ABSTRACT

The variability observed in many complex gamma-ray bursts (GRBs) is inconsistent with causally connected variations in a single, symmetric, relativistic shell interacting with the ambient material (“external shocks”). Rather, either the central site must produce $\sim 10^{50}$ erg s$^{-1}$ for hundreds of seconds (“internal shocks”), or the local spherical symmetry of the shell must be broken on an angular scale much smaller than $\Gamma^{-1}$, where $\Gamma$ is the bulk Lorentz factor for the shell. The observed variability in the external shock models arises from the number of causally connected regions that (randomly) become active. We define the surface filling factor to be the ratio of the area of causally connected regions that become active to the observable area of the shell. From the observed variability in 52 BATSE bursts, we estimate the surface filling factor to be typically $\sim 5 \times 10^{-3}$, although some values are near unity. We find that the surface filling factor, $f$, is $\sim 0.1 \Delta T/T$ in both the constant $\Gamma$ phase (which probably produces the GRB) and the decelerating phase (which probably produces the X-ray afterglows). Here, $\Delta T$ is a typical timescale of variability, and $T$ is the time since the initial signal. We analyze the 2 hr flare seen by ASCA 36 hr after the GRB and conclude that the surface filling factor must be small ($10^{-4}$) in the X-ray afterglow phase as well. Compared with the energy required for an isotropic shell, $E_{\text{iso}}$, explanations for a low surface filling factor can either require more energy ($f^{-1}E_{\text{iso}} \sim 10^{56}$ ergs) or less energy ($[\Delta T/4T]^2E_{\text{iso}} \sim 10^{49}$ ergs). Thus, the low filling factor cannot be used as a strong argument that GRBs must be internal shocks.

Subject headings: gamma rays: bursts — relativity

1. INTRODUCTION

Gamma-ray burst (GRB) spectra often extend to very high energies with no indication of attenuation by photon-photon interactions. This implies substantial relativistic bulk motion of the radiating material with Lorentz factors in the range of $10^2$–$10^3$. At cosmological distances, GRBs require an energy reservoir on the order of $10^{52}$ ergs. The likely sources of such a reservoir would be the rest mass of a compact object released during a merger (e.g., either neutron star–neutron star or neutron star–black hole). However, most detailed calculations of mergers occur on timescales of less than 1 s, which is much less than the observed durations of GRBs (often $10^2$ and sometimes $10^3$ s).

Two classes of models have arisen that explain various (but not all) aspects of the observations. In the “external” shock models (Mészáros & Rees 1993), the release of energy during the merger is very quick, and a relativistic shell forms that expands outward for a long period of time ($10^2$–$10^7$ s). At some point, interactions with the external medium (hence the name) cause the energy of the bulk motion to be converted to gamma rays. Although the shell might produce gamma rays for a long period of time, the shell keeps up with the photons such that they arrive at a detector over a relatively short period of time. If the shell has a velocity $v = \beta c$ with a corresponding bulk Lorentz factor $\Gamma = (1 - \beta^2)^{-1/2}$, then photons emitted over a period $t$ arrive at a detector over a much shorter period, $T = (1 - \beta v) = t/(2\Gamma^2)$. Although this model is consistent with the short energy release expected for a merger and the observed long timescale of GRBs, we have argued that it cannot explain the long, complex time histories except under extreme conditions (Fenimore, Madras, & Nayakshin 1996). In particular, we argue from kinematic considerations that the long gaps often seen in GRBs are inconsistent with local spherical symmetry, the rapid time variability implies that only a small fraction of the shell becomes active, and the observed average envelope of emission is inconsistent with that expected from a relativistic shell. These arguments are contained in Fenimore et al. (1996), Fenimore & Sumner (1997), and Fenimore, Ramirez-Ruiz, & Sumner (1998). In addition, Dar (1998), Sari & Piran (1997b), and Kobayashi, Piran, & Sari (1997) argue from hydrodynamic considerations that the external shock model cannot provide the energy or timescales observed in GRBs.

The alternative theory is that a central site releases energy in the form of a wind or multiple shells over a period of time commensurate with the observed duration of the GRB (Rees & Mészáros 1994). The gamma rays are produced by the internal interactions within the wind; hence these scenarios are often referred to as internal shock models. These models have two weaknesses: first, there is a concern that internal shocks are rather inefficient (although, see Kobayashi et al. 1997), and second, the long, complex time history of a GRB must be postulated at the central site. On the other hand, the free parameters associated with the multiple shells can probably explain any observed time history (Kobayashi et al. 1997).

The need to explain long time histories was further complicated by the discovery of the X-ray afterglows lasting hours (Costa et al. 1997), the optical afterglows lasting
weeks to months (Metzger et al. 1997), and the radio afterglows lasting many months (Frai et al. 1997). These extended timescales appear too long to have been produced by a lingering central site made by a merger. In addition, the observed power-law decay is expected from many external shock models (Wijers, Rees, & Mészáros 1997; Reichart 1997; Tavani 1997; Waxman, Kulikarni, & Frai 1998; Panaitescu & Mészáros 1998; Mészáros, Rees, & Wijers 1998; Rees & Mészáros 1998; Sari, Piran, & Narayan 1998). Piran & Sari (1997) suggested that the initial gamma-ray phase is due to internal shocks from a relativistic wind (or multiple shells) that merge into a single relativistic shell, which then produces the afterglows in a manner similar to the external shock models. This model avoids the difficulty of explaining the rapid variability of the gamma-ray phase with a single shell while retaining the long timescale capability of the single shell for the afterglows.

The purpose of this paper is to demonstrate that the external shock model cannot utilize the full surface of shells. Fundamentally, it arises because of the relatively rapid time variability of the gamma-ray phase. Sari & Piran (1997a) use this as an argument that external shock models would require too much energy. However, this should not be considered a strong argument that the GRB phase is not external shocks, since there are several scenarios that can utilize only a portion of the shell's surface. We also analyze the time variability recently observed by ASCA in the X-ray afterglow of GRB 970828 (Yoshida et al. 1998) and show that its time variability implies low surface utilization in the X-ray afterglow phase as well.

2. RAPID VARIABILITY FROM RELATIVISTIC SHELLS

To understand the problem caused by rapid variability, one must emphasize the difference between arrival time at the detector (which we denote with \( T \)) and coordinate time (equivalent to laboratory time, the detector rest frame time, and the rest frame of the central explosion, denoted by \( t \)). Coordinate time is measured by clocks placed at all locations within the frame and can measure when the photons were produced. In contrast, a detector (such as BATSE) measures when the photons arrive at a single location. These two times are related as \( T = (1 - \beta \cos \theta) t \), where \( \theta \) is the angle between the direction of motion of the emitting material and the direction to the detector. The large bulk Lorentz factor means that the shell is almost certainly moving directly at the observer's detector (i.e., we must be looking at the shell head-on rather than looking at the sides of a jet), so \( \theta \) is small. When \( \theta \) is 0, \( T = t/(2\Gamma^2) \). In contrast, the Lorentz transformation between the time in the rest frame of the detector (\( t \)) and time in a frame moving with the shell (\( t' \)) is \( t = \Gamma t' \). Consider the early behavior of the shell. Assume that clocks are placed along the path of the shell in the detector's rest frame. These clocks report coordinate time; that is, the time the shell passes the location of the clocks. Consider an extreme case, such as \( \Gamma = 10^3 \). Clocks placed with a spacing of \( 3 \times 10^{16} \) cm will see the shell pass their locations at times separated by \( 10^5 \) s. Because the shell is viewed by the detector head-on, its movement causes the shell to keep up with the photons it produces such that photons separated by \( 10^6 \) s in the detector rest frame arrive at the detector separated by a much shorter interval: \( T = (1 - \beta t = t/(2\Gamma^2) \). If the shell expands in the detector rest frame for \( 10^7 \) s, the photons all arrive at the detector within 5 s. Note that this contraction of the photon arrival time is not the result of a Lorentz transformation. Rather, it reflects that a detector at a single location measures when photons arrive, not necessarily the relative time they were emitted.

In the early phase of the shell's expansion, \( \Gamma \) is effectively constant (\( \approx \Gamma_0 \)), such that \( t = T/(2\Gamma_0^2) \). Eventually, the shell begins to decelerate as it sweeps up the interstellar medium (ISM). Without detailed knowledge of the physical process that generates the gamma rays, it is not clear whether the expansion is adiabatic or dominated by radiation losses. However, both solutions lead to a power-law decay (Sari et al. 1998; Mészáros et al. 1998; Rees & Mészáros 1998), and we will show that our conclusions only depend on the power-law nature of the decay and not on the exact index of the power law. For the sake of example, we will assume that the expansion is adiabatic. The energy required to sweep up an ISM mass of \( m_0 \) is \( \Gamma^2 m_0 c^2 \) (Mészáros & Rees 1997). Let \( E_0 \) be the initial energy deposited into the shell such that \( E_0 = \Gamma_0 M_s c^2 \), where \( M_s \) is the original mass of the shell. If \( \rho \) is the density of the ISM, one can solve for the radius of the shell as a function of the coordinate time by solving

\[
E_0 = \Gamma(t)M_s c^2 + \Gamma^2(t)\rho \frac{d\Omega}{3} r^3(t)c^2.
\]

Here, \( d\Omega \) is the angular size of the ejecta, accounting for the situation where the expansion is mechanically confined to a jet. The arrival time of the photons as a function of the coordinate time that the photons were emitted can be found from

\[
\frac{dT}{dt} = 1 - \beta(t) = \frac{1}{2\Gamma^2(t)}
\]

and \( dr/dt = \beta(t) \). Figure 1 shows \( t \) as a function of \( T \) for a typical situation (\( 4\pi E_0 /d\Omega = 10^{44} \text{ ergs}, \rho = 1 \text{ atom cm}^{-3} \)). It is unclear what fraction of the shell's energy is converted to gamma rays. Observations imply that the energy in gamma rays if the explosion is isotropic is on the order of \( 10^{52} \text{ ergs} \). If \( d\Omega \) is much less than \( 4\pi \), much less energy is required (implying many more unseen events). The observations only limit the energy per \( 4\pi \text{ sr} \) to be \( \sim 10^{52} \text{ ergs} \). We scale our solutions to \( 4\pi E_0 /d\Omega = 10^{54} \text{ ergs} \), which assumes that \( \sim 1\% \) of the energy is converted to gamma rays.

There are three phases of the \( t - T \) relationship: the initial constant-\( \Gamma \) phase, a phase where \( t = 1.63 \times 10^4(4\pi E_{54}/\rho_0 d\Omega)^{1/2}T^{1/4} \), and a final phase when the shell becomes nonrelativistic, where \( t = T + \text{constant} \). Here, \( E_{54} \) is \( E_0 \) in units of \( 10^{54} \) ergs and \( \rho_0 \) is the ISM density in units of 1 atom cm\(^{-3}\). The second and third phase are only weakly dependent on the initial conditions. Neither depends on the initial \( \Gamma_0 \), and they depend on the other parameters to the \( \frac{1}{2} \) power. Thus, except for the duration of the first phase, different bursts ought to have similar \( t - T \) relationships. If one assumes the extreme nonadiabatic case that the momentum of the shell is conserved rather than the energy, then the \( \frac{1}{2} \) power becomes \( \frac{1}{4} \). Many other indexes as possible (see Sari et al. 1998; Mészáros et al. 1998; Rees & Mészáros 1998).

Although the exact nature of the conversion of bulk motion to gamma rays is unclear, we argue that the gamma-ray phase must be contained in the first phase, where \( t = T/(2\Gamma_0^2) \). We note that many GRBs show a similar time-of-arrival structure (e.g., observed subpeak widths) at the beginning and end of the burst. In the first phase, the
observed arrival-time structure ($\Delta T$) is related to the emission time as $\Delta t/(2T^2_0)$ such that a constant subpeak width can occur if the emission time is constant in coordinate time. In the second phase, $\Delta T \propto t^3\Delta t$ such that constant subpeak width can only occur if the emission time ($\Delta t$) scales as $t^{-3}$. This is consistent with our scaling of $4\pi E_0/d\Omega = 10^{34}$ ergs, which allows the first phase to last 200 s (see Fig. 1). The afterglows probably occur in the $t \propto T^{1/4}$ phase, where $\Gamma$ is more modest.

As one can see from Figure 1, our view of GRBs becomes quite distorted because we only measure arrival time at the earth rather than the time of emission in the Earth’s rest frame (remember, no Lorentz transformations are involved). It is often said that GRBs are quite quick, say 30 s. Actually, in the external shock models, they occur longer by $2T^2$ or, perhaps, up to $6 \times 10^7$ s if $\Gamma$ is $10^{3}$. The afterglow period seems to cover a very large, dynamic range, from $10^3$ to $10^7$ s. Actually, because $t \propto T^{1/4}$, the dynamic range is only a factor of 10.

3. RELATIONSHIP BETWEEN THE SURFACE FILLING FACTOR AND THE OBSERVATIONS

In the external shock model, there is a single shell, and the rapid variability we observe can only be accounted for by a nonuniform production of gamma rays by the shell. Consider BATSE bursts 1883 and 2993, shown in Figure 2. They have very similar overall envelopes of emission and similar intensities. Burst 2993, however, has stronger deviations from the envelope in the form of multiple peaks. Presumably this structure arises because various patches (or emitting “entities”) on the shell randomly become active. In the case of burst 1883, so many entities become simultaneously active that the overall envelope appears quite smooth, whereas in the case of burst 2993, fewer entities become active, so random fluctuations in the number of simultaneously active entities cause the peak structure to be spiky. In this view, each observed peak in not necessarily caused by a single entity, but the peak structure is caused by random variations in the number of active entities (see Sumner & Fenimore 1998).

---

**Figure 1.**—Effects of the relativistic motion on our perception of GRB timescales. The x axis is the time of arrival of photons at a detector from a relativistic shell moving directly at the detector. This time is not the emission time in the detector’s rest frame, which can be measured only by clocks placed at rest with the detector at the location of the emission. a: The bulk Lorentz factor, $\Gamma$, as a function of the arrival time. At first, the shell expands with a constant $\Gamma$. Eventually, the shell decelerates as it sweeps up the ISM, resulting in $\Gamma(T)$ varying as $T^{-3/8}$. Although we have assumed only small energy losses from radiation, nonadiabatic solutions are similar. b: The differential relationship between the time of emission as measured in the detector’s rest frame time ($t$, $y$ axis) and the time of arrival of the photons at the detector ($T$, $x$ axis). Because a shell moving directly at the detector keeps up with the photons it produces, $dT/dt = 1 - \Gamma(t)/\Gamma(0)/2$. Note that no Lorentz transformations are involved. c: The integral relationship between the detector’s rest frame time and the arrival time at the detector. The left $y$ axis is the time of emission in the detector’s rest frame. The right axis is the radius of the shell at the time of emission (simply $c$ times the left axis). Note that during the gamma-ray phase the event appears to last for a short time (i.e., $T$ is a few hundred seconds). In the detector’s rest frame time, however, this corresponds to a long time ($t \sim 4 \times 10^9$ s). The afterglow phase (where $t \propto T^{1/4}$) appears to cover a wide dynamic range of time ($10^3$–$10^7$ s). Due to deceleration, it actually only covers a factor of 10 in time.

**Figure 2.**—Two similar bursts with different surface filling factors, that is, different fractions of the relativistic shell that become active. A causally connected region is much smaller than the available area of the shell. The smoothness of BATSE burst 1883 implies that many casually connected regions are simultaneously active, whereas the nonstatistical fluctuations in BATSE burst 2993 implies that only a few regions are active. For burst 1883 we estimate that $1 \pm 0.13$ of the shell was active, and for burst 2993 we estimate $0.012 \pm 0.006$. 
In Fenimore et al. (1996) and Fenimore & Sumner (1997), we showed that the expected envelope of the GRB phase has a FRED-like shape (fast rise, exponential decay, although the decays are often just slower, not necessarily exponential). The expected envelope is a power law, $T^{-\alpha}$, where $\alpha$ is the spectral number index (typically $\sim 1.5$). (Note that this power law refers to the GRB phase, in contrast to the well-observed power-law decay found in the afterglows.) In Sumner & Fenimore (1998), we showed that the volume contributing to a set period of time, $\Delta T$, is constant. The $T^{-\alpha}$ envelope comes from the beaming. Thus, if the entities form randomly, one can expect that the rate of entities causing peaks is effectively constant during the decay phase. Indeed, the presence of long gaps in GRBs is one of our key arguments that the GRB phase is not due to a simple single shell. Gaps require large, causally disconnected regions to coordinate their lack of emission (see Fenimore et al. 1996, 1998).

We define the surface filling factor as the fraction of the shell’s surface that becomes active. Let $A_s$ be the area of an entity and $N_s$ be the number of entities that (randomly) become active during the interval $T_{obs}$. If $A_{obs}$ is the area of the shell that can contribute during $T_{obs}$, then the surface filling factor is

$$f = N_s A_s / N_s A_{obs}$$

where $\eta$ is the fraction of the visible area of the shell ($A_s$) that contributes during the interval $T_{obs}$.

The number of entities can be determined from the observed fluctuations in the time history. The overall envelope is due to the beaming, so the residue variations about the overall envelope are due to a combination of the Poisson variations in the number of entities and the Poisson variations associated with the count statistics. Assume for the moment that the contribution due to the count statistics is small such that the variations come from the Poisson variations in the number of contributing entities. We first remove the envelope by fitting a polynomial function to it. The observations are divided by the polynomial function, so the result is a flat envelope with variations due to the number of entities that, on average, are active simultaneously. The rate of occurrence of the entities, $\mu_N$, can be found directly from the variations, because the observed mean level is $K \mu_N$, where $K$ is some constant, and the variance is $K^2 \mu_N$. We define $N$ to be the observed mean level of the flattened envelope and $\delta N$ to be the rms of the flattened envelope. In the case of no contribution from the counting statistics, $\mu_N$ would be $(N/\delta N)^2$. Here we have implicitly assumed that all entities are identical. This is supported by the fact that peaks within GRBs usually are similar to each other. The actual rms of the flattened envelope is a combination of the variance due to the entities and the variance due to the counting statistics, $\sigma_{CS}^2$. We assume that they add in quadrature; that is, $N^2/\delta N^2 = \mu_N + \sigma_{CS}^2$. We estimate $\sigma_{CS}$ to be the rms of many Monte Carlo realizations of the flattened envelope.

The rate of occurrence of entities is

$$\mu_N = N^2 / (\delta N)^2 - \sigma_{CS}^2 .$$

This rate is the number of events per the timescale of the entities. Thus the total number of entities that occur within a period $T_{obs}$ is

$$N_N = \mu_N T_{obs} / \Delta T_p,$$

where $\Delta T_p$ is the timescale for a single entity.

We have considered several scenarios that relate the size of an entity causing a peak to an observed peak duration, $\Delta T_p$ (see Table 1 in Fenimore et al. 1996). Here we consider the two most likely scenarios for the formation of a peak: regions that grow and regions formed by interaction with the ISM.

Consider a gamma-ray-producing region that grows at the speed of sound, $c_s$, for a period $\Delta t'$ in the rest frame of the shell, where presumably the region is symmetric. This growth might be associated with a developing shock. Let $\Delta r ||$ be the radius of the region in the rest frame along the direction of the motion and $\Delta r'\perp$ be the radius in the perpendicular direction such that

$$\Delta r_\perp = c_s \Delta t' = c_s \Gamma^{-1} \Delta t' .$$

The sizes in the rest frame of the detector are related to the sizes in the rest frame of the shell as $\Delta r'_\parallel = \Gamma^{-1} \Delta r ||$ and $\Delta r'_\perp = \Delta r'\perp$.

Combining the effects of the movement of the shell during the growth with the maximum size that the entity can grow in time $\Delta t$, we find that the duration in arrival time is (see Table 1 of Fenimore et al. 1996)

$$\Delta T_p = \frac{\Delta t}{\Gamma^2 \left[ \left( \frac{1}{2} \right)^2 + \left( \frac{c_s}{c} \right)^2 \right]^{1/2}} .$$

The size of the emitting entity is

$$A_N = \pi \Delta r^2 = \pi \left( \frac{c_s \Delta t}{\Gamma} \right)^2 = \frac{\pi c_s^2 \Gamma^2 \Delta T_p^2}{[(1/2)^2 + (c_s/c)^2]} .$$

The alternative cause of a peak resulting from the shell is that the shell interacts with an ambient object, such as an ISM cloud. Presumably the object is symmetric such that $\Delta R = \Delta R || = \Delta R_{amb}$. We use $\Delta R$ for an object that grows in a shell and $\Delta R$ for an ambient object. We assume the case of a “collapsible” object (such as an ISM cloud; see Fenimore et al. 1996). The contribution to the peak duration from the time the shell takes to move through the cloud [i.e., $\Delta R_{amb}/(c\Gamma^2)$] is negligible compared with the time the shell takes to engage the perpendicular size of the object. This engagement time is caused by the curvature of the shell. At an angle $\theta$ from the line of sight, the time to engage the object is $\Delta T_{AR} = \theta \Delta R_{amb}/2c$. Note that Table 2 of Fenimore et al. (1996) was incorrect for $\Delta T_{AR}$; see Sari & Piran (1997a). At a typical angle of $\theta \sim \Gamma^{-1}$,

$$\Delta R_{amb} = \frac{c \Delta T_p \Gamma}{2} ,$$

$$A_N = \pi \Delta R^2 = \frac{\pi c^2 \Gamma^2 \Delta T_p}{4} .$$

Thus, both the case of shocks that grow from a seed and the case of shells running into ambient objects are similar. If $c_s = c/3$, then $A_N$ from equation (8) is $16/13$ times larger than that from equation (10). These two scenarios only differ by a constant of the order of unity.
The final ingredient for the calculation of the surface filling factor is the area of the shell visible to the observer:

\[ A_s = 2\pi R^2(T)(1 - \cos \theta_{\text{max}}), \]

where \( \theta_{\text{max}} \) is either the angular width of the shell (i.e., \( \Delta \Omega = 2\pi(1 - \cos \theta_{\text{max}}) \)) or \( \Gamma^{-1} \), whichever is smaller. Before the shell starts to decelerate, presumably \( \theta_{\text{max}} \) is larger than \( \Gamma^{-1} \). Using \( R(t) = 2\Gamma_0^2 cT \),

\[ A_s = \pi [\Gamma^{-1} R(t)]^2 = 4\pi c^2 \Gamma_0^2 T^2, \quad \text{if } \Gamma(T) = \Gamma_0. \]

Once the shell starts to decelerate,

\[ \Gamma(T) = 463 \left( \frac{4\pi c^2 \Gamma_0^2}{\rho_0 d\Omega} \right)^{1/8} T^{-3/8}, \] \hspace{1cm} (13a)

\[ R(T) = 0.016 \text{ pc} \left( \frac{4\pi c^2 \Gamma_0^2}{\rho_0 d\Omega} \right)^{1/4} T^{1/4} \approx \text{ct}, \]

\[ (13b) \]

such that

\[ A_s = \left\{ \begin{array}{ll}
2.5 \times 10^{-4} d\Omega \text{ pc}^2 \left( \frac{4\pi c^2 \Gamma_0^2}{\rho_0 d\Omega} \right)^{1/2} T^{1/2}, & \text{if } d\Omega < 2\pi(1 - \cos \Gamma^{-1}), \\
3.67 \times 10^{-9} \text{ pc}^2 \left( \frac{4\pi c^2 \Gamma_0^2}{\rho_0 d\Omega} \right)^{1/4} T^{5/4}, & \text{if } d\Omega > 2\pi(1 - \cos \Gamma^{-1}).
\end{array} \right. \]

Combining equations (3)–(5) and (12)–(14), we find three different cases. Case a is the constant factor is beaming angle, and case c is when the deceleration reduces the initial deceleration when the size of the shell exceeds the region of the shell at angle \( \leq 1 \). Using the prolate profile is from the \( \Gamma^{-1} \) that contributes to the signal during time \( T \). Unfortunately, we do not know how, in general, to map the amount of active area of a shell to the observed time history. Only in the particular case of a FRED-like burst can we reliably estimate \( \eta \). For a FRED-like burst, the curvature of the shell delays the photons from the region of the shell at angle \( \Gamma^{-1} \) by \( T_0 \) where \( T_0 \) is the time that the shell expands before producing gamma rays. From Fenimore et al. (1996) and Fenimore & Sumner (1997), one can find \( T_0 \) from the shape of the profile. Given that the profile shape is \( (T_0 / T)^{-\alpha} \), \( T_0 \approx 0.8 T_{50} \), where \( T_{50} \) is the duration of the burst that contains 50% of the event counts (see BATSE catalog; Meegan et al. 1996). Here, \( \alpha \) is the spectral index number, typically \( \approx 1.5 \). Thus, if one uses \( T_{\text{obs}} = T_0 = 0.8 T_{50} \), then the profile is from the region of the shell out to an angle of \( \Gamma^{-1} \), and \( \eta \approx 1 \).

For long complex bursts, we cannot easily detect a FRED-like shape. In lieu of an exact estimate, we will use the rule found for FRED-like bursts: \( T_{\text{obs}} = 0.8 T_{50} \). In Table 1, we list six FRED-like bursts and 46 long, complex bursts that were bright and had sufficient time structure for estimating parameters. The first column gives the BATSE trigger number. The second column gives \( \mu_4 \), as found from equation (4). The third column gives an estimate of \( \Delta T_p \), the temporal width of a single entity. In most cases, this was estimated from the width of the autocorrelation function. For four of the FRED-like bursts, the profiles were rather smooth (presumably due to a larger number of entities being simultaneously active), so we used a typical value for \( \Delta T_p \) (i.e., 0.5 s). The fourth column is \( T_{50} \), and the next two columns are, respectively, the surface filling factor and its propagated error based on equation (15).

In Figure 3 we show the distribution of surface filling factors as a function of burst duration \( T_{50} \). The solid squares are the FRED-like bursts, and the open squares are the long, complex bursts. Although some of the smooth
FRED-like bursts can have surface filling factors near unity, most bursts have values on the order of $5 \times 10^{-3}$.  

### 4.2. Filling Factor during the X-Ray Afterglow

In Figure 4 we show the time history of burst GRB 970828, observed by the *ASCA* satellite, from 28.9 to 48.9 hr after the gamma-ray phase (see Yoshida et al. 1998). At about 36.1 hr there was a flare that lasted about 2 hr. Piro, Matt, & Ricci (1997) also reported variations from a simple power-law decay in the X-ray afterglow of GRB 970508. These two examples of deviations from a power-law decay are quite different. In the case of the Yoshida et al. (1998) observation, the flare is relatively narrow (i.e., $\Delta T/T$ is small). In the case of the Piro et al. (1997) observation, the deviation could be a step function increase in the intensity (i.e., $\Delta T/T \sim 1$, although there is a data gap from 22 to 65 hr after the GRB that obscures the nature of the deviation).
We will emphasize the Yoshida et al. (1998) observation, because the narrow nature of the flare gives much stricter limits on the relativistic shell (see § 5). Except for Earth occultations, *ASCA* continuously observed the afterglow from 28.9 to 48.9 hr after the GRB. The observations have been prorated into samples of equal duration in Figure 4.

Consider how case c in equation (15) applies to GRB 970828. We must estimate the fraction of the shell that contributes to the 20 hr of observations, $\eta$. We can estimate a lower limit on $\eta$ and therefore an upper limit on $f$. In case c we assume that the whole shell was active, but only a fraction ($\eta$) of it contributed during the *ASCA* observation. The narrowest annulus that contributes during any particular observation is one with a width equal to the size of a single entity. At a typical angle of $\Gamma^{-1}$, $\eta$ can be as small as

$$\min \eta = \frac{\pi R_\perp \Delta R_\perp}{\pi R_\perp^2} = \frac{\Delta R_\perp}{R_\perp} = \frac{\Delta T_p}{T}.$$  

At smaller angles, $\eta$ would be smaller, but that is counteracted by the fact that the peak in the *ASCA* observation is clearly much smaller than the whole observation, such that the annulus responsible for the observation must be larger than we used in equation (17).

We only have observations covering a small range of time, so we cannot use equations (4) and (5) to estimate $N_N$. However, there are probably only a few active entities, because we see only one peak in Figure 4. Hence, we will use $N_N = 3$ as the maximum number of entities involved in the *ASCA* flare. Thus, using case c, $f$, is, at most, $\sim 0.02(\Delta T_p/T) = 10^{-3}$, which is a value very similar to that obtained in the gamma-ray phase. In Figure 3, the *ASCA* flare for case c is represented by the solid triangle.

In case b, we do not know $d\Omega$ or $\eta$. However, the smallest $\eta$ when the shell's angular size is limited to $d\Omega$ is

$$\min \eta = \frac{\Delta R_\perp}{R_\perp} = \frac{\Delta T_p}{T \sqrt{d\Omega/\pi}},$$  

such that

$$f = \frac{3.5 \times 10^{-8} N_N}{\sqrt{d\Omega}} \left( \frac{4\pi E_{54} \rho_0}{d\Omega} \right)^{-1/4} \frac{\Gamma \Delta T_p}{T^{1/4}}$$

$$= \frac{1.6 \times 10^{-5} N_N}{\sqrt{d\Omega}} \left( \frac{4\pi E_{54} \rho_0}{d\Omega} \right)^{-1/8} \frac{\Delta T_p}{T^{5/8}}.$$  

Using $N_N = 3$ and $4\pi E_{54}/\rho_0 d\Omega = 1$, the maximum $f$ under case b is $\sim 10^{-4} d\Omega^{-1/2}$.

In summary, $\Delta T/T$ in the *ASCA* flare is $\sim 0.05$, which implies a filling factor of $10^{-3}$. The X-ray afterglow suffers from the same time-history problems that the gamma-ray phase has: peaks do not have the value of $\Delta T/T$ expected from a shell whose surface quasi-uniformly emits photons.

5. DISCUSSION

The rapid variations in GRB time histories imply emitting entities the size of $\Delta R_c \sim \Delta T_p$. Assuming a single expanding shell, these entities must form on a much larger surface, $\sim c\Delta T$. We have defined $f$ (given by eq. [15]) to be the fraction of the surface of the shell that becomes active—that is, generates emitting entities. A crude estimate of $f$ is $\sim 0.1 \Delta T_p/T$. Because both the classic GRB phase and the recently discovered X-ray afterglow phase can have flares with $\Delta T_p/T \sim 0.05$, both time periods can suffer from an inefficient utilization of a single relativistic shell.

In the case of the gamma-ray phase, there is a simple explanation: the gamma-ray source is not a single relativistic shell but rather reflects activity at some central site that produces $\sim 10^{50}$ ergs s$^{-1}$ for up to several hundred seconds. Other observations, particularly gaps seen in the gamma-ray time histories (Fenimore et al. 1998), also argue for a central engine. The X-ray afterglow, however, is widely interpreted as being a single relativistic shell (Wijers et al. 1997; Reichart 1997; Waxman et al. 1998; Sari et al. 1998; Rees & Mészáros 1998). Therefore, we will concentrate our discussion on how to accommodate flares during the X-ray afterglow. These arguments could also be applied to the gamma-ray phase.

Of the two examples of deviations from a power law in the X-ray afterglow, the Yoshida et al. (1998) event (GRB 970828) is much more restrictive than the Piro et al. (1997) event (GRB 970508). GRB 970828 has a narrow flare, whereas GRB 970508 appears to be roughly a step function. A step function is similar to a FRED-like deviation that is approximately what one would expect if most of the observable surface of a relativistic shell interacts with a large ISM cloud (i.e., $\Delta T/T$ should be $\sim 0.25$; Fenimore et al. 1996). Thus GRB 970828 might be an afterglow with a low surface filling factor, whereas GRB 970508 has a high surface filling factor.

A common misconception is that one can just use an ISM cloud that covers most of the shell's surface and, therefore, make a flare that substantially changes the emission. This does not work, because the curvature of the expanding shell prevents the shell from engaging the cloud instantaneously. Rather, the portion of the shell at $\theta \sim \Gamma^{-1}$ requires a time $R(1 - \cos \theta)/c$ longer to reach the cloud. Even if the cloud happens to have a concave shape such that the shell reaches the cloud simultaneously over a wide range of angles, the resulting photons at $\theta \sim \Gamma^{-1}$ must travel farther to the detector, resulting in emission that is delayed by $R(1 - \cos \theta)/c$. Indeed, the delay of the photons due to the curvature is identical to the scenario of a shell that expands without producing photons for a long time and then emits only over a small range of times [the $P = P_0 \delta(t - t_0)$ case in Fenimore et al. 1996]. The result is a FRED-like flare (like GRB 970508) whose duration, $\Delta T$, is about $T_c$ rather than a small fraction of $T$, as is indicated by the *ASCA* observations for GRB 970828.

The shell will have well-defined curvature, because all points on it are moving at nearly the same speed—that is, the speed of light. In any case, there is no reason to believe that variations in the ambient material would cause the shell to develop into a plane wave oriented toward the observer such that the photons produced by an interaction with an ISM cloud or a shock would arrive as a short flare. Only the instantaneous interaction between two plain parallel surfaces oriented perpendicularly to our line of sight can produce a short peak from large surfaces.

We see six possible explanations for flares with small $\Delta T/T$ that produce large changes in the observed intensity. These explanations are explained below and represented in Figure 5. A key issue is the size of the corresponding energy reservoir required in each case. To compare these energy requirements, we assume that a typical peak corresponds to an isotropic luminosity, $L_0$, equal to $10^{51}$ ergs s$^{-1}$. The
Fig. 5.—Source geometries related to producing the chaotic time histories that are typical in GRBs. 

a: A shell that emits uniformly over angles the order of $\Gamma^{-1}$ would have an emitting region with a transverse size of $\sim 2\gamma cT$ and would produce a burst with a FRED-like time history. The time history would be the sum of many causally disconnected regions and, therefore, should be very smooth. The chaotic nature of GRBs indicates that the structure of the shell must be broken on a scale much smaller than $\Gamma^{-1}$. We characterize the chaotic time history by the filling factor ($f$), which is related to the maximum fraction of the causally disconnected regions that can be simultaneously active. A typical value of $f$ based on BATSE data is $\sim 5 \times 10^{-3}$. 

b: Only structures the size of $2\gamma cT$ can produce GRB with time structures of $\gamma c\Delta T$. One possibility is that a few small regions have a much larger efficiency for converting the bulk motion into gamma rays. 

c: Perhaps only a fraction $f$ of the shell becomes active. The dotted portions of the shell represent regions that never produce photons but have a similar energy content. This explanation requires the factor of $f^{-1}$ more energy than the scenario of panel $a$, since most of the bulk energy is never converted to gamma rays. This could require so much energy that the concept of a single relativistic shell might have to be abandoned for the GRB phase. However, other scenarios do not require this much energy. 

d: If the shell consists of many small, narrow jets, and the regions of the shell between the jets are voids such that they do not waste bulk energy as in panel $c$, then the burst requires about the same energy than that of panel $a$. 

e: Perhaps $d\Omega$ is small enough, $\sim (\Delta T/\gamma T)^2$, that it can make time structure on the order of $\Delta T$. The required energy for this scenario can be as small as $10^{50}$ ergs. This is the only explanation that can also explain gaps in the time. 

f: Variations in the initial conditions as a function of $\theta$ could spread the causally disconnected regions out along the line of sight to the observer such that the resulting pulses do not make a FRED-like shape even though most of the shell within $d\Omega$ becomes active.
total energy reservoir is required to have

$$E_{\text{tot}} = L_0 \frac{d\Omega}{4\pi} \epsilon^{-1} T_{\text{dur}},$$

(20)

where $T_{\text{dur}}$ is the average duration of the event (say, 30 s) and $\epsilon$ is the fraction of the total energy that is converted into gamma rays (say, $10^{-1}$). If the emission is isotropic, then the total energy is about $E_{\text{iso}} \approx 3 \times 10^{53}$ ergs. Since we cannot observe the energy released at angles much greater than $\sim \Gamma^{-1}$, any model can reduce the required energy reservoir to $(d\Omega/4\pi)E_{\text{iso}}$. Although the energy requirement can be made small, a single shell that emits nearly uniformly over $d\Omega$ (see Fig. 5a) is not consistent with the observed time structure, since a uniform shell always results in a FRED-like shape. The following scenarios can produce bursts with peaks with small $\Delta T/T$:

1. The simplest way to have short, powerful flares is to invoke a central engine. In a central engine, there are multiple releases of energy at the central site on a timescale commensurate with the observed duration of the GRB. Each release makes a peak in the time history. Indeed, small $\Delta T/T$ and gaps are the key kinematic reasons why the GRB phase is likely to be a central engine (Fenimore et al. 1996, 1998).

2. It is possible for a large fraction of the surface of the shell to be active, and the flare region might still be small, because $\epsilon$ or the energy content of the shell varies strongly as a function of position on the shell. That is, the photons $\text{cm}^{-2} \text{s}^{-1}$ at the flare site must be larger than the rest of the shell by a factor of $\sim f^{-1}$ (see Fig. 5b). For example, some photon production rates scale as $\Gamma^{-3}$, so a single region with a $\Gamma$ that is 10 times larger than other regions could produce emission that doubles the overall count rate. This would give the appearance that $f \approx 10^{-3}$. If the energy content of the shell is constant as a function of angle and $\epsilon$ varies, then the required total energy is about $f^{-1}(d\Omega/4\pi)E_{\text{iso}}$. If $\epsilon$ is constant, and the energy content of the shell varies, then the total required energy is $(d\Omega/4\pi)E_{\text{iso}}$.

3. The third explanation is that the local spherical symmetry of the shell is broken into many emitting regions on a scale much smaller than $\Gamma^{-1}$, where only a fraction $f$ of the surface converts its energy into photons. The regions are small enough ($\Gamma c \Delta T$) to make peaks with small $\Delta T/T$ (see Fig. 5c). Here it is assumed that all regions of the shell have equal energy content, but much of it is wasted, because most regions do not produce gamma rays. Thus this scenario requires an energy reservoir of $f^{-1}(d\Omega/4\pi)E_{\text{iso}}$, that is, $f^{-1} \sim 10^3$ more energy than a simple shell. If $d\Omega$ is large, this scenario could require $E_{\text{tot}}$ to be as large as $10^{56}$ ergs. Sari & Piran (1997a) invoke this explanation and the resulting need for more energy to argue against a single relativistic shell. That argument should not be considered particularly strong, since we have several scenarios that do not waste the energy and/or $d\Omega$ could be small.

4. The fourth explanation is that the local spherical symmetry is broken because the outflow of material is in narrow fingers that occupy only a fraction $f$ of the shell’s surface (see Fig. 5d). This is basically the same geometry as the previous case, except those regions that do not emit gamma rays do not have energy content. Since no energy is wasted in non-emitting regions, this scenario requires an energy reservoir of only $(d\Omega/4\pi)E_{\text{iso}}$.—that is, about what one would expect from a simple shell.

5. The fifth explanation is that $d\Omega$ is very small, so small that the emission lasts only $\Delta T_p$ even if the entire $d\Omega$ region is gamma-ray active. This requires $d\Omega$ to be $\sim (\Delta T_p/4T)^2$. The required total energy is correspondingly small: $(\Delta T_p/4T)^2 E_{\text{iso}} = 10^{49}$ ergs. This explanation has one narrow ejecta that needs to go through a gamma-ray emitting phase for each peak in the GRB time history. In the above explanation (4), there are many narrow ejecta, so each one only has to become gamma-ray active only a few times to make a GRB time history with hundreds of peaks.

6. The last explanation attributes the duration of the GRB to the spreading out in time of the emission because of different conditions as a function of $\theta$ (see Fig. 5f). This is what we called a “thick shell with substructure” in Fenimore et al. (1996). It would have been better to call it a “thick emitting region with substructure,” since the actual emitting shell must always be thin (see Fenimore et al. 1996). Effectively, $f$ is unity, but it appears to be small, because the assumptions involved in equation (11) are not valid. Let $t_j$ be the time that the $j$-th section of a shell at $\theta_j$ expands before becoming active. Let $\beta_j$ be the speed of the material at $\theta_j$. The relative time that the emission from two different regions ($i$ and $j$) arrive at our detector is

$$T = T_j - T_i = t_i(1 - \beta_i \cos \theta_i) - t_j(1 - \beta_j \cos \theta_j).$$

(21)

Explanations 3 and 4 above assume that all $t_i$’s and $\beta_i$’s are equal such that we can estimate the size of the shell from its FRED-like appearance, and peaks arrive at different times because they come from different $\theta$’s. Of course, most bursts are not FRED-like. In this explanation, all $\theta$’s become active (so $f = 1$), but the peaks that would have added up to make the FRED-like envelope are scattered in time because of the different $t_i$’s and $\beta_i$’s. To disrupt the FRED-like shape, $T_{f/2}$ must vary by a factor of at least 2 from what they would be if everything was constant with $\theta$. Consider the effect of different times of emission (i.e., different $t_i$’s but similar $\beta_i$’s). The $t_i$’s would have to vary by about a factor of 2 to change $T_{f/2}$ by a factor of 2. Consider constant $t_i$’s but different $\beta_i$’s. To spread out the emitting regions so their distance from us varies by about $T$ requires that $(\beta_j - \beta_i)ct = T$ or that the radius of the shell is

$$ct = \frac{2\Gamma_i^2 \Gamma_j^2}{\Gamma_j^2 - \Gamma_i^2} cT \sim 2\Gamma_i^2 cT.$$  

(22)

The $\Gamma_i$ and $\Gamma_j$ would have to vary by about $\sqrt{2}$. Such differences could arise from different baryon loading as a function of $\theta$. There are two observations that argue against this explanation. First, even the bursts with FRED-like envelopes show low filling factors (Fig. 3, solid squares). Second, if the $\Gamma_j$’s differ by $\sqrt{2}$, then peaks near the end of the burst should have widths that are at least a factor of 2 wider. Bursts normally do not appear to have such a trend. This explanation requires an energy reservoir like a uniform shell: $(d\Omega/4\pi)E_{\text{iso}}$.

In conclusion, both the gamma-ray phase and the X-ray afterglow show similar rapid variability, and both have values of $f \approx 0.1 \Delta T/T$ that are similar. Such variability implies that either GRBs are central engines or the local spherical symmetry is broken on a scale much smaller than $\Gamma^{-1}$. In the last case, the structure can either be due to variations in $\Gamma$ as a function of $\theta$ (perhaps because of small-scale variations in the baryon loading; explanation 6 above) or only a fraction $f$ of the shell becoming active.
(explanations 3, 4, and 5 above). The required total energy is 
\[ \sim f^{-1} E_{\text{iso}}, E_{\text{iso}}, \text{ and } (\Delta T_p/4T)^2 E_{\text{iso}} \] 
for explanations 3, 4, and 5, respectively. Using typical values, the total energy can vary from \(10^{56}\) to \(10^{49}\) ergs. We note that explanation 5 (a jet much narrower than \(\Gamma^{-1}\)) is the only explanation that can also easily explain gaps in the time history.

We thank Cheng Ho and Chuck Dermer for useful comments. Jay Norris provided the time histories analyzed for Figure 3. This work was done under the auspices of the US Department of Energy and was funded in part by the NASA GRO Guest Investigator program.

### REFERENCES

Costa, E., et al. 1997, Nature, 387, 783
Dar, A. 1998, ApJ, 500, L93
Fenimore, E. E., Madras, C. D., & Nayakshin, S. 1996, ApJ, 473, 998
Fenimore, E. E., Ramirez-Ruiz, E., & Sumner, M. C. 1998, in AIP Conf. Proc. 428, Gamma-Ray Bursts: 4th Huntsville Symposium, ed. C. A. Meegan, R. D. Preece, & T. M. Koshut (New York: AIP), 657
Fenimore, E. E., & Sumner, M. C. 1997, in All-Sky X-Ray Observations in the Next Decade, ed. M. Matsuoka, & N. Kawai (Wako: RIKEN), 167
Frail, D., Kulkarni, S. R., Nicastro, L., Feroci, M., & Taylor, G. B. 1997, Nature, 389, 261
Kobayashi, S., Piran, T., & Sari, R. 1997, ApJ, 490, 92
Meegan, C. A., et al. 1996, ApJS, 106, 65
Mészáros, P., & Rees, M. J. 1993, ApJ, 405, 278
———. 1997, ApJ, 476, 232
Mészáros, P., Rees, M. J., & Wijers, R. A. M. J. 1998, ApJ, 499, 301
Metzger, M. R., et al. 1997, Nature, 387, 878
Norris, J. P., et al. 1996, ApJ, 459, 393
Panaitescu, A., & Mészáros, P. 1998, ApJ, 501, 772
Piran, T., & Sari, R. 1997, ApJ, 485, 270
Piro, L., Matt, G., & Ricci, R. 1997, A&AS, 126, 525
Rees, M. J., & Mészáros, P. 1994, ApJ, 430, L93
———. 1998, ApJ, 496, L1
Reichart, D. 1997, ApJ, 485, L57
Sari, R., & Piran, T. 1997a, ApJ, 485, 270
———. 1997b, MNRAS, 287, 110
Sari, R., Piran, T., & Narayan, R. 1998, ApJ, 497, L17
Sumner, M. C., & Fenimore, E. E. 1998, in AIP Conf. Proc. 428, Gamma-Ray Bursts: 4th Huntsville Symposium, ed. C. A. Meegan, R. D. Preece, & T. M. Koshut (New York: AIP), 765
Tavani, M. 1997, ApJ, 483, L87
Waxman, E., Kulkarni, S. R., & Frail, D. A. 1998, ApJ, 497, 288
Wijers, R. A. M. J., Rees, M. J., & Mészáros, P. 1997, MNRAS, 288, L51
Yoshida, A., et al. 1998, in AIP Conf. Proc. 428, Gamma-Ray Bursts: 4th Huntsville Symposium, ed. C. A. Meegan, R. D. Preece, & T. M. Koshut (New York: AIP), 441