.2} \left(\frac{M}{3M_\odot}\right)^{-0.2} \left(\frac{r}{r_s}\right)^{0.2} \tag{11.5}$$

for a NDAF, and

$$\rho = (6 \times 10^{11}\ \mathrm{g\ cm^{-3}})\, \alpha_{-2}^{-1} \dot{M}_{-1} \left(\frac{M}{3M_\odot}\right)^{-2} \left(\frac{r}{r_s}\right)^{-1.5}, \tag{11.6}$$

$$T_c = (3 \times 10^{11}\ \mathrm{K})\, \alpha_{-2}^{-1/4} \left(\frac{M}{3M_\odot}\right)^{-0.5} \left(\frac{r}{r_s}\right)^{-5/8}, \tag{11.7}$$

$$v_r = (10^{8}\ \mathrm{cm\ s^{-1}})\, \alpha_{-2} \left(\frac{r}{r_s}\right)^{-0.5} \tag{11.8}$$

for an ADAF. Here ρ, T_c, and v_r are the density, temperature at the equatorial plane, and radial velocity of the accretion flow, respectively, M is the BH mass, $\dot{M} = (0.1 M_\odot\ \mathrm{s^{-1}})\dot{M}_{-1}$ is the accretion rate, r is the distance from the BH in the equatorial plane, and $r_s = 2GM/c^2$ is the Schwarzschild radius of the BH.

The disk is in the NDAF regime when $\dot{M} \geq \dot{M}_{\rm ign}$ is satisfied, where $\dot{M}_{\rm ign}$ is the critical accretion rate above which neutrino cooling is ignited. It depends on the BH spin. For a certain accretion rate, there are several characteristic radii (Chen and Beloborodov, 2007), see Fig. 11.1, typically with the order of $r_\alpha > r_{\rm ign} > r_\nu > r_{\rm trap}$, where r_α is the radius inside which most of the α-particles are disintegrated, $r_{\rm ign}$ is the radius inside which neutrino flux rises dramatically, r_ν is the radius inside which the disk becomes ν-opaque, and $r_{\rm trap}$ is the radius inside which neutrinos are trapped and advected into the BH.

For a NDAF, the total pressure in the disk includes five terms,

$$P = P_{\rm rad} + P_{\rm gas} + P_{\rm deg} + P_\nu + P_B, \tag{11.9}$$

where $P_{\rm rad} = (1/3)aT^4$ is the radiation pressure, $P_{\rm gas} = \sum_j n_j kT$ is the gas pressure, $P_{\rm deg} \propto \rho^{4/3}$ is the relativistic electron degeneracy pressure, $P_\nu = (1/3)u_\nu$ is the neutrino pressure, where u_ν is the neutrino energy density, which takes a "bridging" formula connecting the optically thin and optically thick regimes (Di Matteo et al., 2002; Kohri et al., 2005; Liu et al., 2007), and $P_B = B^2/8\pi$ is the magnetic pressure.

The energy balance in the disk gives ($Q^\pm$ has the dimension of power per unit area)

$$Q^+ = Q^-, \tag{11.10}$$

where the heating term

$$Q^+ = Q_{\rm vis} \tag{11.11}$$

originates from viscous heating, and the cooling term,

$$Q^- = Q_\nu + Q_{\rm ph} + Q_{\rm rad} + Q_{\rm adv}, \tag{11.12}$$

includes the terms from neutrino cooling (Q_ν), photo-disintegration ($Q_{\rm ph}$), photon radiation ($Q_{\rm rad}$), and advection ($Q_{\rm adv}$), respectively. The neutrino emission luminosity can be calculated through integrating the neutrino cooling rate across the surface of the disk, i.e.

$$L_\nu = 4\pi \int_{r_{\rm in}}^{r_{\rm out}} Q_\nu r dr. \tag{11.13}$$

Neutrinos and anti-neutrinos annihilate and produce e^+e^- pairs and photons. These pairs and photons, along with a small fraction of baryons entrained, make a fireball. The neutrino annihilation luminosity $L_{\nu\bar{\nu}}$ defines the luminosity of a GRB.

The $\nu\bar{\nu}$ annihilation luminosity can be calculated by considering individual annihilation luminosities from different locations in the disk (e.g. Ruffert et al., 1997; Popham et al., 1999; Rosswog et al., 2003; Liu et al., 2007; Lei et al., 2009; Zalamea and Beloborodov, 2011). One may model the disk as a grid of cells. Each cell k has its neutrino mean energy $\varepsilon^k_{\nu_i}$ and luminosity $l^k_{\nu_i}$. Suppose the neutrinos from cell k annihilate with anti-neutrinos from cell k' at a location d_k above (or below) the disk at an angle $\theta_{kk'}$. Then the neutrino annihilation power at this point may be calculated:

$$\begin{aligned} l_{\nu\bar{\nu}} = {} & A_1 \sum_k \frac{l^k_{\nu_i}}{d_k^2} \sum_{k'} \frac{l^k_{\nu_i}}{d_k^2} (\epsilon^k_{\nu_i} + \epsilon^{k'}_{\bar{\nu}_i})(1-\cos\theta_{kk'})^2 \\ & + A_2 \sum_k \frac{l^k_{\nu_i}}{d_k^2} \sum_{k'} \frac{l^k_{\nu_i}}{d_k^2} \frac{\epsilon^k_{\nu_i} + \epsilon^{k'}_{\bar{\nu}_i}}{\epsilon^k_{\nu_i}\epsilon^{k'}_{\bar{\nu}_i}} (1-\cos\theta_{kk'}), \end{aligned} \tag{11.14}$$

where $A_1 \approx 1.7 \times 10^{-44}$ cm erg^{-2} s^{-1}, and $A_2 \approx 1.6 \times 10^{-56}$ cm erg^{-2} s^{-1}. The two coefficients are the result of the convolution of angle-dependent cross sections and number densities. The power is the strongest in the axis direction, but there is no sharp boundary defining a jet.

The total $\nu\bar{\nu}$ annihilation luminosity from the system can be obtained through integrating across the disk:

$$L_{\nu\bar{\nu}} = 4\pi \int_{r_{\rm in}}^{\infty} \int_H^{\infty} l_{\nu\bar{\nu}} r dr dz. \tag{11.15}$$

The efficiency for neutrino annihilation is defined as

$$\eta_{\nu\bar{\nu}} = \frac{L_{\nu\bar{\nu}}}{L_\nu}. \tag{11.16}$$

The $\nu\bar{\nu}$ annihilation luminosity depends on the mass accretion rate $\dot{M}$, BH mass, and the dimensionless spin a_*. There is no simple analytical derivation of these dependencies. One needs to perform numerical calculations by varying various parameters and to fit the results with some simple functional forms. Different authors reached somewhat different scalings by invoking different degrees of complication in the modeling. The following scalings have been published in the literature:

- Fryer et al. (1999) fitted the numerical results of Popham et al. (1999) and obtained

$$\log L_{\nu\bar{\nu}}({\rm erg\ s^{-1}}) \approx 53.4 + 3.4 a_* + 4.89 \log \dot{m}, \tag{11.17}$$

where $\dot{m} = \dot{M}/(M_\odot\ {\rm s}^{-1})$.

- Zalamea and Beloborodov (2011) derived the NDAF luminosity in different accretion rate regimes, i.e.

$$L_{\nu\bar{\nu}} \approx 5.7\times10^{52}\, x_{\rm ms}^{-4.8}\, m^{-3/2} \left\{ \begin{array}{ll} 0 & \text{for } \dot{m} < \dot{m}_{\rm ign} \\ \dot{m}^{9/4} & \text{for } \dot{m}_{\rm ign} < \dot{m} < \dot{m}_{\rm trap} \\ t m_{\rm trap}^{9/4} & \text{for } \dot{m} > \dot{m}_{\rm trap} \end{array} \right\} \ {\rm erg\ s^{-1}}, \quad (11.18)$$

where $x_{\rm ms} = r_{\rm ms}/r_g$, $r_{\rm ms}$ is radius of the last marginally stable orbit, $r_g = GM/c^2 = (1/2)r_s$, $\dot{m}_{\rm ign} = \dot{M}_{\rm ign}/(M_\odot\ {\rm s^{-1}})$, $\dot{m}_{\rm trap} = \dot{M}_{\rm trap}/(M_\odot\ {\rm s^{-1}})$, $\dot{M}_{\rm ign} = K_{\rm ign}\alpha_{-1}^{5/3}$, $\dot{M}_{\rm trap} = K_{\rm trap}\alpha_{-1}^{1/3}$, and $\alpha \sim 0.1$ is the viscosity parameter. The coefficients $K_{\rm ign}$ and $K_{\rm trap}$ depend on the BH spin. For example, one has $K_{\rm ign} = 0.071 M_\odot\ {\rm s^{-1}}$ and $K_{\rm trap} = 9.3 M_\odot\ {\rm s^{-1}}$ for $a_* = 0$, and $K_{\rm ign} = 0.021 M_\odot\ {\rm s^{-1}}$ and $K_{\rm trap} = 1.8 M_\odot\ {\rm s^{-1}}$ for $a_* = 0.95$ (Chen and Beloborodov, 2007).
- Based on one-dimensional global solutions, Xue et al. (2013) obtained

$$\log L_{\nu\bar{\nu}}({\rm erg\ s^{-1}}) \approx 49.5 + 2.45 a_* + 2.17 \log \dot{m}. \quad (11.19)$$

Including the dependence on the BH mass, Liu et al. (2016a) obtained

$$\log L_{\nu\bar{\nu}}({\rm erg\ s^{-1}}) \approx 52.98 + 3.88 a_* - 1.55 \log m + 5.0 \log \dot{m}, \quad (11.20)$$

where $m = M/M_\odot$.
- Through a grid of simulations, Lei et al. (2017) presented a complete fitting formula for a wide range of accretion rate $\dot{m}$, black hole spin a_*, and mass m_*, which reads

$$L_{\nu\bar{\nu}} \simeq L_{\nu\bar{\nu},{\rm ign}} \left[\left(\frac{\dot{m}}{\dot{m}_{\rm ign}} \right)^{-\alpha_{\nu\bar{\nu}}} + \left(\frac{\dot{m}}{\dot{m}_{\rm ign}} \right)^{-\beta_{\nu\bar{\nu}}} \right]^{-1} \times \left[1 + \left(\frac{\dot{m}}{\dot{m}_{\rm trap}} \right)^{\beta_{\nu\bar{\nu}} - \gamma_{\nu\bar{\nu}}} \right]^{-1}. \quad (11.21)$$

Here

$$\begin{aligned} L_{\nu\bar{\nu},{\rm ign}} &= 10^{(48.0+0.15a_*)} \left(\frac{m_*}{3} \right)^{\log(\dot{m}/\dot{m}_{\rm ign})-3.3} {\rm erg\ s^{-1}}, \\ \alpha_{\nu\bar{\nu}} &= 4.7,\ \beta_{\nu\bar{\nu}} = 2.23,\ \gamma_{\nu\bar{\nu}} = 0.3, \\ \dot{m}_{\rm ign} &= 0.07\text{–}0.063 a_*,\ \dot{m}_{\rm trap} = 6.0\text{–}4.0 a_*^3, \end{aligned} \quad (11.22)$$

where $\dot{m}_{\rm ign}$ and $\dot{m}_{\rm trap}$ are the igniting and trapping accretion rates, respectively. For $m_* = 3$ and $\alpha = 0.1$, $\dot{m}_{\rm ign} = 0.07$ and $\dot{m}_{\rm trap} = 6.0$ for $a_* = 0$, and $\dot{m}_{\rm ign} = 0.01$ and $\dot{m}_{\rm trap} = 2.6$ for $a_* = 0.95$.

The Xue et al. (2013) approximation may be more precise in the low accretion rate regime (say, $\dot{m} < 0.5$) and the Zalamea and Beloborodov (2011) approximation may be more precise in the moderate accretion rate regime (say, $0.1 < \dot{m} < 2$). The Lei et al. (2017) formulae are consistent with both approximations in their respective sensitive regimes.

Baryon loading in the fireball is achieved through neutrino–nucleon weak interaction, either via charged-current interactions to strip protons or via neutral-current interactions

to strip neutrons. Neutrinos have a small probability of transferring momentum to protons/neutrons, giving rise to a neutrino-driven baryon wind. The baryon loading rate is given by (Qian and Woosley, 1996)

$$\dot{M}_\nu = (10^{-6} M_\odot \ \mathrm{s}^{-1})\, L_{\nu,52}^{5/3} \left\langle \left(\frac{\epsilon_\nu}{10\ \mathrm{MeV}} \right)^2 \right\rangle^{5/3} r_6^{5/3} \left(\frac{M}{3M_\odot} \right)^{-2} \left(\frac{h}{r} \right)^{-1}, \tag{11.23}$$

where h is the disk height, r is the radius from the BH in the equatorial plane, and $\langle \epsilon_\nu^2 \rangle = 13.8(kT)^2$ is the mean square neutrino energy. From the baryon loading rate, one may calculate the thermal-energy-to-mass ratio of the $\nu\bar{\nu}$-driven fireball, i.e.

$$\eta = \frac{L_{\nu\bar{\nu}}}{\dot{M}_\nu c^2}. \tag{11.24}$$

This is also the ultimate Lorentz factor of the fireball if $\eta \leq \eta_*$, where η_* is defined in Eq. (7.64).

11.2.3 Blandford–Znajek Mechanism

The Blandford–Znajek (BZ) mechanism (Blandford and Znajek, 1977) describes a mechanism for tapping the spin energy of a BH. The requirement is that there are open magnetic field lines that thread the BH and are connected to a remote astrophysical load, so that the field lines are twisted due to the BH spin and exert a torque on the BH to slow it down. Possible magnetic field configurations of a BH–torus system are shown in the right panel of Fig. 11.1 (van Putten, 2001).

The rotational energy of a BH with angular momentum J can be written as a fraction of the rest mass energy:

$$E_{\rm rot} = 1.8 \times 10^{54} f(a_*) \frac{M}{M_\odot}\ \mathrm{erg}, \tag{11.25}$$

where

$$f(a_*) = 1 - \sqrt{(1+q)/2}, \tag{11.26}$$

$q = \sqrt{1 - a_*^2}$, and $a_* = Jc/GM^2$ is the dimensionless BH spin parameter. For a maximally rotating BH ($a_* = 1$), one has $f(1) = 0.29$.

For a BH with magnetic field strength B near the horizon, the total Poynting flux power from the BZ process may be estimated as (Lee et al., 2000; Li, 2000; van Putten, 2001; Wang et al., 2002; McKinney, 2005)

$$L_{\rm BZ} = (1.7 \times 10^{50}\ \mathrm{erg\ s}^{-1})\, a_*^2 \left(\frac{M}{M_\odot} \right)^2 B_{15}^2 F(a_*), \tag{11.27}$$

where

$$F(a_*) = \left[\frac{1+q^2}{q^2} \right] \left[\left(q + \frac{1}{q} \right) \arctan q - 1 \right], \tag{11.28}$$

$$q = \frac{a_*}{1 + \sqrt{1 - a_*^2}}. \tag{11.29}$$

One has $F(0) = 2/3$ and $F(1) = \pi - 2$.

A major uncertainty in estimating the BZ power is the strength of the magnetic field. Analytically there are two approaches. The first is to assume that the magnetic pressure balances the ram pressure of the accretion flow, i.e.

$$\frac{B^2}{8\pi} \sim P_{\rm ram} = \rho c^2 \sim \frac{\dot{M}_{\rm acc} c}{4\pi r_{\rm H}^2}. \tag{11.30}$$

With this assumption, the magnetic power may be written as a function of mass accretion rate, i.e.

$$L_{\rm BZ} = 9.3 \times 10^{53} a_*^2 \dot{m} X(a_*) \ {\rm erg\ s^{-1}}, \tag{11.31}$$

where

$$X(a_*) = F(a_*)/(1+\sqrt{1-a_*^2})^2. \tag{11.32}$$

It is found that $X(0) = 1/6$ and $X(1) = \pi - 2$. The power does not depend on the mass of the BH in this treatment, but depends only on the dimensionless spin parameter a_* and the accretion rate $\dot{m} = \dot{M}/(M_\odot\ {\rm s^{-1}})$.

Alternatively, one may estimate the magnetic field strength through an equipartition argument, i.e. the comoving magnetic field density is a fraction of the internal energy density of the disk, i.e.

$$\frac{B^2}{8\pi} = \beta P_{\rm gas} \tag{11.33}$$

for the inner NDAF region (where gas pressure dominates). Taking $r = 6r_s$, the BZ power in this treatment gives

$$L_{\rm BZ} = (7.0 \times 10^{53}\ {\rm erg\ s^{-1}})\ \beta \alpha_{-2}^{-1.1} \dot{m} \left(\frac{M}{3M_\odot}\right)^{0.1}, \tag{11.34}$$

which has a shallow dependence on M.

The BZ mechanism has been confirmed from numerical simulations (e.g. McKinney, 2005; Nagataki, 2009; Tchekhovskoy et al., 2010, 2011). Tchekhovskoy et al. (2010) generally wrote

$$L_{\rm BZ} = \frac{\kappa}{4\pi c} \Omega_{\rm H}^2 \Phi_{\rm BH}^2 f(\Omega_{\rm H}), \tag{11.35}$$

where κ is a numerical constant whose value depends on the magnetic field geometry (0.053 for a split monopole geometry and 0.044 for a parabolic geometry), $\Omega_{\rm H} = ac/2r_{\rm H}$ is the angular frequency of the BH horizon, $r_{\rm H} = r_g(1+\sqrt{1-a_*^2})$ is the radius of horizon, $r_g = GM/c^2 = r_s/2$, $\Phi_{\rm BH} = (1/2)\int_\theta \int_\phi |B^r| dA_{\theta\phi}$, and $dA_{\theta\phi} = \sqrt{-g} d\theta d\phi$ is an area element in the θ–ϕ plane. The function $f(\Omega_{\rm H}) \approx 1 + 1.38(\Omega_{\rm H} r_g/c)^2 - 9.2(\Omega_{\rm H} r_g/c)^4$, which can usually be approximated as $\sim$1.

One can also define a BZ efficiency, which is defined as

$$\eta_{\rm BZ} \equiv \frac{\langle L_{\rm BZ}\rangle}{\langle \dot{M}\rangle c^2} \times 100\% = \frac{\kappa}{4\pi c} \left(\frac{\Omega r_g}{c}\right)^2 \left\langle \phi_{\rm BH}^2 \right\rangle f(\Omega) \times 100\%. \tag{11.36}$$

Tchekhovskoy et al. (2011) showed that for a magnetically arrested accretion disk, η can exceed 100% when a_* is close to unity. This suggests that the jet power indeed comes from the BH spin rather than accretion, as originally proposed (Blandford and Znajek, 1977).

The BZ mechanism launches a Poynting-flux-dominated jet. A small amount of baryons are expected to be entrained in the jet. A neutrino-driven baryon wind is still expected from the NDAF accretion disk. Protons, however, cannot penetrate into the magnetically dominated jet due to their small gyration radius. Baryon loading is achieved through neutrons that can penetrate the jet freely. Neutrons may decay to protons to load baryons in the jet, but neutron decay is a relatively slow process. Two efficient proton production mechanisms (Levinson and Eichler, 2003) include positron capture ($n+e^+ \rightarrow p+\bar{\nu}$, with positrons provided by $\nu\bar{\nu} \rightarrow e^-e^+$ annihilation) and inelastic nuclear collisions (e.g. $pn \rightarrow pp\pi^+ \ldots$, $pn \rightarrow pp\pi^- \ldots$, where a small amount of initially entrained protons are needed to trigger the cascade).

A detailed study of baryon loading in BZ jets and $\nu\bar{\nu}$-annihilation jets was carried out by Lei et al. (2013). The results are shown in the top panels of Fig. 11.2. Due to the strong magnetic fields, the BZ jets have much lower baryon loading (as displayed by the large $\Gamma_{\rm max} = \mu_0 = \eta(1+\sigma_0)$ value in the upper right plot; for the definition of μ_0 see Eq. (7.27) in Chapter 7) than $\nu\bar{\nu}$-annihilation jets (as displayed by the relatively small η value in the upper left plot). For both models, one generally has a positive dependence between the potential GRB Lorentz factor (μ_0 or η) and the luminosity, which gives a natural interpretation of several observed positive correlations, e.g. $E_{\gamma,\rm iso}-\Gamma_0$ (Liang et al., 2010), $L_{\gamma,\rm iso}-\Gamma_0$ (Lü et al., 2012), and $L_\gamma-\Gamma_0$ (Yi et al., 2017). On the other hand, $\nu\bar{\nu}$-annihilation jets are relatively "dirty" compared with the BZ jets (Lei et al., 2013).

Besides the BZ mechanism, there might be another mechanism for launching a relativistic jet from a magnetized central engine. If the accretion disk itself is highly magnetized, differential rotation of the disk would lead to accumulation of vorticity and energy within the disk, leading to eruption of magnetic blobs. If these blobs avoid heavy baryon loading during the eruption process, they may be launched with a relativistic speed and power a GRB (e.g. Yuan and Zhang, 2012). Numerical simulations are needed to reveal whether this process can avoid heavy baryon loading from the disk.

11.2.4 Comparison Between Two Jet Launching Mechanisms

If the central engine of a GRB is a BH–torus system, depending on the BH spin and the strength of the magnetic field threading the BH, the jet might be powered by either the $\nu\bar{\nu}$-annihilation mechanism or the BZ mechanism. A comparison between these two mechanisms has been investigated by several groups (e.g. Lei et al., 2013, 2017; Liu et al., 2015). The general conclusions can be summarized as follows:

- For high B and high a_*, the BZ power exceeds the $\nu\bar{\nu}$-annihilation power significantly;
- For high B and low a_*, the $\nu\bar{\nu}$-annihilation power exceeds the BZ power when $\dot{m}$ is large enough, but falls below the BZ power at the low-$\dot{m}$ regime;
- For low B and low a_*, the $\nu\bar{\nu}$-annihilation power dominates;
- In a large range of $\dot{m}$, the BZ power is roughly proportional to $\dot{m}$; the $\nu\bar{\nu}$-annihilation power, on the other hand, has a non-linear dependence on $\dot{m}$. In particular, it drops significantly in the low-$\dot{m}$ regime;

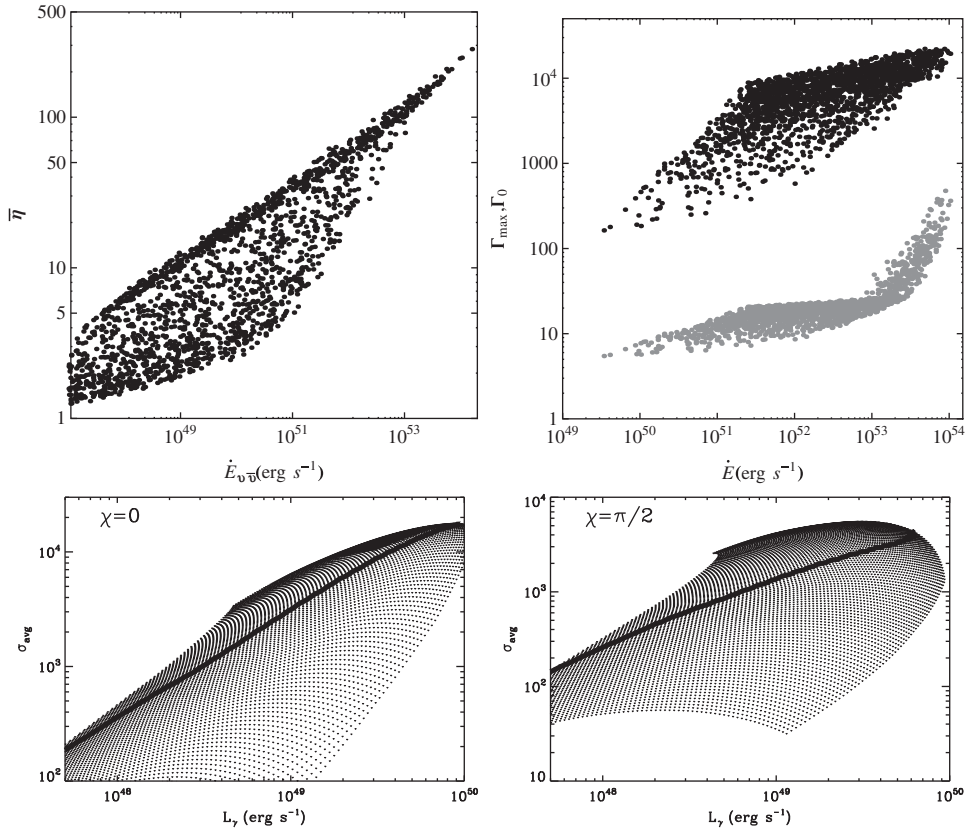

Figure 11.2 *Upper:* Simulated baryon loading of a black hole central engine. *Left:* the case of a $\nu\bar{\nu}$-annihilation-driven jet, with $\bar{\eta}$ denoting the average dimensionless entropy of the launched fireball; *Right:* the case of a Blandford–Znajek jet, with $\Gamma_{\rm max}$ (black) and Γ_0 (grey) denoting the maximum achievable Lorentz factor and the reached Lorentz factor after the rapid acceleration phase, respectively. For both cases, 2000 GRBs are simulated with BH mass, BH spin, accretion rate, and disk mass following certain distributions. From Lei et al. (2013). *Lower:* Simulated baryon loading of millisecond magnetar central engine. $\sigma_{\rm avg}$ is the average central engine magnetization parameter σ_0 during the GRB phase (defined as the phase when σ_0 is between 100 and 1000 by Metzger et al. 2011). GRBs are simulated with a range of distribution of B_p and P_0. Left and right panels are for the magnetic obliquity $\chi = 0$ and $\pi/2$, respectively. From Metzger et al. (2011).

- The BZ mechanism launches a Poynting-flux-dominated jet, whereas the $\nu\bar{\nu}$-annihilation mechanism launches a hot fireball;
- A BZ jet is much cleaner than a $\nu\bar{\nu}$-annihilation jet.

11.2.5 Applications to GRBs

The BH–torus engine has been the leading candidate for powering both massive star GRBs and compact star GRBs.

Massive Star GRBs

The erratic lightcurves of long GRBs may be explained within the BH–torus central engine model by invoking variation of the accretion rate at the central engine. For a massive star GRB, as the jet penetrates through the stellar envelope, additional modulation of the jet due to the jet–envelope interaction is possible. For a matter-dominated jet, a Kelvin–Helmholtz instability would develop near the edge of the jet, which induces variability in the jet even if the jet is continuous. Generally it would smear the short variability time scales from the central engine (Zhang et al., 2004b). However, if small-scale variabilities have a high power in the power density spectrum, they could still be observed along with the seconds-duration modulations introduced by the envelope (the slow variability component) (Morsony et al., 2010). The envelope also collimates the jet to a few degrees as it erupts from the envelope. The jet is surrounded by a hot cocoon. When erupting, the hot cocoon makes a wider, weaker, and less relativistic jet surrounding the central, narrow, stronger, and highly relativistic jet. One therefore naturally expects a two-component jet structure (Zhang et al., 2004b).

Some authors argued that evidence of BH spindown may be retrieved from the GRB lightcurve data (e.g. van Putten, 2012). This interpretation assumes that the lightcurves directly reflect the behavior at the central engine (which is relevant to the photosphere and internal shock models). On the other hand, if a GRB lightcurve reflects the emission history of a fluid unit as it streams in space (see Chapter 9 for a detailed discussion), the connection between the lightcurves and the BH–torus engine becomes irrelevant.

Since the magnetar central engine has a maximum energy budget defined by its initial spin energy (§11.3 below, Eq. (11.37)), direct support for a BH–torus central engine may be collected if a GRB has a total energy budget exceeding the magnetar maximum energy. A challenge of applying this argument is that one has to measure the jet opening angle using the afterglow data to derive the total energy of the system. Some very energetic GRBs indeed exceed the magnetar limit (e.g. Lü and Zhang, 2014), which calls for a BH–torus engine at least for some long GRBs.

Compact Star GRBs

A BH–torus engine is naturally expected for a BH–NS merger system. For a NS–NS merger, a BH–torus system is expected only if the NS equation of state is not too stiff and if the masses of the two pre-merger NSs are large enough. Indeed, the BH–torus engine gives a natural explanation of the duration of short GRBs based on the density argument (§10.2.2).

The challenge of the BH–torus engine is its difficulty in interpreting the internal plateau observed in a fraction of short GRBs (see §11.3 in detail). Attempts to interpret these plateaus within the BH model have been made (Kisaka and Ioka, 2015). However, the model curves are usually much too shallow as compared with the steep drop at the end of the plateau (Fig. 3 of Kisaka and Ioka, 2015). As discussed in §11.3 below, these observations may call for a supra-massive millisecond magnetar as the post-merger product of NS–NS mergers.

11.3 Millisecond Magnetars

11.3.1 General Considerations

Galactic Soft Gamma-ray Repeaters (SGRs) and Anomalous X-ray Pulsars (AXPs) have periods P in the range of 5–12 s, and period derivatives of the order $\dot{P} \sim 10^{-13}$ s s^{-1}. Assuming a standard magnetic dipole to account for the spindown behavior, one may estimate the polar cap surface magnetic field of the order $B_p \sim 10^{14}$–10^{15} G. These objects have been interpreted as "magnetars" (Thompson and Duncan, 1995, 1996).

Based on magnetic flux conservation, a normal star would have a surface dipolar magnetic field of order $\sim 10^{12}$ G when collapsing to a NS. To form a magnetar, one needs to invoke a star with anomalously high B to begin with. Alternatively, it was suggested (Duncan and Thompson, 1992; Thompson and Duncan, 1993) that strong magnetic fields may be generated in a nascent NS through an α–Ω dynamo mechanism. Such a nascent NS has a high magnetic Reynolds number and therefore is convective. The convection overturn time scale is $\sim$1 ms. If a newborn NS is diffentially rotating with a period shorter than the convection overturn time scale, the $\alpha - \Omega$ dynamo would develop efficiently, amplifying the magnetic field to $\gtrsim 10^{15}$ G.

It turns out that a millisecond magnetar has the right parameters to account for a long-duration GRB (Usov, 1992). The total spin energy of a millisecond magnetar with initial spin period $P_0 \sim 1$ ms is

$$E_{\rm rot} \simeq \frac{1}{2} I \Omega^2 \simeq (2.2 \times 10^{52}\ {\rm erg}) \left(\frac{M}{1.4 M_\odot}\right) R_6^2 P_{0,-3}^{-2}, \tag{11.37}$$

where

$$I \simeq \frac{2}{5} M R^2 = (1.1 \times 10^{45}\ {\rm g\ cm^2}) \left(\frac{M}{1.4 M_\odot}\right) R_6^2 \tag{11.38}$$

is the moment of inertia of a $1.4 M_\odot$ NS with radius $R_{\rm NS} \sim 10^6$ cm. This places an upper limit on the total energy budget of a GRB within the magnetar model, if the emission is powered by spin energy of the magnetar. Notice that in the early stage the NS may be differentially rotating, so that the simple estimate (11.37) may not apply. Nonetheless, Eq. (11.37) provides a good order-of-magnitude estimate of the total energy budget of a millisecond magnetar.

An additional energy source is possible if the newborn NS is also subject to accretion from a NDAF, so that the gravitational energy of the accreted materials is released. The total amount of the accretion energy is however limited, in order not to have the NS mass exceed the maximum value. Otherwise, the NS would collapse and one would have a BH–torus engine instead.

11.3.2 Energy Extraction Mechanisms

There are three energy extraction mechanisms for a millisecond magnetar engine.

Spindown

Radio pulsars power broad-band emission (from radio to γ-rays) through consuming their spin energy via magnetic dipole and relativistic wind spindown. So the most straightforward way of tapping the energy of a millisecond magnetar is through spindown.

Assuming rigid rotation (no differential rotation), the simplest neutron star spindown is controlled by magnetic dipole radiation. One can write the spindown law considering $\dot{E} = I\Omega\dot{\Omega} = -(B_p^2 R^6 \Omega^4)/(6c^3) \propto -\Omega^4$ (where $E_{\rm rot} = (I\Omega^2/2)$),

$$\Omega = \frac{\Omega_0}{\left(1 + t/t_{0,\rm em}\right)^{1/2}}, \tag{11.39}$$

and the evolution of the spindown luminosity (Shapiro and Teukolsky, 1983; Usov, 1992; Zhang and Mészáros, 2001a)

$$L(t) = \frac{L_0}{(1 + t/t_{0,\rm em})^2} \simeq \begin{cases} L_0, & t \ll t_{0,\rm em}, \\ L_0 (t/t_{0,\rm em})^{-2}, & t \gg t_{0,\rm em}, \end{cases} \tag{11.40}$$

where

$$t_{0,\rm em} = \frac{3c^3 I}{B_p^2 R^6 \Omega_0^2} \simeq (2.1 \times 10^3 \ {\rm s})\, I_{45} B_{p,15}^{-2} P_{0,-3}^2 R_6^{-6} \tag{11.41}$$

is the characteristic spindown time scale and

$$L_0 = \frac{I\Omega_0^2}{2t_{0,\rm em}} = \frac{B_p^2 R^6 \Omega^4}{6c^3} \simeq (1.0 \times 10^{49} \ {\rm erg \ s^{-1}}) B_{p,15}^2 P_{0,-3}^{-4} R_6^6 \tag{11.42}$$

is the initial spindown luminosity. Here $\Omega_0 = 2\pi/P_0$ and P_0 are the initial angular velocity and initial spin period of the magnetar at birth, B_p is the surface magnetic field strength at the polar cap region, and I and R are the moment of inertia and radius of the NS, respectively. The total spin energy, and hence the total energy budget, of the magnetar depends on I and P_0 only. The magnetic field strength B_p defines the luminosity and the duration that consumes the total energy. To power a GRB, both "millisecond", which defines a large energy budget, and "magnetar", which defines a high luminosity, are needed. If the birth period P_0 of most central engine magnetars is close to the breakup value, then the total energy budget of GRBs would be roughly a constant.

Notice that for magnetic dipole spindown in vacuum, one needs to introduce a factor $\sin\alpha$, where α is the inclination angle between the spin and magnetic axes of the magnetar. However, a magnetar is not surrounded by a vacuum, but rather carries a magnetosphere, inside which a Goldreich–Julian plasma wind streams out of the magnetosphere and carries away angular momentum. It turns out that this component gives a spindown effect comparable to the dipole component, but with a dependence on $\cos\alpha$ (Harding et al., 1999; Xu and Qiao, 2001; Contopoulos and Spitkovsky, 2006). Combining the two effects, the dependence on α is weak, so we do not introduce the α dependence in Eq. (11.42).

Deviation from the simple dipole spindown law (11.40) is possible, especially in the early phase when the neutron star is newly born. There are two possible deviations.

First, a newborn neutron star would launch a strong neutrino-driven wind, or a magnetically driven wind due to the differential rotation of the neutron star, both of which

are different from the Goldreich–Julian wind of a pulsar. These winds carry away angular momentum of the star with different spindown laws (e.g. Metzger et al., 2011; Siegel et al., 2014). Based on numerical simulations, Siegel et al. (2014) derived a semi-analytical spindown formula for a magnetically driven wind:

$$L_{\rm em} \simeq (10^{48}\ {\rm erg\ s^{-1}})\, B_{p,14}^2 R_6^3 P_{-4}^{-1}, \tag{11.43}$$

which may be valid for a limited duration of time before the dipole spindown kicks in (otherwise the angular velocity would turn negative).

Second, the newborn millisecond magnetar may lose significant spin energy via gravitational wave (GW) radiation. If GW spindown dominates over the magnetic dipole spindown (possible if the neutron star ellipticity ϵ is large but the magnetic field strength B_p is relatively low), one has $\dot{E} = I\Omega\dot{\Omega} = -(32GI^2\epsilon^2\Omega^6)/(5c^5) \propto -\Omega^6$, so that the spindown law becomes (Shapiro and Teukolsky, 1983; Usov, 1992; Zhang and Mészáros, 2001a)

$$\Omega = \frac{\Omega_0}{\left(1 + t/t_{0,\rm GW}\right)^{1/4}}, \tag{11.44}$$

where

$$t_{0,\rm GW} = \frac{5c^5}{128GI\epsilon^2\Omega_0^4} \simeq (9.1\times 10^3\ {\rm s})\, I_{45}^{-1} P_{0,-3}^4 \epsilon_{-3}^{-2}. \tag{11.45}$$

More generally, both the dipole and GW components may play a role in spindown, but the luminosity in the EM channel is defined by the dipole component only, since the GW energy escapes the system. The dynamics and spindown luminosity evolution of the millisecond magnetar can be described as follows (Gao et al., 2016; Lasky and Glampedakis, 2016; Sun et al., 2017).

The spindown rate is

$$\dot{E} = I\Omega\dot{\Omega} = -\frac{B_p^2 R^6 \Omega^4}{6c^3} - \frac{32GI^2\epsilon^2\Omega^6}{5c^5}. \tag{11.46}$$

The relation between t and Ω is complicated, but can be derived analytically if ϵ and B_p are constants. The solution reads (Gao et al., 2016)

$$t = \frac{a}{2b^2}\ln\left[\left(\frac{a\Omega_0^2 + b}{a\Omega^2 + b}\right)\frac{\Omega^2}{\Omega_0^2}\right] + \frac{\Omega_0^2 - \Omega^2}{2b\Omega^2\Omega_0^2}, \tag{11.47}$$

where

$$a = \frac{32GI\epsilon^2}{5c^5}, \tag{11.48}$$

$$b = \frac{B_p^2 R^6}{6c^3 I}. \tag{11.49}$$

The spindown luminosity in the EM channel still reads

$$L_{\rm sd}(t) = \frac{B_p^2 R^6 \Omega^4(t)}{6c^3}. \tag{11.50}$$

Considering the competition between $t_{0,\rm em}$ and $t_{0,\rm GW}$, and $\Omega \propto t^{-1/2}$ and $\propto t^{-1/4}$, respectively, in the dipole and GW spindown dominated regimes, the final luminosity scaling laws can be characterized in two regimes:

$$L_{\rm sd} \propto \begin{cases} t^0, & t < t_{\rm sd} = t_{0,\rm GW}, \\ t^{-1}, & t_{\rm sd} = t_{0,\rm GW} < t < t_{0,\rm em}, \\ t^{-2}, & t > t_{0,\rm em} \end{cases} \tag{11.51}$$

for $t_{0,\rm GW} < t_{0,\rm em}$, and

$$L_{\rm sd} \propto \begin{cases} t^0, & t < t_{\rm sd} = t_{0,\rm em}, \\ t^{-2}, & t > t_{\rm sd} = t_{0,\rm em}, \end{cases} \tag{11.52}$$

for $t_{0,\rm em} < t_{0,\rm GW}$. Here we have defined the spindown time scale as

$$t_{\rm sd} = \min(t_{0,\rm em}, t_{0,\rm GW}). \tag{11.53}$$

There are several possibilities for deforming a newborn NS. The bar-mode instability (Andersson, 2003; Corsi and Mészáros, 2009; Lasky and Glampedakis, 2016) may apply to the early phase of the post-merger remnant's life and would be suppressed once differential rotation is quenched. The maximum achievable ellipticity for the bar mode is $\epsilon_f \sim 10^{-3}$. The inertial quadrupolar r-mode (Andersson and Kokkotas, 2001) would induce gravitational wave radiation, but the characteristic time scale is long ($\sim 5 \times 10^9$ s), which does not play a significant role (Lasky and Glampedakis, 2016). The most important contribution to the ellipticity for a newborn millisecond magnetar may be magnetically induced. For a differentially rotating protomagnetar, magnetic fields would be wound up and stored in the toroidal form. This would naturally distort the NS to be non-spherical (e.g. Cutler, 2002; Haskell et al., 2008). The condition for significant gravitational wave radiation is that the magnetic axis misaligns with the rotation axis. Even though initially the magnetic axis is likley aligned with the rotation axis since toroidal field is built up through differential rotation, the aligned geometry is not stable since the orthogonal configuration lowers the energy of the system. As a result, a millisecond magnetar may apply the magnetically induced ϵ to give strong GW radiation through a "spin-flip" instability, which is possible when the internal temperature is high enough (Lasky and Glampedakis, 2016). In general, the ellipticity may be written in the form of (Cutler, 2002; Haskell et al., 2008)

$$\epsilon \simeq 10^{-6} C \left(\frac{R}{10\,{\rm km}}\right)^4 \left(\frac{M}{1.4 M_\odot}\right)^{-2} \left(\frac{\bar{B}}{10^{15}\,{\rm G}}\right)^2, \tag{11.54}$$

where $\bar{B}$ is the volume average B field, and $C \gtrsim 1$ is a model-dependent constant, which may be as large as several hundreds, so that ϵ_B can be as large as 10^{-4}–10^{-3} (Haskell et al., 2008).

The connection between the spindown luminosity with GRB observations is not straightforward. Usov (1992, 1994) proposed that the GRB prompt emission itself is powered by spindown. One drawback of this interpretation is that the GRB lightcurves are often erratic, while the spindown luminosity is continuous without significant fluctuation. The spindown power, on the other hand, is very successful in explaining energy injection in the afterglow phase (Dai and Lu, 1998c,a; Zhang and Mészáros, 2001a), especially the X-ray shallow

decay phase or external plateau as observed in many long GRBs (Zhang et al., 2006; Nousek et al., 2006). These signatures, however, do not necessarily require a long-lasting central engine, since continuous energy injection is possible in a blastwave if the ejected mass has a distribution in Lorentz factor even if it is ejected within a relatively short period of time (Rees and Mészáros, 1998; Sari and Esin, 2001; Uhm and Beloborodov, 2007; Genet et al., 2007; Uhm et al., 2012, see §8.3.3 for details). A smoking gun signature of the existence of a millisecond magnetar at the GRB central engine was revealed when the "internal plateaus" following some long (Troja et al., 2007; Lü and Zhang, 2014) and short (Rowlinson et al., 2010, 2013; Lü et al., 2015) GRBs were discovered. Such a plateau cannot be interpreted via an external shock since the decay slope at the end of plateau is much steeper than allowed from the external shock model due to the curvature effect. It demands internal emission of a central engine lasting for hundreds or even over 10^4 seconds with essentially constant luminosity. A millisecond magnetar is the natural candidate for explaining such a signature. The very steep decay at the end of the plateau is also very intriguing. It suggests a sudden cessation of the central engine wind emission. A natural interpretation within the magnetar model is that the neutron star is "supra-massive" at birth, and collapses to a black hole after it spins down significantly (Zhang, 2014).

Magnetic Bubble Eruption Due to Differential Rotation

The second energy extraction mechanism from a millisecond neutron star is to extract the differential rotational energy of the neutron star through erupting magnetic bubbles due to the wind-up of the seed field (Kluźniak and Ruderman, 1998; Ruderman et al., 2000). Such a process naturally gives rise to an erratic central engine, which is relevant for interpreting GRB prompt emission. The model was also generalized to the models of X-ray flares following short GRBs (Dai et al., 2006). The seed dipole field of these neutron stars is required to be in the range of normal pulsars (10^{10}–10^{12} G). However, the toroidal field after amplification due to differential rotation may reach $\sim 10^{17}$ G, so that they can be considered as millisecond magnetars as well.

The physical picture of this mechanism has been outlined in Kluźniak and Ruderman (1998), Ruderman et al. (2000), and Dai et al. (2006). Differential rotation would lead to wind-up of the seed poloidal field B_r. Consider a toy model of differential rotation between two layers only, a core with a rigid rotation frequency Ω_c and a surrounding shell with a rigid rotation frequency Ω_s. The core rotates faster by a rotation frequency differential

$$\Delta\Omega = \Omega_c - \Omega_s. \tag{11.55}$$

Suppose the neutron star initially has a poloidal field strength B_r. Due to differential rotation, the toroidal magnetic field strength increases at the rate

$$\frac{dB_\phi}{dt} = (\Delta\Omega)B_r. \tag{11.56}$$

The steady growth of B_ϕ forms a magnetically confined "toroid" inside the neutron star, which encloses some neutron star matter. The toroid floats up from the deep interior as B_ϕ is amplified to a large enough value when magnetic buoyancy can overcome the radial

stratification in the neutron star composition. The toroid then erupts from the surface, making a magnetic bubble that powers the GRB. Kluźniak and Ruderman (1998) estimated that the critical magnetic field strength for magnetic buoyancy to dominate is $B_b \sim 10^{17}$ G.

If during the magnetic amplification phase $\Delta\Omega$ is roughly constant (which is usually true since the eruption time scale is usually much shorter than the $\Delta\Omega$ evolution time scale, Kluźniak and Ruderman (1998); Dai et al. (2006)), then the time scale for making a magnetic bubble from the neutron star can be estimated as

$$t_b = \frac{2\pi}{(\Delta\Omega)}\frac{B_b}{B_r} \simeq 20 \text{ s } B_{b,17}B_{r,12}^{-1}(\Delta\Omega)_4. \tag{11.57}$$

Such a time scale agrees with the quiescent time between active episodes in some GRBs. The total energy of each eruption may be estimated as

$$E_b = \frac{B_b^2}{8\pi}V_b \simeq (1.6\times 10^{51} \text{ erg})\frac{V_b}{V_*}, \tag{11.58}$$

where $V_* = (4\pi/3)R^3$ is the volume of the neutron star. This is of the order of the GRB pulse energy, if the bubble volume V_b is not much smaller than V_*. One may also estimate baryon loading rate and bulk Lorentz factor of the bubble from this mechanism based on the condition of buoyancy overcoming gravity in the toroid. This gives (Kluźniak and Ruderman, 1998)

$$\frac{M}{M_*} \simeq 2\times 10^{-5}\frac{V_b}{V_*}[(\Delta\Omega)_4 B_{r,12}]^{2/3}, \tag{11.59}$$

$$\Gamma = \frac{E_b}{Mc^2} \geq 100[(\Delta\Omega)_4 B_{r,12}]^{-2/3}\left(\frac{M}{M_\odot}\right)^{-1}. \tag{11.60}$$

Accretion

A third possible mechanism for powering a GRB with a neutron star engine is through accretion. Similar to a black hole engine, a newborn neutron star may be surrounded by a hyper-accreting disk/torus, so that the NDAF mechanism for the BH engine may be applied to the NS engine. There are, however, noticeable differences.

- Zhang and Dai (2008) showed that a NS disk cools more efficiently and produces a much higher neutrino luminosity than a BH disk. This is due to the existence of an inner surface boundary of the NS which makes the disk denser and hotter in the inner region. For the same reason and subject to parameters, the $\nu\bar{\nu}$-annihilation luminosity of the NS disk is usually also higher than that of the BH disk (Zhang and Dai, 2009). For a magnetar engine, the existence of the strong magnetic field truncates the disk at the Alfvén radius, giving rise to a higher temperature and more concentrated neutrino emission in a ring-like belt region. The neutrino annihilation rate is further enhanced (Zhang and Dai, 2010).
- The newborn NS is also very hot, and therefore launches a dirty neutrino-driven wind to prevent a relativistic outflow (e.g. Dessart et al., 2009; Metzger et al., 2011). A relativistic jet may be launched after a certain waiting time (Zhang and Dai, 2009).

- The total accreted mass should be small enough to prevent collapse of the NS. This places limits on the accretion rate and duration of the burst. For example, if a newborn NS has a mass $\sim 1.4 M_\odot$ and if the maximum NS mass is $\sim 2.4 M_\odot$, the duration should be shorter than 10 s if the accretion rate is $0.1 M_\odot\ \mathrm{s}^{-1}$.

11.3.3 Applications to GRBs

For a millisecond magnetar engine, it may be possible that one or more of the above-mentioned mechanisms operate within the same GRB. In the following we discuss evidence and applications of the millisecond magnetar engine in two types of GRBs.

Massive Star GRBs

In the massive star core-collapse scenario, a millisecond magnetar may be born upon core collapse. Shortly after the launch of the supernova shock, the protomagnetar launches a baryon-loaded neutrino-driven wind. As the magnetar cools, the wind becomes progressively Poynting flux dominated (σ_0 increases with time). The evolution of the wind luminosity (which is also the magnetar spindown rate) and σ_0 has been studied by Metzger et al. (2011). As shown in Fig. 11.3, early on the magnetar wind is too dirty to launch a relativistic jet. After about 10 s, σ_0 rises to above 100, which suggests that the terminating Lorentz factor may be of this order. Metzger et al. (2011) suggested that this is the phase of GRB prompt emission. After tens of seconds, the neutron star becomes "neutrino-thin", baryon loading due to the neutrino-driven wind stops suddenly, and σ_0 becomes extremely high ($\sim 10^9$). Metzger et al. (2011) argued that this corresponds to the cessation of the GRB, since high-σ_0 outflow may not be able to dissipate the magnetic energy efficiently. Later, the magnetar wind injects energy into the blastwave to power the X-ray plateau in the form of a Poynting-flux-dominated outflow. The total baryon loading vs. jet energy can be calculated from the magnetar model. As shown in the lower panels of Fig. 11.2, a rough

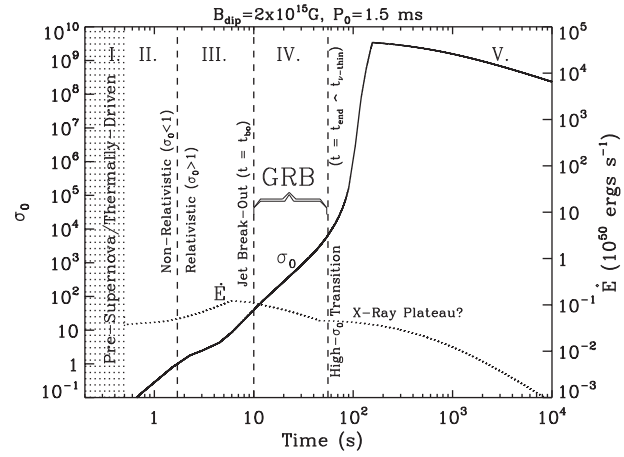

Figure 11.3 Evolution of σ_0 and E of a millisecond-magnetar-powered outflow. From Metzger et al. (2011).

positive correlation between $\sigma_{\rm avg}$ and L_γ is also obtained (similar to the BH model) in the magnetar central engine model.

The magnetar wind is essentially isotropic. However, numerical simulations show that the outflow can be collimated in the direction of the spin axis by the envelope of the massive star to a narrow jet with opening angle 5–10$^{\rm o}$ (Bucciantini et al., 2009). Since the magnetic field configuration of a pulsar-like rotator is striped-wind-like, the jet is wrapped by magnetic field lines with alternating polarity. Reconnection may be facilitated through kink instabilities. Metzger et al. (2011) suggested that magnetic dissipation may occur mostly below the photosphere, so that the observed emission is from a dissipative photosphere.

Even though this picture gives a plausible scenario for GRB emission for long GRBs, there are many uncertainties. First, there might be an accretion phase that would complicate the picture. Second, there might not be a direct connection between σ_0 and the GRB prompt emission Γ. According to the dynamical evolution of a Poynting-flux-dominated flow (§7.4.2) the jet is quickly accelerated to $\Gamma \sim (1 + \sigma_0)^{1/3}$ at the sonic point. Beyond this radius, the acceleration is very delicate. If such acceleration is insignificant, the Lorentz factor in the "GRB" phase defined by Metzger et al. (2011) may be too low to power a GRB, and the later high-σ phase during the neutrino-thin phase may instead allow a GRB with a Γ of several hundreds. Another complication is that a high-σ_0 outflow is subject to strong inverse Compton drag by the hot photons trapped in the envelope, so that the terminal Γ may be much smaller (Ceccobello and Kumar, 2015). Finally, erratic behavior at the central engine (e.g. the magnetic bubble mechanism, Kluźniak and Ruderman 1998) would give rise to magnetic field configurations other than a striped wind. It is possible that magnetic bubbles collide and trigger ICMART events in an optically thin region far above the photosphere radius. The internal plateau in some GRBs lasts up to 10^4 s (Troja et al., 2007), suggesting that a Poynting-flux-dominated jet may indeed dissipate internally and power bright emission in X-rays.

From the prompt emission data there is essentially no smoking gun evidence for a magnetar central engine. The strongest support for the magnetar engine is from X-ray afterglow data that show the existence of an external plateau in a large sample of GRBs and, more importantly, an internal plateau in a small fraction of long GRBs and a good fraction of short GRBs. A systematic confrontation of the magnetar model with the long GRB afterglow and prompt emission data was carried out by Lü and Zhang (2014). They classified GRBs into several samples based on the likelihood that they harbor a magnetar central engine. The "Gold" sample includes GRBs that possess an internal plateau; the "Silver" sample includes GRBs with an external plateau whose energy injection index q (convention $L(t) \propto t^{-q}$) is consistent with being 0, a prediction of the magnetar model; the "Aluminum" sample includes GRBs with an external plateau whose q index is not 0; and those GRBs without a plateau or a shallow decay phase are included in the "Non-magnetar" sample. Based on the plateau duration and luminosity, they estimated P_0 and B_p of the magnetars. The results suggest that the derived magnetar parameters for the Gold and Silver samples are indeed consistent with theoretical expectations of the magnetar model, and the total energy budget of these GRBs indeed does not exceed the maximum energy of a magnetar (Eq. (11.37)). The Non-magnetar sample GRBs, on the other hand, are on average more energetic than the magnetar samples. Some Non-magnetars even have jet-corrected energy

exceeding Eq. (11.37). Even if the results do not prove, they nonetheless suggest, that both types of central engines (BH and magnetar) might be operating in GRBs.

Compact Star GRBs

A millisecond magnetar central engine is relevant only for NS–NS mergers, since if one of the members in the merger system is a BH, the central engine must be a BH–torus system.

For NS–NS mergers there are in principle four outcomes, depending on the total mass that enters the merger system and the *equation of state (EoS)* of neutron stars. Given a stiff enough EoS (which suggests a relatively large maximum NS mass) the following four outcomes, with the sequence of reducing total mass in the merger remnant, are in principle possible:

- *Prompt BH*: If the total mass is large enough, the merger will leave behind a BH immediately after the merger;
- *Hyper-massive NS*: For a lower total mass, the merger product will be a temporary NS supported by differential rotation which collapses into a BH after a short period of time, e.g. ∼100 ms;
- *Supra-massive NS*: For an even lower total mass, the merger product may not collapse even after the NS enters the rigid rotation phase. The total mass is greater than the maximum mass for a non-rotating NS, but can survive for an extended period of time as long as the NS is rapidly spinning. The NS collapses to a BH after it loses significant angular momentum due to dipole and gravitational wave radiation;
- *Stable NS*: If the total mass is smaller than the maximum mass of a non-rotating NS, then the merger remnant will be a stable NS lasting forever.

Since the NS EoS is not well constrained from observational data, the maximum mass of a non-rotating NS, i.e. $M_{\rm TOV}$, is unknown. Nonetheless, the existence of a $\sim 2M_\odot$ NS (Demorest et al., 2010) suggests that $M_{\rm TOV} > 2M_\odot$. This rules out all soft EoSs with $M_{\rm TOV} < 2M_\odot$ and suggests a population of stiff NS EoSs. Observations of Galactic NS–NS binaries, on the other hand, suggest that the sum of the gravitational masses of the two NSs is in the range of 2.5–2.7$M_\odot$ (Kiziltan et al., 2013; Martinez et al., 2015). The baryon mass M_b of a NS may be inferred from the gravitational mass M_g as (Burrows and Lattimer, 1986; Timmes et al., 1996)

$$\frac{M_b}{M_\odot} = \frac{M_g}{M_\odot} + 0.075 \left(\frac{M_g}{M_\odot}\right)^2 \tag{11.61}$$

for each NS. The total baryon mass in the merger product can be calculated as

$$M_{{\rm rem},b} = M_{1,b} + M_{2,b} - \Delta M_{\rm ej}, \tag{11.62}$$

where $\Delta M_{\rm ej} \sim 10^{-3}$–$10^{-1} M_\odot$ is the mass of the ejecta launched during the merger (Freiburghaus et al., 1999; Rezzolla et al., 2010; Hotokezaka et al., 2013; Rosswog et al., 2013). The final gravitational mass of the remnant $M_{{\rm rem},g}$ can be calculated using Eq. (11.61) according to $M_{{\rm rem},b}$. This mass can be compared against the maximum NS mass sustained by spin. For a supra-massive NS (or even a quark star (QS)), the maximum

Table 11.1 The parameters of supra-massive NSs or QSs for some EoSs. Here $M_{\rm TOV}$, $R_{\rm eq}$ are the static gravitational maximum mass and the corresponding equatorial radius, respectively; α, β are the fitting parameters for $M_{\rm max}$ in Eq. (11.63). From Lasky et al. (2014); Lü et al. (2015) (first five) and Li et al. (2016a) (the rest; see Fig. 11.4 lower panel for relevant mass–radius relations)

	EoS	$M_{\rm TOV}$ ($M_\odot$)	$R_{\rm eq}$ (km)	α	β
	SLy	2.05	9.99	1.60	−2.75
NS	APR	2.20	10.0	0.303	−2.95
	GM1	2.37	12.05	1.58	−2.84
	AB-N	2.67	12.9	0.112	−3.22
	AB-L	2.71	13.7	2.92	−2.82
	BCPM	1.98	9.941	0.03859	−2.651
NS	BSk20	2.17	10.17	0.03587	−2.675
	BSk21	2.28	11.08	0.04868	−2.746
	Shen	2.18	12.40	0.07657	−2.738
	CIDDM	2.09	12.43	0.16146	−4.932
QS	CDDM1	2.21	13.99	0.39154	−4.999
	CDDM2	2.45	15.76	0.74477	−5.175

NS mass allowed by the maximum rigid rotation (near breakup or with the Keplerian frequency) is about 20% more than $M_{\rm TOV}$ (Lasky et al., 2014; Breu and Rezzolla, 2016; Li et al., 2016a). For a range of EoS, the maximum mass can be parameterized as

$$M_{\rm max} = M_{\rm TOV}\left[1+\alpha\left(\frac{P}{\rm ms}\right)^{\beta}\right], \tag{11.63}$$

with α and β taking different values for different EoSs (Table 11.1). Comparing $M_{{\rm rem},g}$ distributions from Galactic NS–NS binary systems with $M_{\rm max}$ of some stiff NS/QS EoSs, one can see that indeed a good fraction of NS–NS mergers can leave behind supra-massive and even stable NSs (Lasky et al., 2014; Lü et al., 2015; Gao et al., 2016; Li et al., 2016a).

Equating $M_{{\rm rem},g}$ and $M_{\rm max}$, one can solve for the characteristic spin period of the NS when it collapses to a BH, i.e.

$$P_c = \left(\frac{M_{{\rm rem},g}-M_{\rm TOV}}{\alpha M_{\rm TOV}}\right)^{1/\beta}. \tag{11.64}$$

For a NS with an initial spin period P_0 (which is essentially the shortest possible one for NS–NS mergers, since both NSs were in Keplerian motion before the merger), one has the following situations:

- If $P_c < P_0$, the merger makes a prompt BH or a hyper-massive NS that shortly collapses to a BH;
- If $P_c > P_0$, the merger makes a supra-massive NS, which collapses after it spins down to period P_c;
- If $P_c < 0$, the merger product will never collapse, so the merger makes a stable NS.

For the standard dipole and GW spindown of a rigidly rotating NS and assuming constant B_p and ϵ during the spindown phase, based on Eq. (11.47), one can derive the collapse time of a supra-massive NS, i.e. (Gao et al., 2016)

$$t_c = \frac{a}{2b^2} \ln \left[\left(\frac{a\Omega_0^2 + b}{a\Omega_c^2 + b} \right) \frac{\Omega_c^2}{\Omega_0^2} \right] + \frac{\Omega_0^2 - \Omega_c^2}{2b\Omega_0^2\Omega_c^2}, \tag{11.65}$$

where a and b are defined in Eqs. (11.48) and (11.49). If GW spindown is negligible, the collapse time due to dipole spindown can be estimated as (Lasky et al., 2014)

$$t_c = \frac{3c^3 I}{4\pi^2 B_p^2 R^6} \left[\left(\frac{M_{\rm rem,g} - M_{\rm TOV}}{\alpha M_{\rm TOV}} \right)^{2/\beta} - P_0^2 \right]. \tag{11.66}$$

Based on the assumption that the end of the internal plateau (at t_b) following a short GRB (Rowlinson et al., 2010, 2013; Lü et al., 2015) marks the time when the supra-massive NS/QS engine collapses, one can give constraints on the possible EoSs of NSs and QSs.

Since the density covers many orders of magnitude from the core to the surface of a NS, a global solution of the NS depends on a proper description of the EoS of neutron-rich matter in different density ranges. There are a lot of uncertainties, especially in the core region where density exceeds the nuclear density. As a result, many models exist to describe NS mass as a function of radius (Fig. 11.4). The existence of massive NSs rules out all the soft EoSs that cannot produce $2M_\odot$ NSs. However, there are many stiff EoSs with maximum mass greater than $2M_\odot$.

Lasky et al. (2014) worked on eight individual short GRBs and considered five different EoSs and suggested that one of them (GM1, with $M_{\rm TOV} = 2.37M_\odot$, $R = 12.05$ km) matches the data the best in terms of the observed collapsing time $t_c = t_b$. The same EoS was found to be consistent with more short GRBs with an internal plateau (Lü et al., 2015). Gao et al. (2016) worked on the entire short GRB sample and confronted the five EoSs with the following observational constraints: (1) the fraction of short GRBs with an internal plateau (observationally about 22%; in view that some short GRBs may originate from other progenitors such as BH–NS mergers, this fraction is the minimum fraction of supra-massive NSs); (2) the distribution of the collapse time t_c; and (3) the distribution of the plateau luminosity. Through Monte Carlo simulations, they found that the other four EoSs (SLy, APR, AB-N, AB-L) considered by Lasky et al. (2014) cannot satisfy the constraint of the supra-massive NS fraction. The GM1 EoS gives 40%, 30%, and 30% BHs, supra-massive, and stable NSs, respectively. In order to reproduce the short GRB data one requires $P_0 \sim 1$ ms, $B_p \sim 10^{15}$, and $\epsilon \sim 0.004$–0.007. The results suggest that these merger-produced neutron stars are indeed millisecond magnetars. The relatively large value of ϵ is close to the maximum value allowed by various distortion mechanisms for a newborn magnetar, but is theoretically allowed. Of course, GM1 is not the only EoS that can satisfy the observational constraints. Li et al. (2016a) investigated four more NS EoSs and three more QS EoSs, and found that all except one can satisfy the minimum fraction 22% for supra-massive NSs/QSs. The conclusions regarding the P_0, B_p, and ϵ distributions are similar to the case for GM1. However, QS EoSs produce a narrower t_c distribution, which matches the observations better (§11.4 for more discussions). Piro et al.

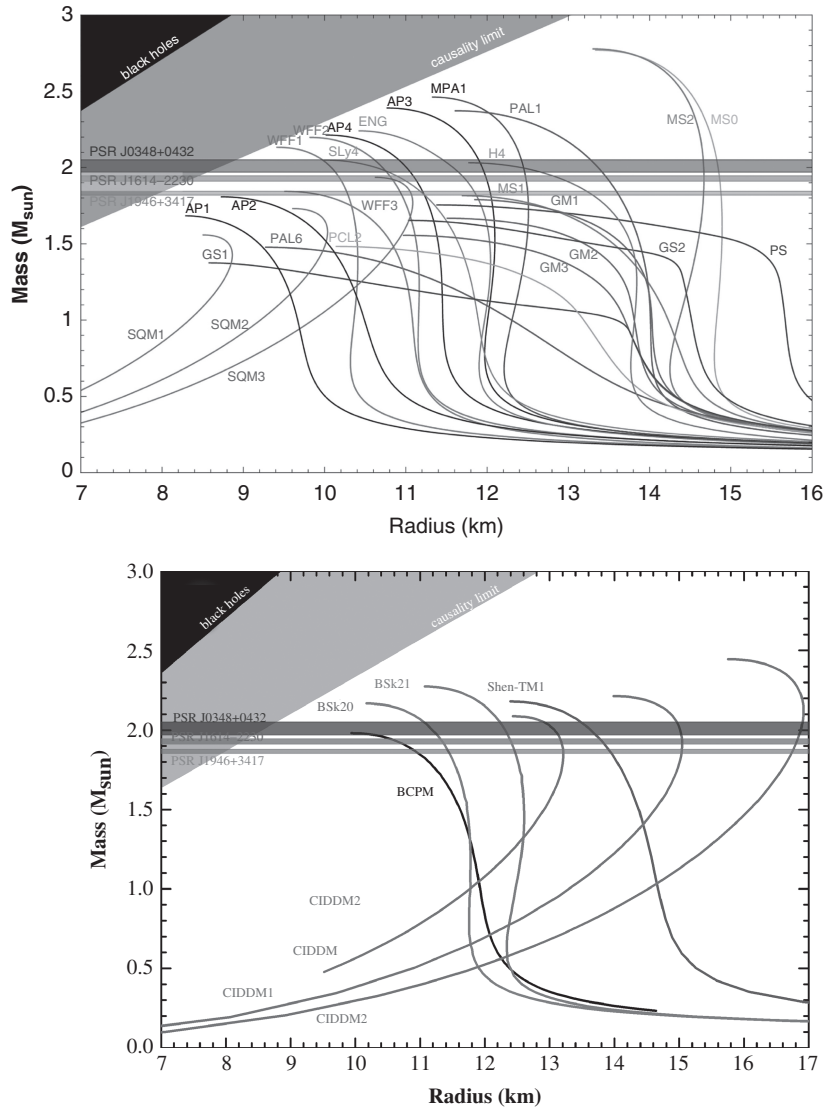

Figure 11.4 Various NS/QS EoSs and available constraints. *Upper:* Figure courtesy Norbert Wex, data from Lattimer and Prakash (2001). From http://www3.mpifr-bonn.mpg.de/staff/pfreire/NS_masses.html. *Lower:* EoSs studied in Li et al. (2016a). Figure courtesy Ang Li.

(2017) investigated a list of realistic NS EoSs and reached a similar conclusion that a large fraction of post-merger products are supra-massive. They also ruled out some of the stiffest EoSs based on the short GRB data.

One constraint on the magnetar central engine model for short GRBs is its energetics. A straightforward expectation is that the newborn magnetar is likely spinning at the maximum energy, so that a total energy close to the maximum, a few 10^{52} erg (Eq. (11.37)), is expected. Since the millisecond magnetar energy is likely deposited quasi-isotropically in

all directions, the brightness of the late-time radio afterglow would significantly constrain this energy. Upper limits on the radio afterglow fluxes in a few short GRBs (Horesh et al., 2016; Fong et al., 2016) suggest that the total energy of some GRBs cannot exceed a few times 10^{51} erg. This seems to be inconsistent with the magnetar model for short GRBs. However, according to the global fit to the short GRB X-ray plateau data, even though magnetars are typically born with $P_0 \sim 1$ ms, not all the spin energy (Eq. (11.37)) goes to the electromagnetic channel. In fact, the energy will likely be distributed in three channels: (1) the EM channel; (2) the GW channel (which includes the GW emission during the merger phase (Radice et al. 2016) and the GW emission during the supra-massive NS spindown phase (Fan et al. 2013c; Gao et al. 2016; Lasky and Glampedakis 2016); and (3) the BH channel (the magnetar collapses before spinning down, so that most of the spin energy of the magnetar falls into the BH). Matching the short GRB X-ray plateau data, Gao et al. (2016) found that the total energy in the EM channel can range from a few times 10^{50} erg to the maximum energy (Eq. (11.37) with $M = M_{\rm rem,g}$). Also considering the possibility of low values of shock microphysics parameters as inferred from GRB afterglow observations, the radio afterglow upper limits (Horesh et al., 2016; Fong et al., 2016) may not necessarily pose severe constraints on the validity of the magnetar model.

One major issue of the magnetar model for short GRBs is the origin of the short-duration prompt emission itself. Within the BH central engine model, the duration of the short GRB is controlled by the accretion time scale and possibly the duration of the hyper-massive NS (e.g. Rosswog et al., 2003). Within the magnetar central engine model, however, since the magnetar lasts much longer than 2 s, another physical time scale is needed to account for the duration of the prompt emission. One may still appeal to the accretion time scale to define the short GRB emission duration (e.g. Metzger et al., 2008). However, a nascent NS is likely so hot that the neutrino-driven wind would be too dirty to produce a clean short hard GRB. Another plausible time scale would be the differential-rotation-damping time scale due to magnetic braking and viscosity. Shapiro (2000) estimated that the time scale for magnetic braking of differential rotation by Alfvén waves is given by

$$t_{\rm A} = \frac{R}{v_{\rm A}} \simeq (1\ {\rm s})\, B_{14}^{-1} \left(\frac{R}{20\ {\rm km}}\right)^{-1/2} \left(\frac{M}{3M_\odot}\right)^{1/2}. \tag{11.67}$$

So for a NS with a magnetar field strength, the differential rotation damping time scale would be comparable to the duration of the short GRB. However, the issue of a dirty neutrino-driven wind still exists. It is unclear whether a clean fireball can be launched during this phase.

In order to avoid the problem, Rezzolla and Kumar (2015) and Ciolfi and Siegel (2015) proposed the so-called "time-reversal" scenario. In such a model, a NS–NS merger produces a supra-massive magnetar which ejects an energetic wind that dissipates through the interaction of an inner pair-rich wind and an outer baryon wind to produce X-rays. The supra-massive NS collapses later and produces the short GRB prompt emission through accretion. Since it takes time for X-ray photons to diffuse out, the GRB signal may reach the observer earlier than X-rays, so that the short GRBs would be followed by X-rays due to internal dissipation. One important prediction of such a scenario is that there should be X-ray emission prior to the short GRB detection. This scenario however suffers two

difficulties. First, when a supra-massive NS collapses, there is essentially no left-over material for forming a debris disk to power a GRB jet via accretion (Margalit et al., 2015). Second, since X-rays are diffused out through an interaction far from the central engine, it is difficult to produce an internal plateau with the post-plateau decay index much steeper than -3.

11.4 Quark Stars

11.4.1 Concept of Quark Stars

The concept of quark stars (QSs) is based on the so-called Bodmer–Witten hypothesis (Bodmer, 1971; Witten, 1984), namely, at an extremely high density and a nearly zero temperature, a three-flavor (*uds*) quark–gluon plasma (QGP) is more stable than baryons and a two-flavor (*ud*) QGP. The rationale of this hypothesis lies in that the strange quark (95 MeV/c^2) is the third lightest quark after the up (2.3 MeV/c^2) and down (4.8 MeV/c^2) quarks. The next lightest, the charm quark, is much more massive: 1.275 GeV/c^2. At extremely high densities, protons (*uud*) and neutrons (*udd*) are dissolved due to quark deconfinement and make a two-flavor (*ud*) QGP. Up and down quarks would convert to other flavors of quarks (strange, charm, etc.) via weak interaction to lower the Fermi energy by increasing the degeneracy. In practice, given the chemical potential involved (roughly 300 MeV), only up, down, and strange quark flavors occur in the QGP, which is more stable than the *ud* two-flavor QGP. Such a QGP is called strange quark matter. The relevant weak interactions include

$$d \to u + e^- + \bar{\nu}_e, \tag{11.68}$$

$$u + e^- \to d + \nu_e, \tag{11.69}$$

$$s \to u + e^- + \bar{\nu}_e, \tag{11.70}$$

$$u + e^- \to s + \nu_e, \tag{11.71}$$

$$s + u \leftrightarrow d + u. \tag{11.72}$$

On the astrophysical scale, it is hypothesized that, since the three-flavor QGP is more stable, an entire NS may be converted to a star made of strange quark matter. Such a star is called a *strange quark star*, *strange star*, or *quark star* (Alcock et al., 1986; Haensel et al., 1986). The properties of a strange star include:

- Since the stars are bound via strong interaction rather than gravity, their M–R relation is opposite to that of neutron stars, i.e. the larger the mass, the larger the size. Neutron stars, on the other hand, need to reduce size to increase degeneracy pressure in order to balance gravity when mass increases (Fig. 11.4).
- According to the simple MIT bag model that treats strange quark matter as gas, the pressure reads

$$P = \frac{1}{3}(\rho - 4B), \tag{11.73}$$

where $B \simeq (145\ \mathrm{MeV})^4 \simeq 57\ \mathrm{MeV\,fm^{-3}}$ is the vacuum energy density (Alcock et al., 1986), but with a large uncertainty. Based on such a treatment, strange quark stars have a maximum mass below $2M_\odot$, so are essentially ruled out for interpreting Galactic pulsar-like objects.

- Detailed studies of the phase structures of QSs suggest that their properties can be very different from what the simple model suggest. A first-principles calculation is unachievable due to the complicated non-linear and non-perturbative nature of quantum chromodynamics (QCD). Nonetheless, in the so-called confined-density-dependent-mass (CDDM) models, density dependence of quark masses and the inclusion of leading-order QCD perturbative interactions make it possible to produce QSs more massive than $2M_\odot$ (e.g. Li et al. 2016a and references therein). It is also possible that multiple quarks may form "clusters" or large "strange nucleons" due to local interactions. The QSs in such a case would carry the properties of a solid (Xu, 2003). The maximum mass of such QSs may also substantially exceed $2M_\odot$ (Lai and Xu, 2009).
- If the quark matter is exposed at the QS surface, the QS is called "naked". A naked QS would serve as a "membrane", since matter can only enter the strange quark matter body but cannot get out due to the strong interaction binding of the star (Alcock et al., 1986; Cheng and Dai, 1996; Paczyński and Haensel, 2005). Since there is a sharp boundary of the quark matter, while electrons are distributed in a more extended layer, there exists a strong electric field emerging from the QS surface. A bare QS could be a bright emitter of electron–positron pairs (Usov, 1998). It is possible that QSs may carry a normal matter crust (with a smaller mass than the typical NS crust) (Alcock et al., 1986; Huang and Lu, 1997). However, in astrophysical situations it may not be easy to form a crusted QS through fallback or accretion (Xu et al., 2001).
- Since a three-flavor QGP is more stable than a two-flavor QGP and normal matter, as normal matter is converted to strange quark matter, energy is released. The energy release per baryon when normal matter is converted to strange quark matter is about $\sim$10 MeV (Dai et al., 1995). As a result, converting an entire NS (with baryon number $\sim 10^{57}$) to a QS would release an energy of the order of $\sim 10^{52}$ erg (Cheng and Dai, 1996; Ouyed et al., 2002). Since a QS is more stable than a NS, gravitational energy would also be released, which is of the order of 10^{53} erg (Bombaci and Datta, 2000).

11.4.2 Quark Star as GRB Central Engine

So far there is no compelling evidence for the existence of QSs in the universe. However, there is also no compelling reason why they should not exist. Being speculative in nature, QSs have been proposed as the engine of GRBs by many authors based on the following arguments:

- Cheng and Dai (1996) proposed the conversion of NSs to QSs as the engine of GRBs. They emphasized that QSs can produce clean fireballs due to their membrane nature. This scenario was revived later by different authors (e.g. Bombaci and Datta, 2000; Ouyed et al., 2002; Paczyński and Haensel, 2005).

- Dai and Lu (1998c) pointed out that a QS engine can naturally give rise to continuous energy injection into the blastwave to power the afterglow. However, this feature is not unique to QSs, since a NS (magnetar) engine can also provide the energy injection signature to the blastwave.
- Ouyed and Sannino (2002) introduced the onset of exotic phases of quark matter at the surface of bare QSs, in particular, the so-called two-flavor color superconductivity. They argued that such a phase would unstably generate particles erratically, a crucial feature of a GRB central engine.
- Xu and Liang (2009) suggested that if a QS is solidified some time after formation, quakes in the solid QS would power X-ray flares.
- Drago et al. (2016) considered a NS–NS merger producing a QS to power a short GRB, which can produce both a short hard spike and an internal X-ray plateau.
- Li et al. (2016a) applied the short GRB X-ray plateau data to test various NS/QS EoSs. Observationally, the end of the internal X-ray plateau (t_b) clusters around several hundred seconds. Interpreting it as the supra-massive NS collapse time, the NS model usually produces too wide a distribution of $t_b = t_c$ (Gao et al., 2016). Li et al. (2016a) found that for QS EoSs the t_b distribution is much narrower, which matches the data better. This is due to the more sensitive dependence of $M_{\rm max}$ on spin for QSs than NSs. This hints that NS–NS mergers may make QSs rather than NSs.

The discovery of an r-process-powered macronova/kilonova (Coulter et al., 2017; Shappee et al., 2017; Pian et al., 2017; Evans et al., 2017; Tanvir et al., 2017; Nicholl et al., 2017; Chornock et al., 2017) associated with the NS–NS merger gravitational wave source GW170817 (Abbott et al., 2017d) suggested that the two merging compact objects with NS masses are indeed NSs. The required ejecta mass to power the kilonova is typically in the range of $0.03–0.05 M_\odot$, which is much greater than the available neutron-rich material in standard QSs. As a result, this observation essentially rules out the existence of the standard QSs at the NS density.[1] The possibility of forming a QS during the merger process may not be ruled out, even though the very small energy budget in the EM channel of the GW170817/GRB 170817A system also posed some constraints on the scenario (the phase transition from neutron matter to strange quark matter is expected to release a significant amount of energy, which seems to have exceeded the energy budget of GRB 170817A).

11.5 Late Central Engine Activity

Swift observations suggest that GRB central engine activities last much longer than the duration of the prompt emission. There are two types of extended engine activities: the

[1] The QSs made of strange nucleons (or strangeons) may be still allowed (Lai et al., 2018).

erratic one manifested as late X-ray flares and the steady one manifested as the internal X-ray plateau. These observations raise more constraints on the possible central engines.

As discussed in §11.3, the internal X-ray plateau may be related to the spindown of a millisecond magnetar central engine (could be either a NS or a QS). A BH engine may also produce steady X-ray emission. Indeed, Kumar et al. (2008b,a) interpreted the entire canonical X-ray lightcurves (steep decay, shallow decay, and the follow-up normal decay) as internal jet emission powered by the fallback of a massive stellar envelope with distinct internal density boundary layers. However, these features can also be interpreted within the framework of the external shock model (Zhang et al., 2006; Nousek et al., 2006; Panaitescu et al., 2006b; Wang et al., 2015b). The BH models may encounter difficulties explaining those GRBs showing an internal X-ray plateau with an extremely rapid drop at the end of the plateau (Kisaka and Ioka, 2015). More detailed investigations are needed to see whether a rapid transition of accretion mode (e.g. from NDAF to ADAF) may be consistent with the data.

The X-ray flares need to quench and restart the central engine multiple times. Prompted by observations, the following ideas have been proposed for interpreting X-ray flares following GRBs:

- King et al. (2005) suggested that the massive star progenitor of the long GRBs may fragment into multiple clumps during core collapse, which are subsequently accreted into the newborn BH. This can power distinct X-ray flares observed following long GRBs.
- Perna et al. (2006) argued that, since X-ray flares are also observed to follow some short GRBs, the flare mechanism must apply to both types of bursts. Instead of invoking fragmentation of the star, they suggested that fragmentation within the BH accretion disk would be the origin of X-ray flares. The condition for gravitational instability is

$$Q_{\rm T} \equiv \frac{c_s \kappa}{\pi G \Sigma} < 1, \tag{11.74}$$

  where $Q_{\rm T}$ is the Toomre parameter, Σ is the surface density of the disk, c_s is sound speed, and κ is the epicyclic frequency, i.e. the frequency at which a radially displaced parcel would oscillate. Perna et al. (2006) suggested that this condition is readily satisfied for massive star GRBs and may be marginally satisfied in compact star GRBs as well. Dall'Osso et al. (2017) showed that such a model can well explain many observed properties of X-ray flares.
- Proga and Zhang (2006) introduced the concept of magnetic barriers seen in the MHD simulations of BH accretion (Proga and Begelman, 2003) to GRBs and conjectured that the accretion flow may be temporarily stopped by a magnetic barrier near the BH for a highly magnetized accretion flow. Accretion may resume later after enough materials are accumulated to break the barrier, powering the X-ray flares. Such a scenario was studied in more detail by different authors, who reached similar conclusions (McKinney et al., 2012; Cao et al., 2014).
- Following the spirit of the prompt emission model of Kluźniak and Ruderman (1998), Dai et al. (2006) suggested that if a NS–NS merger produces a differentially rotating NS with a dipole field of a few 10^{10} G, magnetic buoyancy would launch magnetic bubbles

generated from differential rotation of the NS. These magnetic bubbles would power the X-ray flares following short GRBs (Dai et al., 2006).

- Lee et al. (2009) considered He synthesis in the disk of short GRBs. They suggested that recombination of free nucleons into α-particles in the disk would modulate the accretion rate through launching powerful winds, which shut off accretion episodically to power X-ray flares following short GRBs.
- Wu et al. (2013) interpreted the bright X-ray flares with extreme re-brightening and extended decay as the result of fallback of the He envelope of a massive star after the intial explosion.