THE ASTROPHYSICAL JOURNAL, 563:582–591, 2001 December 20

THE ENERGY DISTRIBUTION OF GAMMA-RAY BURSTS

DAVID L. BAND[1]
X-2, Los Alamos National Laboratory, Los Alamos, NM 87545
Received 2001 May 15; accepted 2001 August 21

ABSTRACT

The distribution of the apparent total energy emitted by a gamma-ray burst reflects not only the distribution of the energy actually released by the burst engine, but also the distribution of beaming angles. Using the observed energy fluences, the detection thresholds, and burst redshifts for three burst samples, I calculate the best-fit parameters for lognormal and power-law distributions of the apparent total energy. Two of the samples include a small number of bursts with spectroscopic redshifts, while the third sample has 220 bursts with redshifts determined by the proposed variability-luminosity correlation. I find different sets of parameter values for the three burst samples. The Bayesian odds ratio cannot distinguish between the two model distribution functions for the two smaller burst samples with spectroscopic redshifts, but does favor the lognormal distribution for the larger sample with variability-derived redshifts. The data do not rule out a distribution with a low-energy tail that is currently unobservable. I find that neglecting the burst detection threshold biases the fitted distribution to be narrower with a higher average value than the true distribution; this demonstrates the importance of determining and reporting the effective detection threshold for bursts in a sample.

Subject headings: gamma rays: bursts — methods: statistical

1. INTRODUCTION

The growing number of gamma-ray bursts with redshifts has not only established that most, if not all, bursts are at cosmological distances, and that up to 10^{54} ergs are radiated by a burst, but permits us to determine relevant intrinsic physical distributions. Since we sample a burst's radiation pattern at only one point, the observed energy fluence can only be related to the energy flux emitted in our direction; this energy flux can be expressed as the total energy the burst would have emitted if it radiated isotropically. Here I consider the distribution of this apparent total energy.

The distribution of the apparent total energy is a convolution of the distribution of the actual energy emitted and the distribution of the angle into which the emission is beamed, both quantities of crucial importance in understanding the physics of the progenitor, the frequency of burst occurrence, and the impact of a burst on its environment. Frail et al. (2001) have recently determined the beaming angles for a burst sample by modeling the late-time breaks in the temporal decay of the afterglows; accounting for the beaming angle provides the actual total energy, which they found is clustered around $E \sim 5 \times 10^{50}$ ergs. If the analysis of Frail et al. is indeed correct (there are competing models for the evolution of the afterglow), then methods similar to those I develop here will be necessary to properly determine the beaming angle and total energy distributions. Here I do not attempt to disentangle these two distributions.

This study extends the work of Jimenez, Band, & Piran (2001), which considered lognormal distributions for the apparent total gamma-ray energy, the peak gamma-ray luminosity, and the total X-ray afterglow energy. Jimenez et al. expanded the database of bursts with spectroscopic redshifts by adding bursts for which a redshift probability distribution could be inferred from the host galaxy brightness. Currently, the bursts with only redshift probability distributions do not augment the relevant burst database sufficiently to warrant their inclusion in my study, although in the future it may be advantageous to include these bursts if the determination of spectroscopic redshifts does not keep pace with the detection of bursts and their host galaxies.

Recently correlations have been proposed between burst properties and their peak luminosities. Norris, Marani, & Bonnell (2000) proposed that more luminous bursts have smaller time lags between energy channels, while Fenimore & Ramirez-Ruiz (2001) and Reichart et al. (2001) reported that the light curves of more luminous bursts were more variable; Schaefer, Deng, & Band (2001) found that applying both correlations to the same burst sample resulted in consistent burst luminosities. The redshift is determined from the derived luminosities and the observed peak fluxes. Thus, these correlations can give us a large burst sample with redshifts from which the energy distribution can be determined.

The methodology presented here demonstrates that accurate estimation of physical parameters, such as those for the burst energy distribution, requires well-defined samples. Such samples are not yet available. While this study does characterize the apparent energy distribution, it is more significantly an argument for characterizing and reporting the threshold for including a burst in a sample. Such a threshold may result from the detection of the burst, the localization of the burst or its afterglow, or the determination of the redshift. Consequently, clear criteria must be established and reported for follow-up observations after a burst.

In § 2 I discuss the methodology for finding the best parameter values for a given functional form of the distribution, and for comparing different functional forms. I also evaluate the sensitivity to small burst samples, and demonstrate the importance of considering the detection threshold in determining the distributions. The energy distributions resulting from two small burst samples with spectroscopic

[1] Current address: GLAST SSC, Code 661, NASA Goddard Space Flight Center, Greenbelt, MD 20771; dband@lheapop.gsfc.nasa.gov.

redshifts and from a large sample with redshifts from the variability-luminosity correlation are described in § 3. Finally, § 4 summarizes my conclusions.

2. METHODOLOGY

2.1. *The Likelihood Function*

I begin with an assumed distribution $p(E \mid \boldsymbol{a}_j, M_j, I)$, where E is the apparent total burst energy, $\boldsymbol{a}_j$ is the set of parameters that characterize the jth model distribution function represented by M_j, and I specifies general assumptions about the distribution function. I use $p(a \mid b)$ to mean the probability of a given b. Below I present the distribution functions used in this study. A distribution is assumed to be universal and not a function of redshift (i.e., no evolution), or of properties of the host galaxy, burst, etc. Of course, once a sufficiently large sample is available, the energy distributions for burst subsets can be investigated; alternatively, the possible dependencies of these distributions on other physical parameters can be modeled using the entire data set. I use normalized distributions, since the burst rate is not of interest here. The observed energy fluence F and the burst energy E are related by $F = E(1+z)/4\pi D_L(z)^2 = EC(z)$, where $D_L(z)$ is the luminosity distance. It is in calculating the luminosity distance that the burst redshift and a cosmological model are required. Here I assume $H_0 = 65$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

This energy distribution is converted into $p(F \mid \boldsymbol{a}_j, z, M_j, I)$, the probability of obtaining the energy fluence F given the parameters $\boldsymbol{a}_j$ for model M_j, the burst redshift z, and other assumptions I (e.g., the choice of cosmological model). However, the observed fluences are not drawn from this probability distribution but from $p(F \mid F_T, \boldsymbol{a}_j, z, M_j, I)$, the normalized distribution that is truncated below F_T, the minimum fluence at which that particular burst would have been included in the burst sample.

I now form the likelihood

$$L_j = \prod_{i=1}^{N} p(F_i \mid F_{T,i}, \boldsymbol{a}_j, z_i, M_j, I) \, , \tag{1}$$

where the ith burst has energy fluence F_i, fluence threshold $F_{T,i}$, and redshift z_i.

2.2. *Parameter Estimation*

In the "frequentist" framework, best-fit parameters are typically found by maximizing L_j.

The Bayesian analysis is based on the posterior probability for the parameters, $p(\boldsymbol{a}_j \mid D, M_j, I)$; here D is the set of observed fluences and fluence thresholds. Note that $L_j = p(D \mid \boldsymbol{a}_j, M_j, I)$. Thus, an estimate of the parameters is

$$\begin{aligned}\langle \boldsymbol{a}_j \rangle &= \int d\boldsymbol{a}_j \, \boldsymbol{a}_j \, p(\boldsymbol{a}_j \mid D, M_j, I) \\ &= \frac{\int d\boldsymbol{a}_j \, \boldsymbol{a}_j \, p(D \mid \boldsymbol{a}_j, M_j, I) p(\boldsymbol{a}_j \mid M_j, I)}{\int d\boldsymbol{a}_j \, p(D \mid \boldsymbol{a}_j, M_j, I) p(\boldsymbol{a}_j \mid M_j, I)} \\ &= \frac{\int d\boldsymbol{a}_j \, \boldsymbol{a}_j \prod_{i=1}^{N} p(F_i \mid F_{T,i}, \boldsymbol{a}_j, z_i, M_j, I) p(\boldsymbol{a}_j \mid M_j, I)}{\int d\boldsymbol{a}_j \prod_{i=1}^{N} p(F_i \mid F_{T,i}, \boldsymbol{a}_j, z_i, M_j, I) p(\boldsymbol{a}_j \mid M_j, I)} \, , \end{aligned} \tag{2}$$

where $p(\boldsymbol{a}_j \mid M_j, I)$ is the prior for the parameters $\boldsymbol{a}_j$. If $\Lambda_j = p(D \mid \boldsymbol{a}_j, M_j, I) p(\boldsymbol{a}_j \mid M_j, I)$ is sharply peaked, then the expectation value of the parameters occurs at the peak of Λ_j. Note that Λ_j is the likelihood in equation (1) times the priors for the parameters, and is proportional to the posterior probability $p(\boldsymbol{a}_j \mid D, M_j, I)$. The posterior probability is also used to determine the acceptable parameter range, which is often more meaningful than the "best" parameters values.

The choice of priors can be regarded as a judgement as to the "natural variables" for the particular functional form; the priors are constant for these variables. Here the burst energies vary over a number of decades, and thus I assume that the priors for average or cutoff energies are logarithmic. Consequently, maximizing Λ_j is equivalent to maximizing the likelihood in terms of the logarithm of the energy parameters. This is the methodology I use here; I do not attempt to integrate the integrals in equation (2).

2.3. *The Cumulative Probability*

For each burst the cumulative probability is

$$P(F_i \mid F_{T,i}, \boldsymbol{a}_j, z_i, M_j, I) = \int_{F_i}^{\infty} p(F \mid F_{T,i}, \boldsymbol{a}_j, z_i, M_j, I) dF \, . \tag{3}$$

If the assumed energy distribution function is an acceptable characterization of the observations (which would be the case if the model M_j is correct) and all the assumptions are valid (e.g., the cosmological model is correct), then the cumulative probabilities $P(F_i)$ should be uniformly distributed between 0 and 1, and have an average value of $\langle P(F_i) \rangle = \frac{1}{2} \pm (12N)^{-1/2}$ for N bursts in the sample.

2.4. *Model Comparison*

The Bayesian framework provides a clear prescription for comparing models through the odds ratio. Let $p(M_j \mid D, I)$ be the posterior probability for the jth model M_j given the data D. Then the odds ratio comparing the jth and kth models is $O_{jk} = p(M_j \mid D, I)/p(M_k \mid D, I)$. However, by Bayes' theorem $p(M_j \mid D, I) \propto p(M_j \mid I) p(D \mid M_j, I)$, where the normalizing factor is independent of M_j, and thus cancels in forming the odds ratio. I assume that no model is favored a priori; therefore, the "priors" $p(M_j \mid I)$ are the same for all M_j, and cancel in forming the odds ratio. Here D consists of the observed energy fluences (and the detection thresholds). Thus, $p(M_j \mid D, I) \propto \prod_i p(F_i \mid F_{T,i}, z_i, M_j, I)$. However, we begin with the more fundamental probabilities $p(F_i \mid F_{T,i}, z_i, M_j, \boldsymbol{a}_j, I)$, which are functions of the model parameters $\boldsymbol{a}_j$. In determining the preferred model we are uninterested in the specific model parameters, and therefore we "marginalize" over the parameters $\boldsymbol{a}_j$: $p(F_i \mid F_{T,i}, z_i, M_j, I) = \int d\boldsymbol{a}_j \, p(\boldsymbol{a}_j \mid M_j, I) \prod_i p(F_i \mid F_{T,i}, z_i, M_j, \boldsymbol{a}_j, I)$, where $p(\boldsymbol{a}_j \mid M_j, I)$ is the prior for the parameters of the jth model. Thus,

$$O_{jk} = \frac{\int d\boldsymbol{a}_j \, p(\boldsymbol{a}_j \mid M_j, I) \prod_i^N p(F_i \mid F_{T,i}, z_i, M_j, \boldsymbol{a}_j, I)}{\int d\boldsymbol{a}_k \, p(\boldsymbol{a}_k \mid M_k, I) \prod_i^N p(F_i \mid F_{T,i}, z_i, M_k, \boldsymbol{a}_k, I)} \, . \tag{4}$$

It will be noted that the odds ratio is the ratio of the likelihoods for each model (eq. [1]), marginalized over the model parameters. Derived from a different philosophy of inference, the frequentist likelihood ratio test is the same as the Bayesian odds ratio with the prior for the parameters set equal to a delta function at the parameter values that maximize the likelihood. In the Bayesian approach such a prior is the result of circular reasoning; the prior should be based on information available before the new data were

obtained. Below I present priors defined in terms of the expected parameter ranges, in which case the integrals in equation (4) must be integrated. I present values of both the frequentist likelihood ratio test and the Bayesian odds ratio. Different conclusions may result from these tests, as is indeed the case here.

As mentioned above, the average of the cumulative probability $\langle P(F_i)\rangle$ should be $\frac{1}{2}$ (within a quantifiable uncertainty) if the distribution function describes the observations satisfactorily; this statistic for model M_j is a measure of the acceptability of that model. Furthermore, a comparison of the values for different models compares the relative merits of these models; of course, this comparison should account for the expected uncertainty in the value of this statistic. A model that is consistent with the data by this statistic would be favored over one that is not. However, the statistic's ability to discriminate between models is not addressed by its derivation.

2.5. *Distribution Functions*

2.5.1. *Lognormal Distribution*

I assume that E has a lognormal distribution

$$p(E\,|\,E_0,\,\sigma)d(\ln E) = \frac{1}{\sqrt{2\pi}\sigma} \times \exp\left[-\frac{(\ln E_0 - \ln E)^2}{2\sigma^2}\right] d(\ln E)\ . \tag{5}$$

Thus, the fluence F also has a lognormal distribution. Note that σ is a width in logarithmic space, and the linear change of variables from E to F does not affect this width. As discussed above, we need to consider the range over which the fluence could actually have been observed, i.e., for fluences above the threshold F_T. The resulting normalized fluence probability distribution is

$$p_{\rm obs}(F\,|\,F_T,\,E_0,\,\sigma,\,z)d(\ln F) = \frac{1/(\sqrt{2\pi}\sigma)\exp\left(-\{\ln[E_0\,C(z)] - \ln F\}^2/2\sigma^2\right)\theta(F - F_T)d(\ln F)}{(1/2)[1 + {\rm erf}(\{\ln[E_0\,C(z)] - \ln(F_T)\}/\sqrt{2}\sigma)]}\ , \tag{6}$$

where $\theta(x)$ is the Heaviside function (1 above $x = 0$, and 0 below), and $C(z) = (1 + z)/4\pi D_L(z)^2$ converts energies in the burst's frame to fluences in our frame; the denominator results from integrating over the numerator from F_T to infinity.

For the Bayesian formulation we also need the priors for the model parameters. There is no reason to favor one energy over another over many energy decades, and thus I assume that the prior is constant in logarithmic space: $p(E_0)dE_0 = (E_0 \ln[E_u/E_l])^{-1}dE_0 = (\ln[E_u/E_l])^{-1}d\ln E_0 = (\log_{10}[E_u/E_l])^{-1}d\log_{10} E_0$, where E_u and E_l are the upper and lower limits, respectively, of the permitted range (because of the logarithmic dependence, the result is not very sensitive to the precise values). I use $E_l = 10^{51}$ ergs and $E_u = 10^{54}$ ergs. Similarly, I have no a priori information about σ, and therefore assign it a uniform prior between 0 and 5. Note that the distribution function is explicitly a function of ln E_0 and σ, and I have chosen priors for which these are the natural parameters for this functional form.

2.5.2. *Single-Component Power-Law Distribution*

I assume that the underlying energy probability distribution is

$$p(E\,|\,E_1,\,E_2,\,\alpha)dE = \frac{(1-\alpha)E_2^{\alpha-1}}{1-(E_1/E_2)^{1-\alpha}}\,E^{-\alpha}dE;\quad E_1 \le E \le E_2,\quad \alpha \ne 1$$

$$p(E\,|\,E_1,\,E_2,\,\alpha)dE = \frac{E^{-1}}{\ln(E_2/E_1)}\,dE\ ;\quad E_1 \le E \le E_2,\quad \alpha = 1\ . \tag{7}$$

If E_1 is extended to 0 or E_2 to infinity, then we must restrict α to be less than or greater than 1, respectively. The expected fluence probability distribution, accounting for the fluence threshold F_T, is

$$p(F\,|\,F_T,\,E_1,\,E_2,\,\alpha)dF = \frac{(1-\alpha)[E_2\,C(z)]^{\alpha-1}\theta(F-F_T)}{1-\{\max[E_1,\,F_T/C(z)]/E_2\}^{1-\alpha}}\,F^{-\alpha}dF\ ;$$

$$E_1 \le F/C(z) \le E_2,\quad \alpha \ne 1$$

$$p(F\,|\,F_T,\,E_1,\,E_2,\,\alpha)dF = \frac{F^{-1}\theta(F-F_T)}{\ln\{E_2/\max[E_1,\,F_T/C(z)]\}}\,dF\ ;$$

$$E_1 \le F/C(z) \le E_2,\quad \alpha = 1\ , \tag{8}$$

where again $C(z) = (1 + z)/4\pi D_L(z)^2$ converts burst energies to fluences.

As with E_0 for the lognormal distribution, I have no reason to prefer any value of the energy limits E_1 and E_2 over many energy decades, and therefore again I use logarithmic priors for these two energies. For definiteness, I assume that E_1 can have a value between 10^{49} and 10^{52} ergs, and E_2 between 10^{52} and 10^{55} ergs. The spectral index is assumed to have a uniform prior between -2.5 and 2.5.

2.6. *Data*

The methodology discussed above requires the energy fluence F, the fluence threshold F_T, and the redshift z for each burst. I consider three samples. The B9 sample consists of nine bursts with BATSE data and spectroscopic redshifts. The fluences were calculated by fitting the BATSE spectrum accumulated over the entire burst with the "GRB" function (Band et al. 1993), and then integrating the resulting fit over the 20–2000 keV energy range in the burst's rest frame and over the time during which the spectrum was accumulated. The resulting fits are presented in Jimenez et al. (2001). There are some bursts for which the high-energy power law in the GRB function has an index $\beta < -2$ (where $N \propto E^{\beta}$), and thus the fluence depends crucially on the high-energy cutoff (or roll-off) that must exist for a finite fluence but that could not be determined from the BATSE data. The spectroscopic redshifts are taken from Frail et al. (2001).

The limiting fluence F_T is more difficult to determine. The bursts in our sample must have been intense enough to be first detected and then localized properly. In addition, a decision was made to attempt to observe the afterglow and thus determine the redshift. These threshold quantities have generally not been reported. Note also that detectors almost never trigger on the fluence, but usually trigger on the peak count rate sampled over a time bin of a specified duration.

TABLE 1

THE BATSE GAMMA-RAY BURSTS SAMPLE

Burst	F_{obs}[a] (erg cm^{-2})	F_{burst}[b] (erg cm^{-2})	z_{obs}	C_{max}/C_{min}	E_{obs}[c] (10^{51} ergs)	E_{burst}[d] (10^{51} ergs)
970508.....	3.18×10^{-6}	2.59×10^{-6}	0.835	3.3[e]	6.734	5.482
970828.....	9.57×10^{-5}	7.88×10^{-5}	0.958	20[e]	267.1	219.3
971214.....	9.44×10^{-6}	7.59×10^{-6}	3.412	7.32[f]	261.3	210.7
980703.....	2.26×10^{-5}	2.13×10^{-5}	0.966	3.08[f]	63.94	60.18
990123.....	2.68×10^{-4}	1.93×10^{-4}	1.600	80.1[f]	1996	1438
990506.....	1.94×10^{-4}	1.69×10^{-4}	1.2	50[e]	838.8	854.0
990510.....	2.26×10^{-5}	2.32×10^{-5}	1.619	19.3[f]	172.0	176.8
991216.....	1.93×10^{-4}	1.70×10^{-4}	1.02	144.[f]	611.1	534.1
000131.....	4.18×10^{-5}	2.71×10^{-5}	4.5	3[e]	1791	1159

[a] Fluence over 20–2000 keV in the observer's frame.
[b] Fluence over 20–2000 keV in the burst's frame.
[c] Gamma-ray energy over 20–2000 keV in the observer's frame, assuming isotropic emission, $H_0 = 65$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$.
[d] Gamma-ray energy over 20–2000 keV in the burst's frame.
[e] Estimated from light curve.
[f] From the on-line BATSE catalog.

Here I assume that the ratio of the observed to fluence threshold, F/F_T, is the same as the ratio of the observed to threshold peak count rate, C_{max}/C_{min}, for the BATSE data. These thresholds are most likely underestimates of the true thresholds. The C_{max}/C_{min} ratio is a standard part of the BATSE catalog;[2] however, in some cases the BATSE team did not calculate this quantity because of data gaps, in which case I estimated this ratio from the light curves. This sample is described by Table 1.

The C17 sample consists of the 17 bursts with spectroscopic redshifts and fluences in Frail et al. (2001). This sample is basically a superset of the B9 sample with an additional BATSE burst for which there are no spectra, and bursts observed by *BeppoSAX* and *Ulysses*. For those bursts without reported detection thresholds I use a fluence threshold of 10^{-6} erg cm^{-2}.

Finally, the F220 sample uses the 220 bursts in Fenimore & Ramirez-Ruiz (2001), with redshifts determined by the variability-luminosity correlation. This sample was selected to have a peak count rate accumulated on the 256 ms timescale of greater than 1.5 counts s^{-1}, which provides a well-defined detection threshold. The fluences were taken from the BATSE catalog without any k-corrections. Note that Bloom, Frail, & Sari (2001) find that the k-correction for the 20–2000 keV energy range is of order unity.

It should be noted that for the first two samples the redshifts are reliable but the detection threshold is very uncertain. Even when the detection threshold is known for the gamma-ray portion of the burst, the effective threshold for optical follow-up observations has not been reported. On the other hand, the detection threshold for the third sample is known, but the validity of the variability-luminosity correlation is still not well established, and the uncertainty in the resulting redshifts is not considered.

2.7. *Simulations*

To determine the sensitivity of the methodology to the number of bursts and to demonstrate the importance of considering the fluence threshold, I ran a series of simulations. For each simulation I first created between 100 and 500 simulated databases, to which I then applied the methodology described above to determine the parameters of their energy distribution. For some simulations I found the parameters with the fluence thresholds used in creating the database or with much smaller thresholds, effectively comparing parameter estimation with and without considering the detection thresholds.

For each simulated burst I needed a redshift, a burst energy, and a fluence threshold. The redshifts were drawn from a distribution that is qualitatively similar to the proposed cosmic star formation rate (e.g., rising steeply with redshift to $z \sim 1.5$ and flat for higher redshifts). The fluence threshold was drawn from a uniform logarithmic distribution over one decade. The burst energy was drawn from a lognormal distribution with specified central energy E_0 and

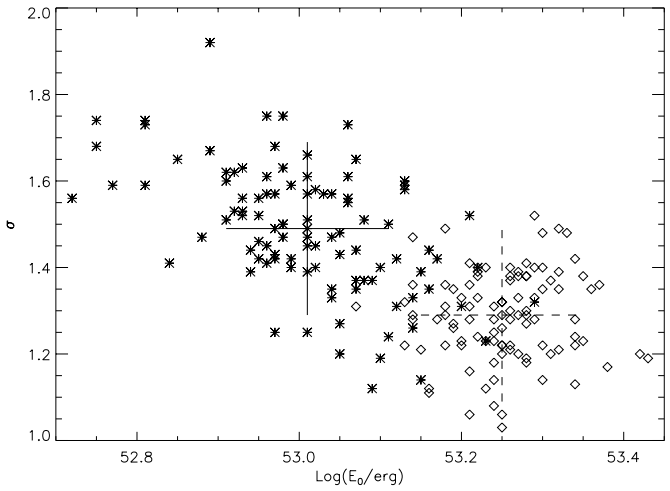

FIG. 1.—Best-fit parameter values including (*asterisks*) or neglecting (*diamonds*) the fluence threshold in calculating the likelihood for 100 simulated databases with 80 bursts each. The model lognormal energy distribution had a central energy of $E_0 = 10^{53}$ ergs and a logarithmic width of $\sigma = 1.5$. The median values for the fits including (*large solid cross*) or neglecting (*large dashed cross*) the fluence threshold are indicated. The bursts were drawn from a redshift distribution similar to that of star formation, and the fluence threshold was between 10^{-6} and 10^{-5} erg cm^{-2}.

[2] BATSE current catalog (2001) is available at: http://www.batse.msfc.nasa.gov/batse/grb/catalog/current/.

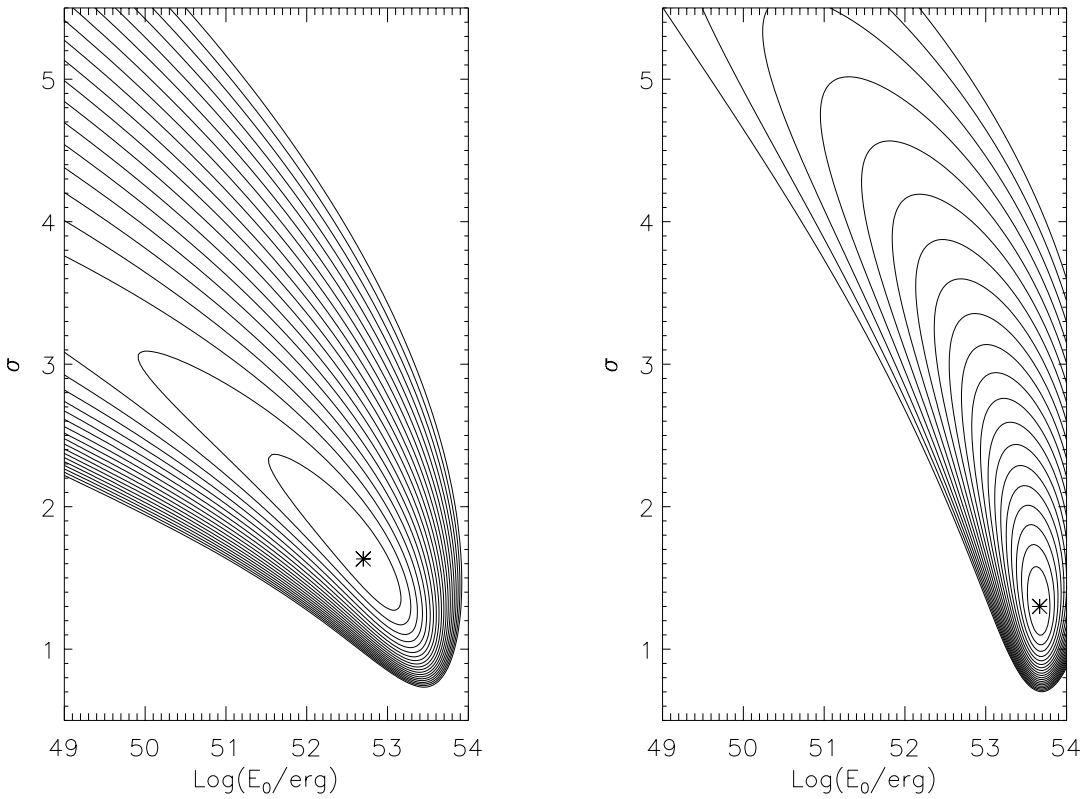

FIG. 2.—Comparison of the likelihood contours for a sample analyzed with the correct fluence thresholds (*left*) and thresholds a factor of 10 smaller (*right*). The sample of nine bursts was drawn from a lognormal distribution with $E_0 = 10^{53}$ ergs and $\sigma = 1.5$, a redshift distribution similar to star formation, and a uniform fluence threshold between 10^{-5} and 10^{-4} erg cm^{-2}. The contours are spaced by $\Delta \log_{10} L = 0.1$ starting from the maximum value.

logarithmic width σ, as long as the resulting fluence was greater than the fluence threshold.

Figure 1 shows the importance of considering the fluence threshold in calculating the likelihood. As can be seen, the best-fitted parameter values cluster around the input central energy $E_0 = 10^{53}$ ergs and logarithmic width $\sigma = 1.5$ when the fluence threshold is considered (Fig. 1, *asterisks*), but cluster around a higher energy and narrower width when the fluence threshold is neglected (*diamonds*). The fluence threshold removes low-energy bursts, resulting in a narrower apparent distribution, which is shifted to higher energy. Each of the 100 simulated data sets had 80 bursts, and a fluence threshold between 10^{-6} and 10^{-5} erg cm^{-2}. Figure 2 shows that the likelihood contours for a sample analyzed with the correct fluence thresholds (*left panel*) and thresholds a factor of 10 smaller (*right panel*) differ significantly.

Table 2 gives the width of the distributions of the parameters of the lognormal distribution, log E_0 and σ, for databases with 9, 20, 40, and 80 bursts. As expected, the distributions become narrower as the number of bursts increases. As can be seen, a database with 40 bursts should give satisfactory best-fit parameter values.

TABLE 2

WIDTH OF PARAMETER DISTRIBUTION

Number of Bursts in Sample	Width[a] of log E_0 Distribution	Width[a] of σ Distribution
9	0.38	0.47
20	0.40	0.52
40	0.20	0.31
80	0.13	0.19

NOTE.—In these simulations, 100 samples were constructed with the indicated number of bursts per sample. The burst energies were drawn from a lognormal distribution with central energy $E_0 = 10^{53}$ ergs and logarithmic width $\sigma = 1.5$. The redshift distribution qualitatively mimics the cosmic star formation rate, and the fluence threshold was between 10^{-6} and 10^{-5} erg cm^{-2}. The best-fit parameters were found by maximizing the likelihood.

[a] The width given is the range within which half the simulated bursts fell.

3. RESULTS

As can be seen from Table 3, the parameter values at the likelihood maximum for the two distribution functions differ for the three burst samples, and the 90% confidence ranges from one burst sample do not always include the parameters from the other samples. Note that the lower energy cutoff E_1 for the simple power-law distribution is not fitted. The fits are insensitive to E_1 values smaller than the lowest energy for which any burst in the sample would have been detected, and E_1 cannot be greater than the smallest observed burst energy; the difference between these two limits is very small. Figures 3–8 show the likelihood contours for the fitted parameters. A ridge of somewhat lower likelihood values curves toward lower values of E_0 and higher values of σ for the lognormal distributions. Thus, the data do not strongly exclude a broader distribution that includes lower energy bursts that are not detected because their fluence is below the threshold. Consequently, the 90% confidence bounds for the parameters are quite broad. For

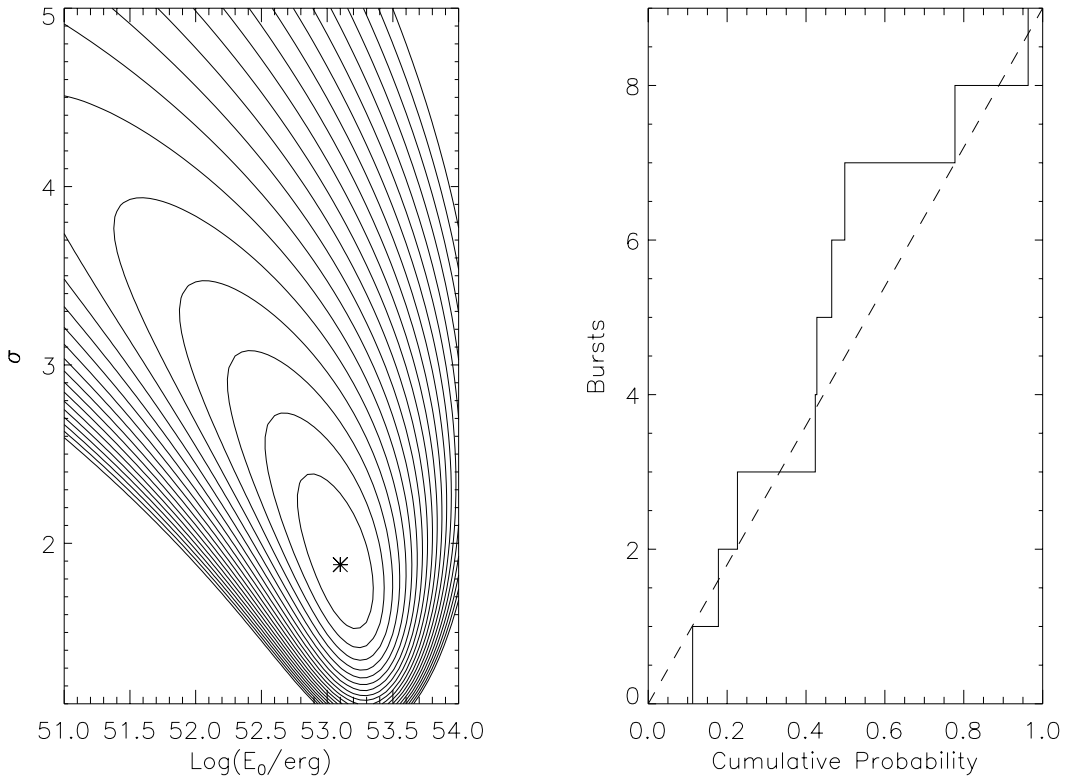

FIG. 3.—Contour plot of the likelihood for the lognormal energy distribution for the B9 sample (*left panel*). The parameters are the central energy E_0 and the logarithmic width σ. The asterisk indicates the location of the maximum likelihood, while contours are spaced by $\Delta\log_{10} L = 0.1$ starting from the maximum value. Cumulative distribution of the cumulative probability for each burst assuming their energies are drawn from the best-fit lognormal energy distribution for the B9 sample (*right panel*). The dashed line shows the expected distribution.

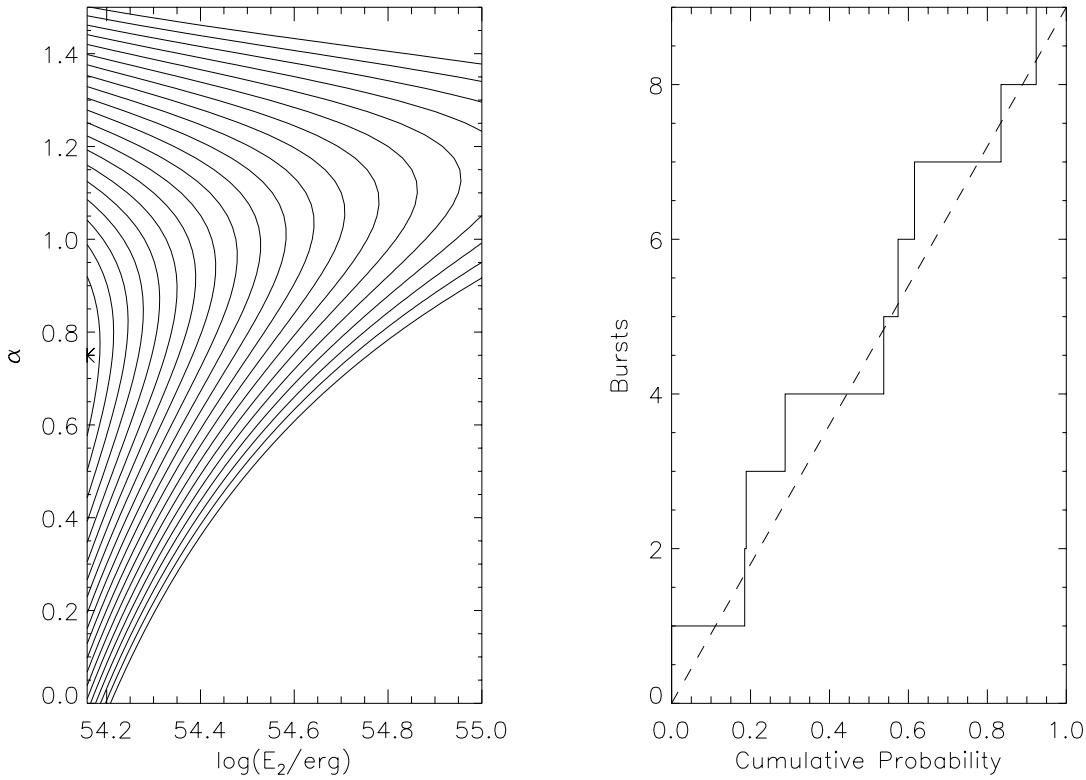

FIG. 4.—Same as Fig. 3, but for a simple power-law energy distribution for the B9 sample. The parameters are the upper cutoff energy E_2 and the power-law index α.

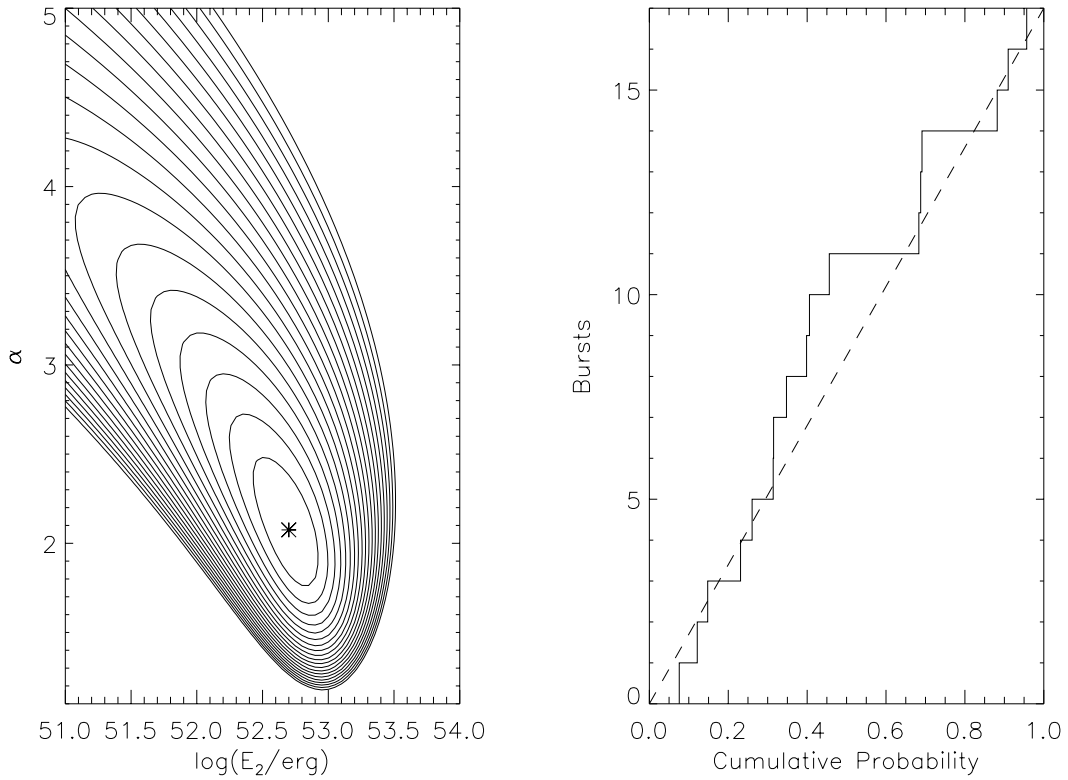

FIG. 5.—Same as Fig. 3, but for the lognormal distribution and the C17 sample. Note that the parameter ranges plotted differ from Fig. 3 to focus on the region of large likelihood values.

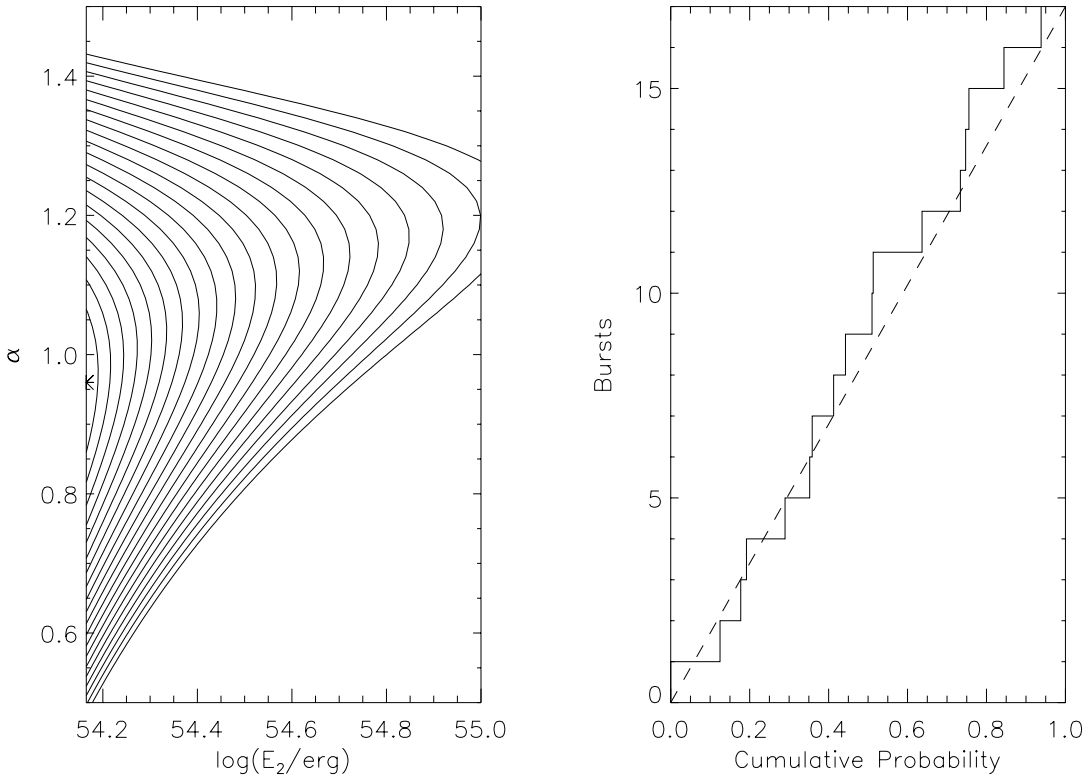

FIG. 6.—Same as Fig. 4, but for the power-law distribution and the C17 sample. Note that the parameter ranges plotted differ from Fig. 4 to focus on the region of large likelihood values.

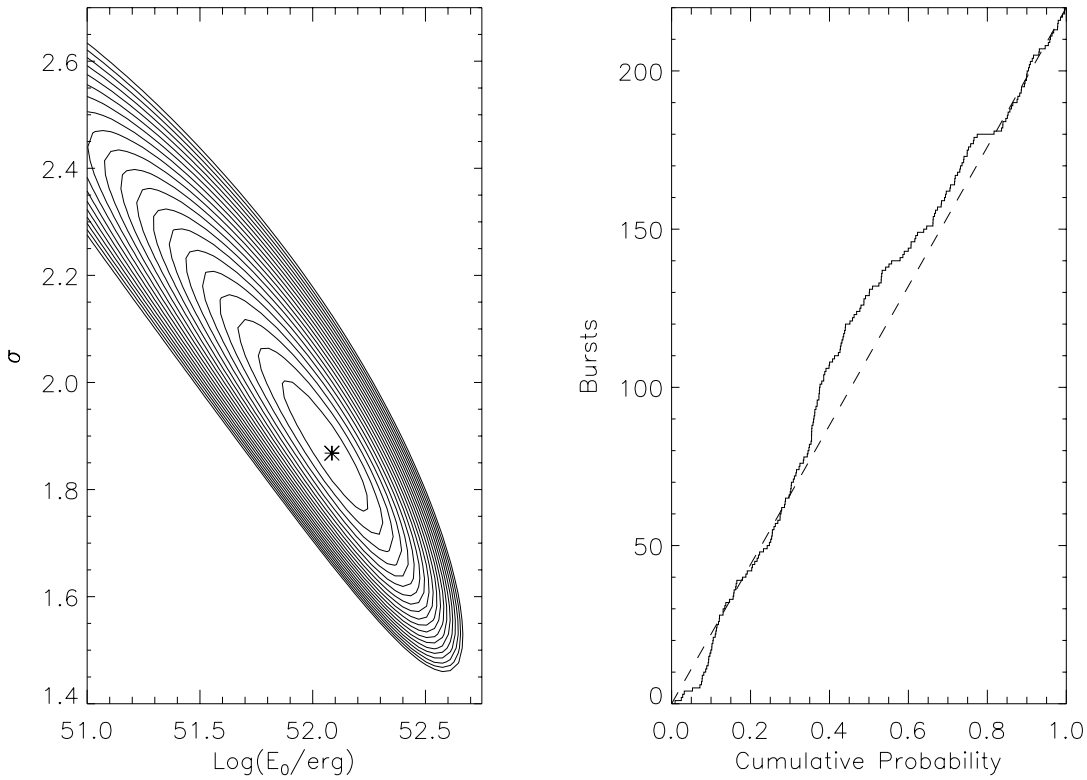

FIG. 7.—Same as Figs. 3 and 5, but for the lognormal distribution and the F220 sample. Note that the parameter ranges plotted differ from Figs. 3 and 5 to focus on the region of large likelihood values.

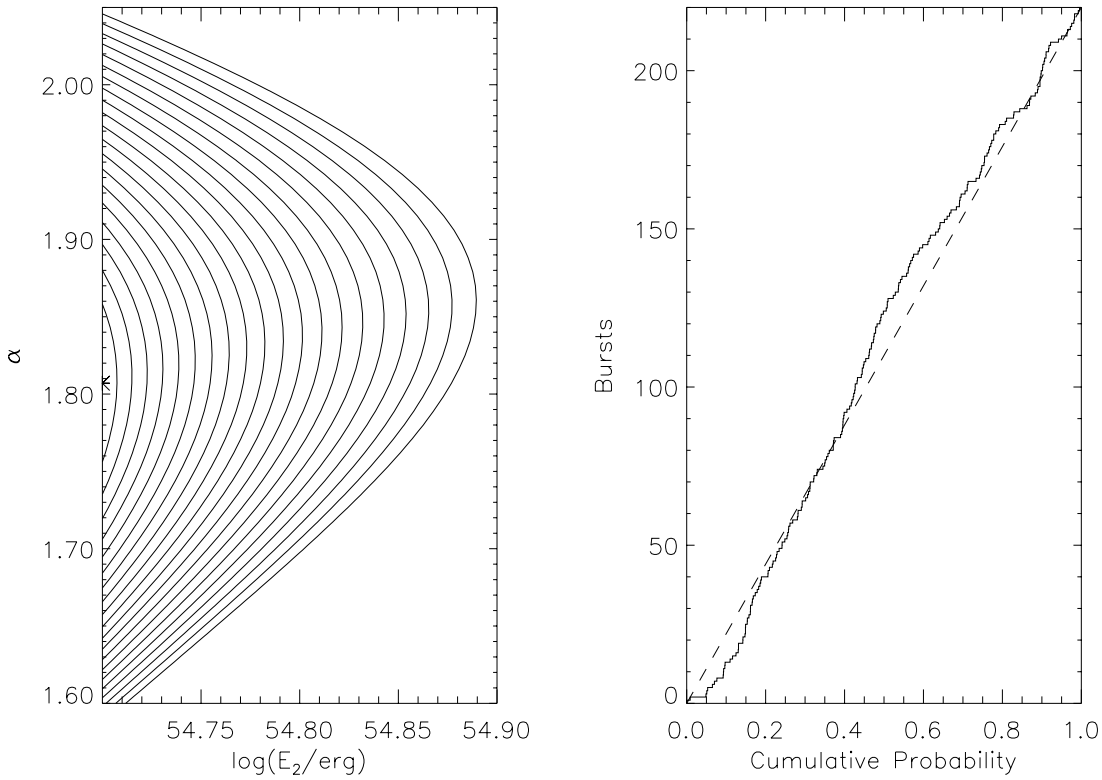

FIG. 8.—Same as Figs. 4 and 6, but for the power-law distribution and the F220 sample. Note that the parameter ranges plotted differ from Figs. 4 and 6 to focus on the region of large likelihood values.

TABLE 3

COMPARISON OF MODEL DISTRIBUTIONS

Quantity	B9[a]	C17[b]	F220[c]
E_0[d]	125.9	52.48	11.79
	1.6–315	1.6–100	2.–23.4
σ[e]	1.87	2.06	1.88
	1.5–4.5	1.7–4.15	1.65–2.3
E_1[f]	1.6	0.55	0.12
E_2[g]	1440	1460	5000
	1440–3550	1460–3350	5000–6100
α[h]	0.74	0.96	1.81
	0.4–1.2	0.75–1.25	1.7–1.94
$\langle p \rangle_{\ln}$[i]	0.4525	0.4638	0.4753
$\langle p \rangle_{\rm pl}$[j]	0.4608	0.4723	0.4892
$\sigma_{\langle p \rangle}$[k]	0.0962	0.0700	0.0195
Likelihood ratio test[l]	4.29×10^{-2}	5.61×10^{-2}	4.38×10^{2}
Odds ratio[m]	1.19	1.65	9.92×10^{3}

[a] Sample of nine BATSE bursts with spectroscopic redshifts and fitted spectra (Table 1).

[b] Sample of 17 bursts with spectroscopic redshifts (Frail et al. 2001).

[c] Sample of 220 bursts with redshifts derived from variability redshifts (Fenimore & Ramirez-Ruiz 2001).

[d] The central energy of the lognormal distribution, in units of 10^{51} ergs. The following line in the table provides the 90% confidence range.

[e] The logarithmic width (in units of the energy's natural logarithm) for the lognormal distribution. The following line provides the 90% confidence range.

[f] The low-energy cutoff of the power-law distribution, in units of 10^{51} ergs. This energy has been set to the lowest threshold energy for the sample.

[g] The high-energy cutoff of the power-law distribution, in units of 10^{51} ergs. The following line provides the 90% confidence range.

[h] The power-law index of the power-law distribution, $p(E) \propto E^{-\alpha}$. The following line provides the 90% confidence range.

[i] Average of the cumulative probabilities for the lognormal distribution; $\frac{1}{2}$ is expected.

[j] Average of the cumulative probabilities for the power-law distribution; $\frac{1}{2}$ is expected.

[k] The standard deviation $[12N]^{-1/2}$ of the average of the cumulative probabilities for a sample of N bursts.

[l] Likelihood ratio test comparing the lognormal and power-law distributions. A value greater than 1 favors the lognormal distribution.

[m] Odds ratio comparing the lognormal to power-law distributions. A value greater than 1 favors the lognormal distribution.

the power-law distributions, the favored high-energy cutoff is at the highest observed energy in the burst sample; if the bursts truly have a power-law distribution, this cutoff is most likely somewhat greater.

Table 3 also provides the average of the cumulative probabilities for each sample and distribution. As can be seen, these averages are within $\sim 1\ \sigma$ of the expected value of $\frac{1}{2}$, indicating that the distributions are both acceptable descriptions of the data. Comparing the values of this statistic does not discriminate between distributions; however, this statistic was created as a goodness-of-fit measure, and thus a comparison of its values for different models is not necessarily very powerful in discriminating between these models. The actual distributions of these cumulative probabilities are also presented by Figures 3–8.

Finally, Table 3 presents both the frequentist likelihood ratio test and the Bayesian odds ratios comparing the lognormal to power-law distributions. As noted above, the likelihood ratio test is quantitatively the same as the Bayesian odds ratio using delta function priors set to the parameters that maximize the likelihoods (which violates the definition of a prior). The Bayesian odds ratios given here use the priors described in § 2.5. As can be seen, the likelihood ratio test and the odds ratio give quantitatively different results. Based on the odds ratio, the B9 and C17 samples are insufficient to discriminate between the two distribution functions. On the other hand, the odds ratio favors the lognormal distribution for the F220 sample.

4. DISCUSSION AND CONCLUSIONS

The parameters and parameter ranges differ for the three different burst samples. It is not clear whether we yet have a sufficiently large, properly defined burst sample from which to calculate the energy distribution. The two samples with spectroscopic redshifts do not have correct detection thresholds: the threshold for detecting the burst itself is usually reported, but the intensity threshold that triggers further localization and spectroscopic redshift determination has not been reported. Indeed, there may not yet be a formal definition of such a follow-up threshold. Thus these two samples are flawed. On the other hand, the validity of the variability-determined redshifts has not yet been proven, although the detection threshold was defined in choosing the burst sample. The importance of the detection thresholds for statistical studies of burst samples argues for well-defined (and reported) intensity thresholds for triggering the follow-up observations of the expected large number of *HETE-2* and *Swift* burst localizations.

The distributions of cumulative probabilities and the averages of these distributions indicate that the two functional forms used here are sufficient to describe the distribution of energies; consequently, I concluded that the data do not justify trying more complicated distribution functions at this time. The Bayesian odds ratio does not distinguish between these two functional forms for the two samples with spectroscopic redshifts, although it does favor the lognormal distribution for the large F220 sample with redshifts derived from the variability of the burst light curves.

The energy distribution cannot be determined for energies below the lowest energy threshold (i.e., the lowest burst energy corresponding to the fluence thresholds). Indeed, Hakkila et al. (1996) make a distinction between the "observed" and "intrinsic" luminosity functions in their study of luminosity functions for cosmological bursts based on the shape of the peak flux distribution; they point out that the observed distribution may be much narrower than the intrinsic distribution. In my study the true low-energy cutoff of the power-law distribution cannot be determined. Similarly, the likelihood contours for the lognormal distribution do not rule out broader distributions with lower central energies.

We anticipate that the *HETE-2* and *Swift* missions will result in the construction of a large burst sample with spectroscopic redshifts. Properly defined subsets can be studied to identify trends with burst redshift, duration, and other properties. As we move from the study of individual bursts to the study of burst ensembles, we must define and report the criteria (e.g., detection thresholds) by which the burst samples are collected. The promise of future missions can be realized fully only with the careful consideration of the burst database's selection criteria.

I thank the referee, C. Graziani, for his comments, which have improved the text. This work was performed under the auspices of the US Department of Energy by the Los Alamos National Laboratory under contract W-7405-Eng-36.

REFERENCES

Band, D., et al. 1993, ApJ, 413, 281
Bloom, J. S., Frail, D. A., & Sari, R. 2001, AJ, 121, 2879
Fenimore, E., & Ramirez-Ruiz, E. 2001, ApJ, submitted (preprint astro-ph/0004176)
Frail, D. A., et al. 2001, Nature, submitted (preprint astro-ph/0102282)
Hakkila, J., Meegan, C. A., Horack, J. A., Pendleton, G. N., Briggs, M. S., Mallozzi, R. S., Koshut, T. M., Preece, R. D., & Paciesas, W. S. 1996, ApJ, 462, 125
Jimenez, R., Band, D., & Piran T. 2001, ApJ, 561, 171
Norris, J., Marani, L., & Bonnel, J. 2000, ApJ, 534, 248
Reichart, D. E., Lamb, D. Q., Fenimore, E. E., Ramirez-Ruiz, E., Cline, T. L., & Hurley, K. 2001, ApJ, 552, 57
Schaefer, B., Deng, M., & Band, D. 2001, ApJ, submitted (preprint astro-ph/0101461)