Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **420,** 405–415 (2012) doi:10.1111/j.1365-2966.2011.20045.x

On multiple classes of gamma-ray bursts, as deduced from autocorrelation functions or bivariate duration/hardness ratio distributions

C. Koen[1⋆] and A. Bere[1,2]

[1]*Department of Statistics, University of the Western Cape, Private Bag X17, Bellville 7535, South Africa*
[2]*Department of Mathematics and Physics, Cape University of Technology, Bellville 7535, South Africa*

Accepted 2011 October 20. Received 2011 October 18; in original form 2011 August 16

ABSTRACT

Autocorrelation functions (ACFs) of 119 gamma-ray bursts (GRBs) monitored by the Burst Alert Telescope (BAT) on *Swift* are calculated. Contrasting with previous results for smaller numbers of bursts from other missions, the widths of the ACFs are not bimodally distributed. Although the distribution appears slightly asymmetrical, underlying mixtures of distributions can also probably be ruled out. Factors contributing to differences between the results presented here, and those in the literature, may include the differences in energy passbands used, and the superior sensitivity of the BAT instrument (which affects e.g. the redshift distribution of the detected GRB). The second part of the paper is concerned with the fitting of mixtures of bivariate Gaussians to the joint duration/hardness ratio data of 325 GRBs. A careful analysis confirms that a three-component mixture is the statistically most acceptable, but it is shown that the implied marginal distribution of the hardness ratios does not fit the data very well. It is also stressed that mixture components cannot automatically be assumed to represent different classes of GRBs. The point is illustrated by showing two substantially different, but statistically almost equivalent, mixture models for the distribution of 571 BAT-determined GRB durations.

Key words: methods: statistical – gamma-ray burst: general.

1 INTRODUCTION

A number of authors have studied the autocorrelation functions (ACFs) of gamma-ray bursts (GRBs), e.g. Borgonovo (2004, hereafter B04), Borgonovo et al. (2007, hereafter B07) and Vasquez & Kawai (2011). References to earlier work can be found in these papers. The three cited works considered the distribution of ACF widths across long-duration GRBs and showed it to be bimodal, if corrected for time dilation.

There are few formal statistical tests for the presence of multimodality in data; perhaps the most widely cited is the 'dip test' devised by Hartigan & Hartigan (1985). The test statistic resembles the well-known Kolmogorov–Smirnov (KS) statistic in that the maximum difference D between two cumulative distribution functions (CDFs) is considered. The first CDF is that of the actual data, while the other is chosen to be the particular unimodal distribution which minimizes D. Large values of D indicate multimodality. Significance levels are determined by simulations based on samples drawn from a flat (i.e. uniform) distribution – see Hartigan & Hartigan (1985) for the motivation. For the 22 ACF widths determined by B07, $D = 0.1246$, with significance level 0.3 per cent (i.e. 0.3 per cent probability of such a large value of D occurring by chance). (It is noted in passing that the formal p-value will be somewhat larger, since the precise detail of the width calculations were devised to *maximize* evidence for bimodality.)

The data used by B04 and B07 were taken from a number of sources – Gamma-ray Burst Monitor (GRBM), Burst and Transient Source Experiment (BATSE) and *Konus* – with the concomitant problems of different sensitivities, time resolutions, energy ranges, etc. In this paper, ACFs are calculated exclusively from *Swift* 'Burst Alert Telescope' (BAT) GRB data, since these are both abundant and homogenous. Light curves were downloaded from 'http://gcn.gsfc.nasa.gov/swift_gnd_ana.html' in the form of 64-ms (observer frame) time resolution light curves, corrected to give on-axis count rates. Background-subtracted sums of the counts in the four energy channels, covering the energy range 14–195 keV, are used. The calculations are described in Section 2, and the results are given and interpreted in Section 3.

Bimodality in GRB durations, as measured by the parameter T_{90} (the time interval over which 90 per cent of the burst energy is observed), is generally accepted as fact (e.g. Gehrels, Ramirez-Ruiz & Fox 2009). More recently, it has been claimed that there may be a third type of GRB, intermediate between the so-called short and long bursts. The claim is based on the fact that the distribution of $\log T_{90}$ values can be modelled by a mixture of three Gaussians. Furthermore, there is evidence that the bivariate distribution of T_{90}

⋆E-mail: ckoen@uwc.ac.za

and radiation hardness is best described as a mixture of three distributions – see Řípa et al. (2009) and Horváth et al. (2010), which also contain overviews of earlier work.

Although hypothesis tests have been carried out in these studies, in order to ascertain the correct number of components in the mixture of distributions, these tests have unfortunately not had the properties assumed for them. In particular, use has been made of likelihood ratio statistics (LRSs), which compare the Gaussian-mixture likelihoods under null and alternative hypotheses. Subject to certain regularity conditions, the LRSs have χ^2 distributions (e.g. Andrews 2001). However, in the case of this particular problem, the regularity conditions are not satisfied – for example, the parameters estimated under the alternative hypothesis are not identified at all under the null. The implication is that the true distribution of the LRS is non-standard – see e.g. the references in Miloslavsky & Van der Laan (2003). Similar considerations apply to the F-test which has been used to compare χ^2 statistics of different models (Řípa et al. 2009) – see Protassov et al. (2002).

In Section 5 below, the likelihood ratio tests are repeated using simulated percentage points of the statistics. Attention is also given to the evaluation of the fit of the models, since finding the optimal number of mixture components does not automatically guarantee that an adequate representation of the data has been found.

Values of *Swift* BAT T_{90}, peak fluxes and redshifts used below were taken from 'http://swift.gsfc.nasa.gov/docs/swift/archive/grb_table/' which is gratefully acknowledged .

Please note that in this paper all logarithms are referred to the base 10, unless explicitly indicated otherwise.

2 THE FORM OF THE ACF

We start by motivating a form of the ACF which differs slightly from that in the recent literature. The ACF at lag $\ell = k\Delta t$ is defined by

$$A(k\Delta t) = \frac{1}{N}\sum_{j=1}^{N-k} s_j s_{j+k} \,/ A(0)$$
$$= \frac{1}{N}\sum_{j=1}^{N-k} (m_j - b_j)(m_{j+k} - b_{j+k}) \,/ A(0), \quad (1)$$

where s_j, b_j and $m_j = s_j + b_j$, respectively, denote the source, background and total (source plus background) count rates, the time interval between measurements (or bin width) is Δt and the duration of the burst is $N\Delta t$. Neither source nor background count rates are directly observable during the burst. The background level can be estimated by e.g. fitting a low-order polynomial to pre- and post-burst light curves, and interpolating across the burst, giving estimates $\widehat{b}_j$. For practical application, equation (1) is then replaced by

$$A(k\Delta t) = \frac{1}{N}\sum_{j=1}^{N-k} (m_j - \widehat{b}_j)(m_{j+k} - \widehat{b}_{j+k}) \,/ A(0). \quad (2)$$

The variance of the source signal is

$$A(0) = \frac{1}{N}\sum_{j=1}^{N} [(m_j - \widehat{b}_j)^2 - V] \,, \quad (3)$$

where V is the variance of the background; V is assumed to be constant, independent of j. It can be estimated from the out-of-burst light curve:

$$V = \frac{1}{L}\sum_{i} (b_i - \widehat{b}_i)^2 \,, \quad (4)$$

where L is the number of background-only measurements used. Alternatively, if the individual background variances V_j are available, equation (3) could be replaced by

$$A(0) = \frac{1}{N}\sum_{j=1}^{N} [(m_j - \widehat{b}_j)^2 - V_j] = \frac{1}{N}\sum_{j=1}^{N} (m_j - \widehat{b}_j)^2 - \overline{V} \,.$$

Equation (1) differs from previous forms in that the upper limit to the summation here is $N - k$, rather than N. The latter limit requires that data values be *defined* at times $N + 1, N + 2, \ldots$; clearly, this would contribute to greater uncertainty in $A(\tau)$. Although not spelt out, it appears that the definition in the literature was motivated by the desire to have the number of terms in the summation be the same for all τ. However, it is known that, at least for stationary series, the ACF as defined in (1) is unbiased, i.e. it is not necessary to make any adjustment for the dependence of the summation limit in (1) on k.

In B04 and B07, the analogue of (3) is

$$A(0) = \frac{1}{N}\sum_{j=1}^{N} [(m_j - \widehat{b}_j)^2 - m_j] \,. \quad (5)$$

To see clearly the difference between (3) and (5), consider the following expected value (ensemble average):

$$\begin{aligned} E(m_j - \widehat{b}_j)(m_{j+k} - \widehat{b}_{j+k}) &= E(s_j + b_j - \widehat{b}_j)(s_{j+k} + b_{j+k} - \widehat{b}_{j+k}) \\ &= Es_j s_{j+k} + \mathrm{E}(b_j - \mathrm{E}b_j) \\ &\quad \times (b_{j+k} - \mathrm{E}b_{j+k}) \\ &= Es_j s_{j+k} + \mathrm{cov}(b_j, b_{j+k}), \end{aligned}$$

where it has been assumed that source and background are uncorrelated, and also that $\widehat{b}_j$ accurately estimates the mean background level at each time point $j\Delta t$ across the burst. For uncorrelated background noise,

$$\begin{aligned} E(m_j - \widehat{b}_j)(m_{j+k} - \widehat{b}_{j+k}) &= Es_j s_{j+k} + \delta(k,0)\mathrm{cov}(b_j, b_{j+k}) \\ &= Es_j s_{j+k} + V\delta(k,0), \end{aligned} \quad (6)$$

where δ is the Kronecker delta. It follows that for $A(0)$ as defined in (3), the expected values of the individual terms are $\mathrm{E}s_j^2$ and

$$EA(0) = \frac{1}{N}\sum_{j} \mathrm{E}s_j^2,$$

which is clearly the correct normalization in (1). For $A(0)$ as defined in (5), on the other hand,

$$E[(m_j - \widehat{b}_j)^2 - m_j] = Es_j^2 + [\mathrm{var}(b_j) - Es_j - Eb_j] \,.$$

If the background counts are Poisson distributed, the first and last terms in the square brackets cancel, but the term in Es_j remains.

The normalization by $A(0)$ in (3) works reasonably well, but is not perfect. Fig. 1 shows the light curves of three GRBs, with the associated ACFs in Fig. 2. Close-ups of the ACFs at short lags (Fig. 3) show that values of $A(0)$ are too small for the GRB 100814A and GRB 081008 data, and too large for GRB 091024.

B04 and B07 use the widths of the ACFs to derive characteristic time-scales of the bursts. The time-scale τ of a burst is essentially defined to be the lag at which the ACF has declined to the value 0.5. Clearly incorrect normalization will affect the results: if $A(0)$ is too small, τ will be overestimated, whereas if $A(0)$ is too large, τ is underestimated. A simple remedy is to fit a low-order polynomial to the ACF over small lags, extrapolate to find the value $A'(0)$ at zero lag and then to normalize by this value. In practical application, a quadratic is fitted to the low-lag ACF values larger than 0.8.

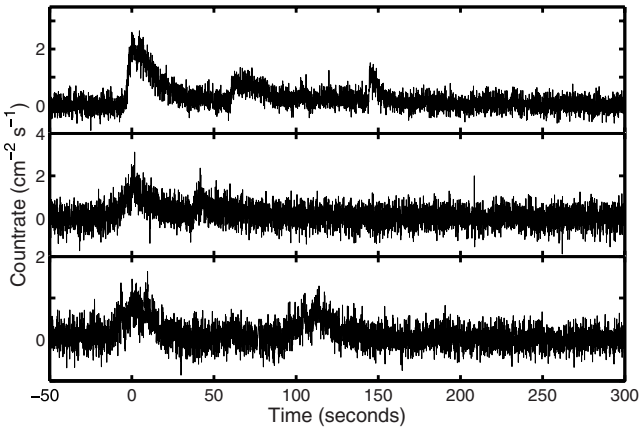

Figure 1. Light curves of GRB 100814A (top), GRB 091024 (middle) and GRB 081008 (bottom).

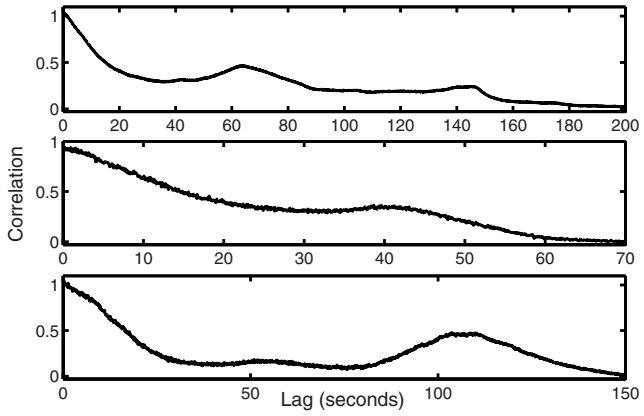

Figure 2. ACFs of the bursts in Fig. 1. The time intervals over which ACFs were calculated were $(-10, 250)$, $(-20, 55)$ and $(-20, 150)$ s, respectively.

The detail of the determination of the ACF widths by B04 and B07 is the following: the ACF is assumed to decay roughly exponentially. A quadratic $g(\tau)$ is therefore fitted to $\log A(\tau)$ over the interval $0.4 \leq A(\tau) \leq 0.6$, and the lag τ_0 such that $g(\tau_0) = \log 0.5$ is noted.

The top and bottom ACFs in Fig. 2 demonstrate that the algorithm cannot be applied blindly, since the respective secondary peaks near $\sim$65 and $\sim$110 s also satisfy $0.4 \leq A(\tau) \leq 0.6$. The plots furthermore raise questions regarding the aim of the analysis, since this would determine the most suitable measure of the width of the ACF. B07 state " ... it was found empirically ... that at half-maximum the separation between the two sets is most significant", i.e. their choice of methodology gives the clearest indication of bimodality. It would seem also to be useful to consider measures of the overall time-scale of a burst. A reasonable alternative is the mean $\tau_m \equiv \overline{\tau}$ of the lags such that $0.4 \leq A(\tau) \leq 0.6$, and this statistic will also be used in what follows. In the case of τ_0, if ACFs have multiple maxima higher than 0.4, then only the small-lag branch of the ACF is considered in applying the B04 and B07 algorithm.

Nothing has been said about the choice of the burst time interval over which ACFs are calculated. It was found that results were quite sensitive to the assumed burst starting point, but insensitive to the end point. This is perhaps not surprising, given the typical burst shapes (rapid onset, slow decline). Selecting a starting point a few seconds before the onset of the burst obviated the problem, giving more stable results. Selection of the end points was guided by the values of T_{90}. Within these constraints, ACFs were calculated for several beginning and end points; the consequent range of widths found for a given burst is reported below.

3 WIDTHS OF THE ACF

At the time this project was started, redshifts could be found for 162 *Swift* GRBs, the latest being GRB 100906A. Of these, 12 were short bursts ($T_{90} < 2$ s), and a further 31 were either incompletely observed or so noisy/faint that no analysis was attempted. This left 119 GRBs for which burst widths corrected for time dilation could be estimated. Results are given in Table 1. Fig. 4 summarizes the information in columns 2–4 of the table.

The uncertainties in τ_0 and τ_m, as reflected in the ranges quoted, are, with the exception of GRB 081028, relatively small. The range $\Delta\tau_0$ exceeds 1 s for only four GRBs, and $\Delta\tau_0 > 1.8$ s only for GRB

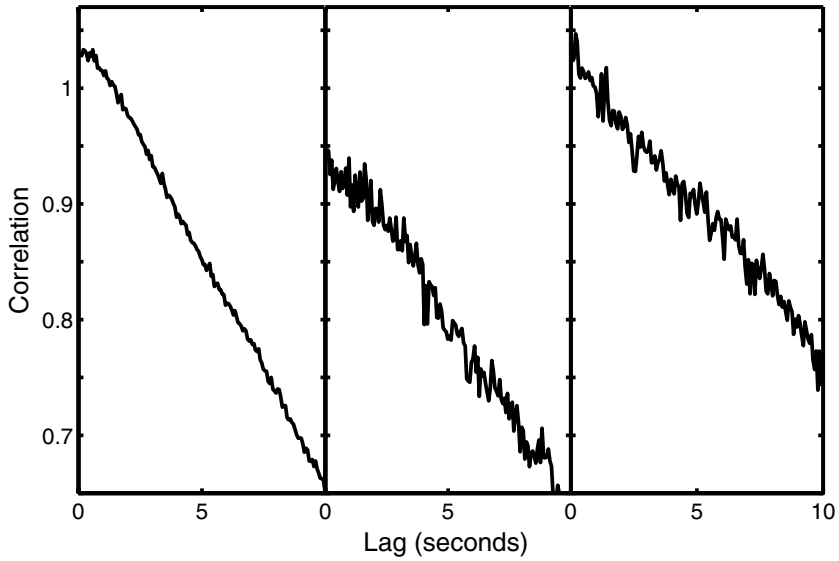

Figure 3. Details of the ACFs in Fig. 2, at short lags. From left to right, GRB 100814A, GRB 091024 and GRB 081008. Note the slightly incorrect normalizations in all three cases.

Table 1. Estimated rest-frame ACF widths of 119 *Swift* GRBs. For each burst, a range of values is given – these resulted from varying the burst time interval used for the ACF calculations.

GRB	Duration T_{90} (s)	Peak flux P (count s^{-1})	Redshift z	τ_0 range (s)		τ_m range (s)	
050126	24.8	0.7	1.29	5.81	6.34	6.26	6.74
050223	22.5	0.7	0.59	8.29	8.33	8.51	8.77
050315	95.6	1.9	1.95	8.55	8.61	8.38	8.51
050319	152.5	1.5	3.24	1.03	1.06	1.15	1.19
050401	33.3	10.7	2.90	0.75	0.75	1.09	1.23
050416A	2.5	4.9	0.65	0.47	0.47	0.43	0.45
050505	58.9	1.9	4.27	1.85	2.00	1.96	2.35
050525A	8.8	41.7	0.61	1.03	1.03	2.17	2.17
050603	12.4	21.5	2.82	0.30	0.30	0.26	0.27
050730	156.5	0.6	3.97	12.97	14.76	13.18	15.69
050802	19.0	2.8	1.71	2.50	2.65	2.60	2.87
050820A	26.0	0.5	2.61	3.79	4.40	3.86	4.28
050826	35.5	0.4	0.30	4.69	5.58	4.81	5.72
050904	174.2	0.6	6.30	13.96	14.91	13.90	14.99
050908	19.4	0.7	3.34	2.21	2.32	2.19	2.34
050922C	4.5	7.3	2.20	0.54	0.54	0.55	0.55
051016B	4.0	1.3	0.94	0.49	0.53	0.83	0.93
051109A	37.2	3.9	2.35	1.15	1.28	1.14	1.30
051109B	14.3	0.6	0.08	6.04	6.10	6.18	6.51
051111	46.1	2.7	1.55	4.39	4.47	4.51	4.63
060115	139.6	0.9	3.53	3.16	3.19	11.40	11.85
060206	7.6	2.8	4.05	0.55	0.55	0.55	0.55
060210	255.0	2.7	3.91	8.89	8.91	8.78	8.83
060223A	11.3	1.4	4.41	0.78	0.79	0.79	0.82
060319	10.6	1.1	1.15	1.43	1.52	1.51	1.52
060418	103.1	6.5	1.49	8.15	8.22	7.00	7.47
060502A	28.4	1.7	1.51	4.54	4.56	4.54	4.56
060505	4.0	2.7	0.09	1.82	1.82	1.88	2.29
060510B	275.2	0.6	4.90	24.80	25.86	24.20	24.54
060512	8.5	0.9	2.10	1.98	2.09	2.02	2.23
060522	71.1	0.6	5.11	5.54	5.72	5.30	5.58
060526	298.2	1.7	3.21	1.34	1.47	1.31	1.84
060605	79.1	0.5	3.78	2.21	2.41	2.20	2.33
060607A	102.2	1.4	3.08	4.39	4.49	4.06	4.09
060614	108.7	11.5	0.13	33.42	33.64	31.60	31.83
060707	66.2	1.0	3.43	3.58	4.03	3.71	4.12
060714	115.0	1.3	2.71	5.04	5.05	12.27	13.07
060814	145.3	7.3	0.84	8.14	8.28	23.83	24.03
060904B	171.5	2.4	0.70	2.25	2.25	2.30	2.39
060908	19.3	3.0	2.43	1.68	1.72	1.97	1.99
060912A	5.0	8.6	0.94	0.63	0.63	0.64	0.66
060926	8.0	1.1	3.21	0.65	0.67	0.74	0.78
060927	22.5	2.7	5.60	0.38	0.41	0.48	0.51
061007	75.3	14.6	1.26	10.31	10.31	8.89	8.89
061021	46.2	6.1	0.35	2.32	2.32	2.42	2.47
061110A	40.7	0.5	0.76	13.09	13.60	13.59	14.12
061110B	134.0	0.5	3.44	3.20	3.55	3.18	3.69
061121	81.3	21.1	1.31	2.94	2.96	2.72	2.73
061222A	71.4	8.5	2.09	2.09	2.09	2.23	2.23
061222B	40.0	1.6	3.36	5.35	5.38	5.26	5.27
070110	88.4	0.6	2.40	8.98	9.07	8.80	8.85
070306	209.5	4.7	1.50	4.25	4.25	4.34	4.35
070318	74.6	1.8	0.84	6.37	6.50	6.80	6.97
070411	121.5	0.9	2.95	8.45	8.79	12.28	12.56
070506	4.3	1.0	2.31	0.89	0.89	0.87	0.94
070508	20.9	24.1	0.82	4.25	4.25	4.17	4.17
070521	37.9	6.5	0.55	6.89	7.02	7.08	7.31
070529	109.2	1.4	2.50	4.19	4.53	7.73	8.46
070611	12.2	0.8	2.04	1.73	1.75	1.79	1.92
070612A	368.8	1.5	0.62	23.47	24.50	67.68	73.48
070714B	64.0	2.7	0.92	0.40	0.40	0.37	0.37

Table 1 – *continued*

GRB	Duration T_{90} (s)	Peak flux P (count s^{-1})	Redshift z	τ_0 range (s)		τ_m range (s)	
070721B	340.0	1.5	3.63	2.60	2.63	2.68	2.68
070802	16.4	0.4	2.45	2.93	3.73	2.97	4.41
070810A	11.0	1.9	2.17	1.23	1.25	1.24	1.27
071003	150.0	6.3	1.10	5.52	5.52	5.14	5.20
071010B	>35.7	7.7	0.95	2.69	2.72	2.74	2.77
071020	4.2	8.4	2.10	0.59	0.59	0.58	0.59
071031	180.0	0.5	2.69	3.75	4.04	4.08	5.00
071117	6.6	11.3	1.33	0.63	0.63	0.63	0.65
080210	45.0	1.6	2.64	3.25	3.48	3.31	3.71
080310	365.0	1.3	2.43	2.94	3.06	3.51	4.02
080319C	34.0	5.2	1.95	2.04	2.08	2.09	2.16
080330	61.0	0.9	1.51	0.46	0.56	1.23	1.33
080411	56.0	43.2	1.03	1.10	1.10	8.55	8.59
080413A	46.0	5.6	2.43	0.75	0.76	2.08	2.24
080413B	8.0	18.7	1.10	0.67	0.67	0.69	0.69
080430	16.2	2.6	0.77	2.28	2.31	2.30	2.37
080520	2.8	0.5	1.55	0.68	0.78	0.65	0.78
080603B	60.0	3.5	2.69	0.66	0.66	1.71	1.71
080605	20.0	19.9	1.64	2.81	2.81	2.70	2.70
080607	79.0	23.1	3.04	1.63	1.63	1.62	1.62
080707	27.1	1.0	1.23	1.15	1.18	7.80	8.67
080721	16.2	20.9	2.59	1.85	1.85	1.59	1.59
080804	34.0	3.1	2.20	2.90	2.90	3.03	3.06
080805	78.0	1.1	1.51	7.96	7.96	8.39	8.67
080810	106.0	2.0	3.35	6.14	6.16	4.98	5.04
080905B	128.0	0.5	2.37	1.61	1.69	1.76	1.80
080913	8.0	1.4	6.70	0.51	0.52	0.40	0.43
080916A	60.0	2.7	0.69	10.68	10.68	8.71	8.77
080928	280.0	2.1	1.69	3.90	4.26	3.51	4.13
081007	10.0	2.6	0.53	1.92	2.55	1.95	2.48
081008	185.5	1.3	1.97	5.67	5.71	26.00	26.36
081028	260.0	0.5	3.04	26.68	36.93	26.12	28.08
081121	14.0	4.4	2.51	2.97	3.10	2.96	3.14
081203A	294.0	2.9	2.10	9.66	9.66	10.55	10.80
081222	24.0	7.7	2.77	1.27	1.27	1.28	1.28
090102	27.0	5.5	1.55	4.99	4.99	4.94	4.96
090418A	56.0	1.9	1.61	9.15	9.27	10.04	10.10
090423	10.3	1.7	8.00	0.64	0.64	0.64	0.65
090424	48.0	71.0	0.54	1.12	1.12	1.54	1.54
090618	113.2	38.9	0.54	16.00	16.00	16.31	16.35
090715B	266.0	3.8	3.00	2.03	2.03	2.10	2.11
090812	66.7	3.6	2.45	6.75	7.09	5.74	5.78
090926B	109.7	3.2	1.24	10.31	10.40	10.56	10.63
090927	2.2	2.0	1.37	0.19	0.24	0.23	0.26
091018	4.4	10.3	0.97	0.84	0.84	0.84	0.88
091020	34.6	4.2	1.71	2.67	2.67	2.66	2.67
091024	109.8	2.0	1.09	9.68	9.71	14.61	15.02
091029	39.2	1.8	2.75	4.61	4.61	4.56	4.58
091127	7.1	46.5	0.49	0.82	0.82	0.69	0.71
091208B	14.9	15.2	1.06	0.47	0.50	0.50	0.50
100316B	3.8	1.3	1.18	0.97	1.03	1.00	1.01
100418A	7.0	1.0	0.62	2.76	3.03	3.05	3.23
100425A	37.0	1.4	1.76	1.30	1.30	4.74	6.46
100621A	63.6	12.8	0.54	18.30	18.39	16.35	16.37
100728B	12.1	3.5	2.11	1.21	1.32	1.27	1.32
100814A	174.5	2.5	1.44	5.80	5.80	16.44	17.27
100816A	2.9	10.9	0.80	0.64	0.64	0.62	0.64
100906A	114.4	10.1	1.73	3.28	3.31	3.25	3.25

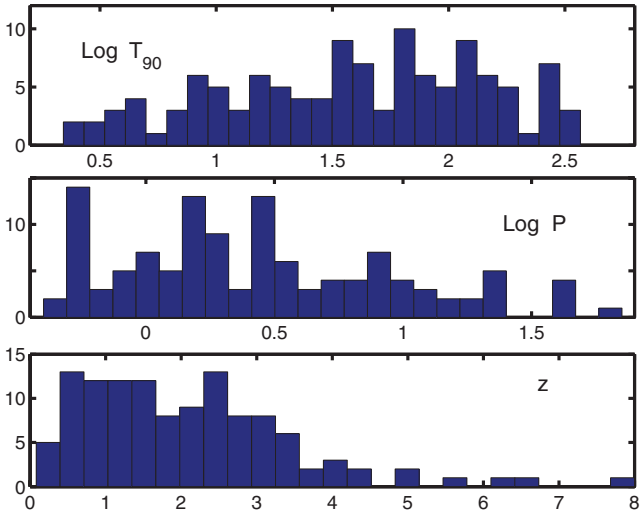

Figure 4. Histograms of the data in columns 2–4 of Table 1.

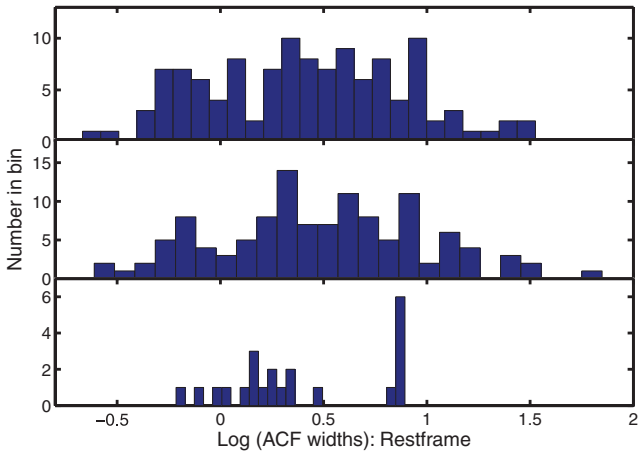

Figure 5. Histograms of the logarithms of the estimated rest-frame ACF widths; $\log(\tau_0)$ (top), $\log(\tau_m)$ (middle) and B07 results (bottom).

081028 ($\Delta\tau_0 = 10.25$). Similarly, $\Delta\tau_m > 1$ s for six GRBs, the maximum being $\Delta\tau_m = 5.8$ s for GRB 070612A. Below, τ_0 and τ_m will, respectively, be equated to the mean values of columns 5 and 6, or columns 7 and 8, i.e. to the mid-points of their ranges.

Generally speaking, there is reasonable agreement between τ_0 and τ_m. For 13 GRBs, the ratio $|\tau_m - \tau_0|/\tau_0 > 1$; the greatest discrepancies are for GRB 080411 and GRB 080707. There are few cases where $\tau_m < \tau_0$, the most notable being GRB 081028, which has a very noisy light curve (peak flux 0.5 s^{-1}). These results could, of course, have been anticipated from Fig. 1, which suggests that bursts with multiple peaks will have $\tau_m > \tau_0$.

In what follows, ACF widths are used in logarithmic form: $\tau_{L0} \equiv \log(\tau_0)$, $\tau_{Lm} \equiv \log(\overline{\tau})$. Histograms of the ACF widths are shown in Fig. 5. For comparison, observer frame distributions of the ACF widths are given in Fig. 6. There are suggestions of bimodality in Fig. 5 – there are dips in the histograms near $\tau_{L0} \approx 0.15$ and $\tau_{Lm} \approx 0$, respectively. Of course, this may simply be due to the choices of bin positions and/or statistical fluctuations, so a quantitative evaluation is desirable.

The dip statistics for the distributions τ_{L0} and τ_{Lm} are $D = 0.022$ and 0.028, respectively, with significance levels of 98 and 78 per cent, indicating that there is no evidence for multimodality.

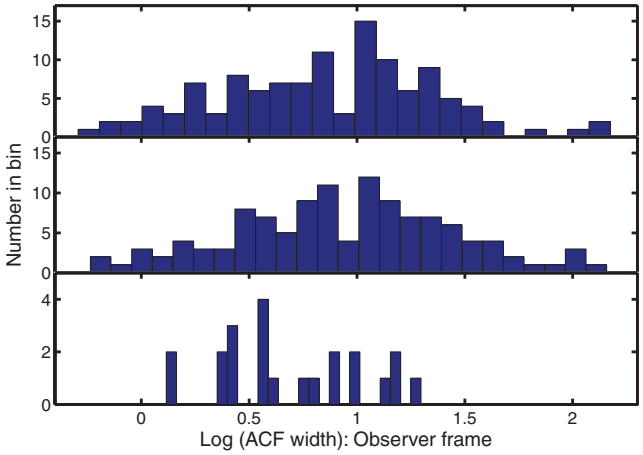

Figure 6. Histograms of the logarithms of the estimated observer frame ACF widths; $\log(\tau_0)$ (top), $\log(\tau_m)$ (middle) and B07 results (bottom).

Kernel smoothers provide more sophisticated estimators of the probability density functions (PDFs) of data than do histograms. Extensive descriptions can be found in e.g. Silverman (1986) and Wand & Jones (1995); a brief summary, repeated here, is given by Koen & Lombard (2003). The kernel estimator of the PDF f in the point x is defined by

$$\widehat{f}(x) = \frac{1}{Nh}\sum_{j=1}^{N} K\left(\frac{x - w_j}{h}\right),$$

where $N = 119$ is the number of data, w_j is the jth ACF width, K is a suitable kernel function and h is a smoothing bandwidth. The commonly used Epanechnikov and Triangular kernels are defined by

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2) & |x| \le 1, \\ 0 & |x| > 1, \end{cases}$$

and

$$K(x) = \begin{cases} (1 - |x|) & |x| \le 1, \\ 0 & |x| > 1, \end{cases}$$

respectively. The 'normal scale' bandwidths given by

$$h_{\rm E} = 2.34N^{-1/5}s \quad \text{and} \quad h_{\rm T} = 2.58N^{-1/5}s$$

are suitable for the two respective kernels; s is an estimate of the spread of the w_j. The outlier resistant estimator,

$$s = 0.741[w_{(0.75)} - w_{(0.25)}],$$

of the standard deviation is used: $w_{(0.75)}$ and $w_{(0.25)}$ are the 75th and 25th percentiles of the distribution of w_j values.

We also apply a density estimator recently proposed by Botev, Grotowski & Kroese (2010): it uses a Gaussian kernel with a non-parametric bandwidth and has good performance for multimodal data.

The results are given in graphical form in Figs 7 and 8. There is excellent agreement between the three different kernel estimators. The estimated PDF of τ_{Lm} is close to symmetrical, but there is a slight bump on the estimated PDF of τ_{L0}.

A statistical model which may explain the PDF shapes in Fig. 7 is a mixture of Gaussians. Denote the Gaussian PDF of the random variable w by $n(w; \mu, \sigma)$, where μ and σ are the mean and standard deviation. The mixture PDF is then

$$f(w) = \alpha n(w; \mu_1, \sigma_1) + (1 - \alpha)n(w; \mu_2, \sigma_2), \qquad (7)$$

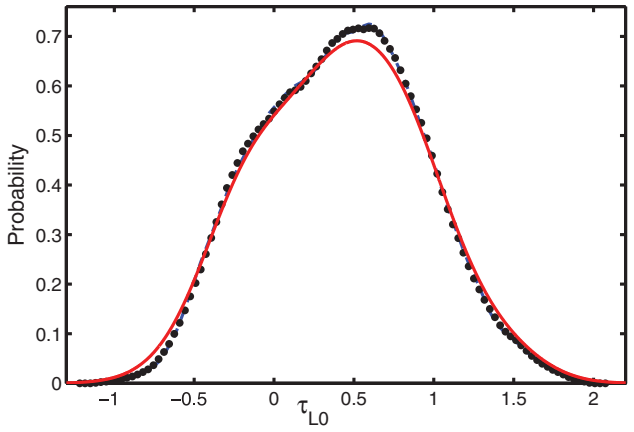

Figure 7. Kernel density estimates of the PDF of τ_{L0}. The dots and broken line (obscured by the dots) denote the triangular and Epachnikov kernel results and the solid line the Botev et al. (2010) estimate.

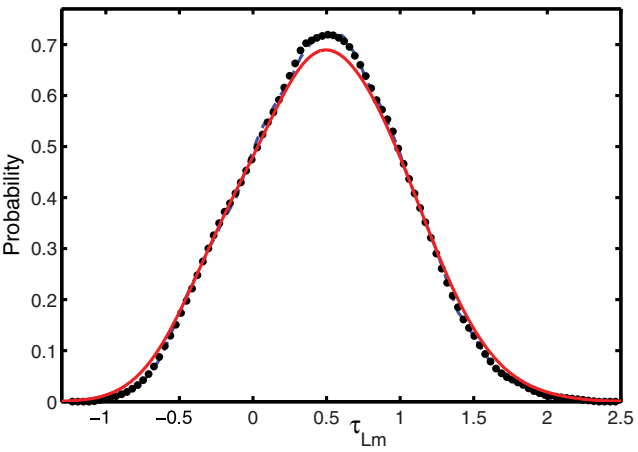

Figure 8. As for Fig. 7, but for τ_{Lm}.

where α is the fraction of w values with density $n(w;\ \mu_1,\ \sigma_1)$. Equation (7) was fitted to the τ_{L0} and τ_{Lm} data sets, with the results shown in Table 2. Centring of the second component at τ_{L0}, $\tau_{Lm} \approx -0.2$ is not surprising, given the shapes of the histograms in Fig. 5 (see especially the middle panel).

It is interesting to compare current results with those obtained by B07 for 22 burst ACFs (last line of Table 2). There is little correspondence between the two sets of values. This may be due to differences in the nature of the data (*Swift* versus earlier missions), slight differences in the analysis techniques and/or the fact that our data set is substantially larger. Further discussion follows in Section 4.

Table 2. Parameters estimated from fitting the Gaussian mixture model (7) to the logarithms of the ACF widths. Results for logarithmically transformed B07 ACF widths are also shown for comparison.

Data set	μ_1	σ_1	α	μ_2	σ_2	$1-\alpha$
τ_{L0}	0.50	0.44	0.88	−0.22	0.083	0.12
τ_{Lm}	0.54	0.47	0.93	−0.22	0.081	0.07
B07	0.87	0.022	0.32	0.16	0.18	0.68

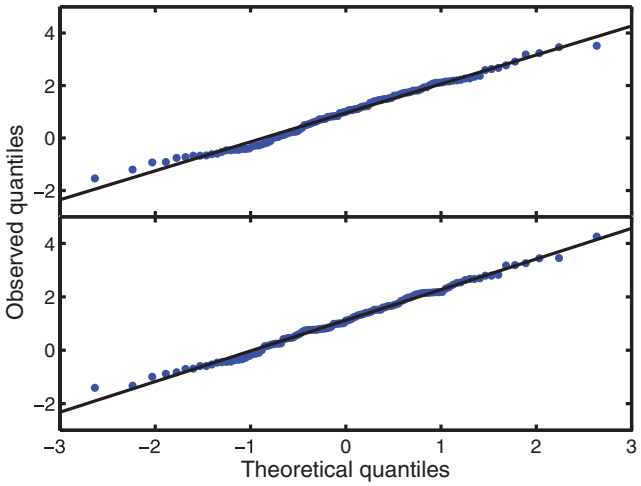

Figure 9. Normal distribution quantile–quantile plots for the τ_{L0} (top panel) and τ_{Lm} (bottom panel) distributions. Theoretical quantiles are based on the assumption that the data are Gaussian.

The question of the statistical significance of describing the PDFs of the data as a mixture of two components, as opposed to one component, still needs to be addressed. This can be done by calculating the LRS Λ_{21}, which essentially compares the probability of the mixture model (7) to the representation of the data by a single Gaussian. It was pointed out in Section 1 that the LRS has a non-standard distribution in this context. McLachlan (1987) suggested establishing the distribution of the LRS by simulation under the null hypothesis (i.e. a single Gaussian fit to the data). This leads to significance levels of $p = 0.12$ and 0.49 for the τ_{L0} and τ_{Lm} distributions, respectively, i.e. favouring the single-Gaussian description. (It is noted in passing that, similarly to Feng & McCulloch 1994, a distribution of the LRS between χ^2_5 and χ^2_6 is obtained.)

It remains to be seen whether the single-Gaussian fit to the data is in fact adequate. As a first informal test, normal quantile–quantile plots may be examined: these are displayed in Fig. 9. There is clearly no gross deviation from normality. This is followed up by two formal tests, namely the D'Agostino–Pearson and Anderson–Darling tests (see e.g. D'Agostino & Stephens 1986). The first of the two tests combines third (skewness) and fourth (kurtosis) moment information into a single χ^2_2 statistic: significance levels are 17 and 63 per cent for the τ_{L0} and τ_{Lm} data sets, respectively. The Anderson–Darling statistic is significant at levels of 27 and 79 per cent, respectively.

It may be concluded that the distributions of τ_{L0} and τ_{Lm} are both well described by single Gaussians. The parameter values are $(\mu = 0.42, \sigma = 0.48)$ for τ_{L0} and $(\mu = 0.49, \sigma = 0.50)$ for τ_{Lm}. The implication is of course that τ_0 and τ_m are lognormally distributed.

A further point which deserves mention is the significant correlation of the ACF widths with the peak fluxes – see Fig. 10. The regression lines are given by

$$\begin{aligned} \tau_{L0} &= 0.50(0.056) - 0.21(0.081)\log(P) \quad \sigma = 0.47, \\ \tau_{Lm} &= 0.58(0.058) - 0.22(0.085)\log(P) \quad \sigma = 0.49. \end{aligned} \tag{8}$$

The implication is that the distributions would generally be shifted to smaller values if bright cut-offs were introduced.

The possibility was investigated that the data in Fig. 10 would be better modelled by a mixture of regressions (see e.g. Turner 2000; Hurn, Justel & Robert 2003). If this were the case, then it

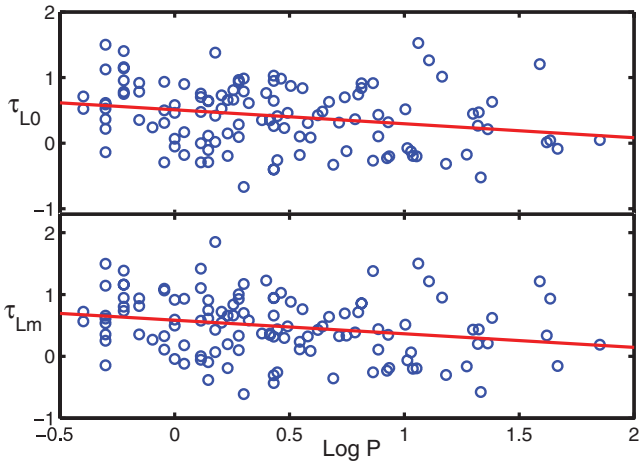

Figure 10. The dependence of ACF widths on the peak flux P: a probable selection effect.

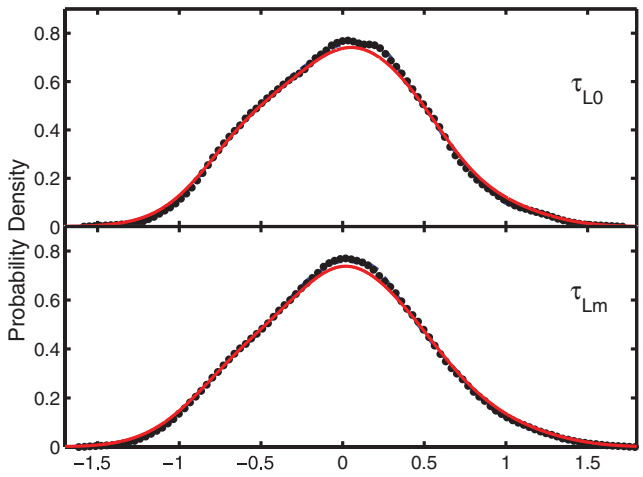

Figure 11. As for Figs 7 and 8, but showing kernel estimates of the PDFs of τ_{L0} and τ_{Lm}, after correction for the dependence on the peak count rate P.

may happen that subsets of the ACF width data, selected on the basis of P (e.g. bright, high-P bursts), could exhibit bimodality. The analysis was done in R , using the package MIXREG. The Gaussian LRS was used to test the null hypothesis of a single linear regression against the alternative of two regressions. A parametric (Gaussian) bootstrap procedure was used to find the significance levels of the LRSs for the τ_{L0} and τ_{Lm} data sets. The p-values obtained were 0.57 and 0.93, respectively, implying that a single regression line provides an adequate fit. This is perhaps not too surprising, as the origin of (8) may be just a selection effect: fainter narrow bursts are missed due to the increased role of background noise.

More directly, examination of kernel density estimates of bright subsets of τ_{L0} and τ_{Lm} did not provide any evidence for bimodality.

Fig. 11 shows the kernel density estimates of τ_{L0} and τ_{Lm}, corrected for the dependence on P (as given by equation 8). The D'Agostino–Pearson and Anderson–Darling test statistics indicate that both corrected data sets have distributions consistent with normality.

4 FACTORS WHICH MAY CONTRIBUTE TO DIFFERENCES BETWEEN SECTION 3 RESULTS AND THOSE IN THE LITERATURE

There are several factors which may contribute to the differences between the B07 results and those derived above. Given the relatively slight difference in distributions of τ_{L0} and τ_{Lm}, it seems unlikely that the differences in the methodology for estimating ACF widths plays a major role in the difference between our results and those in the literature. The differences in the observing platforms (*Swift* BAT versus earlier missions), with the concomitant differences in selection effects, appear much more likely to lie at the heart of the differences. Prime differences are those in energy ranges covered, and sensitivities, which again have a bearing on the redshift distributions of the observed GRBs.

B07 adopted BATSE as their 'reference instrument' and summed the counts from its channels 2 and 3 (together covering the 55–320 keV range). The rest of their data came from the 'GRBM' on *BeppoSAX*, and *Konus*, with energy ranges 40–700 and 50–200 keV, respectively. By comparison, BAT covers 14–24, 24–50, 50–100 and 100–195 keV in its four channels. Fluxes are generally higher at lower energies, and bursts are wider at lower energies [Fenimore et al. 1995; compare also the greater widths obtained by B07 for the softer (2–26 keV) *BeppoSAX* Wide Field Camera observations]. It follows that, for a given GRB, pulse widths – and hence ACF widths – measured by BAT would be larger than those observed by BATSE, GRBM and *Konus*. None the less, a formal comparison, by the two-sample KS statistic, finds a non-significant ($p = 0.13$) difference between the observer-frame distributions in the top and bottom panels of Fig. 6.

A visual comparison of the redshift distributions is shown in Fig. 12. Given that the respective fractions of GRB with $z > 2$ are 48 and 14 per cent, it is hardly surprising that the two samples are significantly different ($p = 0.0019$ for the two-sample KS statistic). This implies that many of the *rest-frame* energy ranges observed by BAT may be comparable to those observed by the older instruments: this will counteract to some extent the effects of the softer *observer* frame energy range, as discussed in the previous paragraph.

Repeating the calculations of τ_{L0} and τ_{Lm}, but summing only the two higher energy BAT channels (i.e. 50–195 keV), would be an interesting exercise.

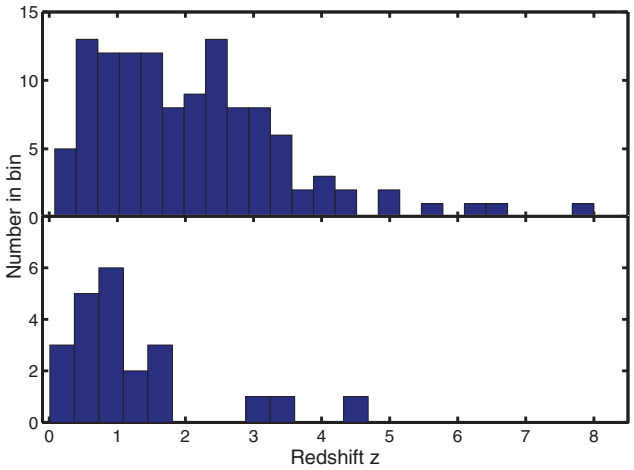

Figure 12. A comparison of the redshift distribution of the GRBs of this paper (top) and those of B07 (bottom).

5 MIXTURE MODELLING OF T_{90} AND HARDNESS RATIO DATA

Dr Istvan Horváth kindly supplied the data analysed in Horváth et al. (2010). These are logarithms of the durations T_{90} and the hardness ratios HR (50–100 keV fluence divided by 25–50 keV fluence) of 325 GRBs observed by the BAT on *Swift*.

Mixtures of two, three and four bivariate Gaussians were fitted to the paired ($\log T_{90}$, $\log$ HR) data. The LRSs are $\Lambda_{32} = 35.85$ and $\Lambda_{43} = 12.81$. Percentage points of Λ_{32} were calculated by simulating 1000 data sets of size 325, using the parameters of the optimal two-component bivariate model fitted to the observations. The process was also carried out for Λ_{43}, using parameters of the best-fitting three-component bivariate model. The upper percentage points of the two simulated distributions are quite similar, as can be seen from inspection of Fig. 13. The significance levels of the statistics are $p < 0.001$ (Λ_{32}) and $p = 0.15$ (Λ_{43}), confirming the conclusion of Horváth et al. (2010) that the best model is one with three components.

The next task is to establish whether the three-component model provides an adequate representation of the model, i.e. a goodness of fit test needs to be carried out. This is considerably simpler for univariate data; hence we first test whether the two marginal distributions, each consisting of a mixture of three Gaussians, respectively, provide adequate fits to the duration and hardness ratio data.

A goodness of fit test such as the KS test can be used to establish whether the fitted distribution gives a good representation of the empirical distribution. However, the distribution of the KS statistic is known only in the case that the theoretical distribution is fully specified, i.e. parameters values are all known ab initio, and need not be estimated. The same problem is encountered with other test statistics. Use will therefore be made of the bootstrapping procedure described by Stute, González-Manteiga & Presedo Quindimil (1993). For convenience, it is assumed that the test statistic T is based on a comparison of the empirical CDF F_n, and a partially specified theoretical CDF $F(\boldsymbol{\theta})$ (as in the case of the KS statistic).

(i) The theoretical PDF f (mixture of Gaussians, in the present case) depends on a number of unknown parameters (means, variances and mixture proportions). Let $\boldsymbol{\theta}$ be the vector of unknowns: estimate these, and denote the estimate by $\widehat{\boldsymbol{\theta}}$.

(ii) Calculate the statistic of interest, $T_0 = T[F_n, F(\widehat{\boldsymbol{\theta}})]$.

(iii) Simulate a sample of size n from the PDF $f(\widehat{\boldsymbol{\theta}})$ corresponding to $F(\widehat{\boldsymbol{\theta}})$. Determine the empirical CDF F_{n*} of these synthetic data.

(iv) From the simulated sample, estimate the parameter values, in exactly the same manner in which $\widehat{\boldsymbol{\theta}}$ was estimated from the real observations. Let the vector of estimates be $\widehat{\boldsymbol{\theta}}_*$.

(v) Calculate the statistic $T_* = T[F_{n*}, F(\widehat{\boldsymbol{\theta}}_*)]$.

(vi) Repeat steps (iii)–(v) many (preferably a few thousand) times and determine the percentile of T_0 with respect to the collection of T_* values.

The procedure was carried out for the three-component marginal distributions of the durations and hardness ratios, using the Anderson–Darling statistic A^2. A thousand simulated data sets were generated for each of the two tests. The respective significance levels $p = 0.57$ (durations) and $p = 0.023$ (hardness ratios) were obtained: it follows that a three-component mixture of Gaussians provides an adequate description of the distribution of durations, but *not* of the hardness ratio distribution. Inspection suggests that the discrepancy is in the central part of the HR distribution, rather than in the tails.

Bivariate models describe the simultaneous distribution of the two variables of interest. It is instructive to also consider the two individual data sets of durations and hardness ratios. Univariate tests for the number of mixture components in the distributions of $\log T_{90}$ and $\log$ HR give significance levels for the LRS Λ_{32} of $p = 0.021$ and 0.48, respectively, indicating three- and two-component mixtures, respectively. The Anderson–Darling goodness of fit statistics are not significant ($p = 0.88$ for the three-Gaussian fit to the T_{90} data and $p = 0.42$ for a two-component fit to the HR data).

The implication of the preceding material is the following: although a three-component fit to the bivariate (T_{90}, HR) distribution is formally preferred, it does not fit the data very well. The distribution of the HR values alone is best described by a two-component Gaussian. It follows that a statistically acceptable model of the data should consist of three- and two-component mixture distributions for the two marginals, while the interdependance of T_{90} and HR needs to be described by e.g. a copula (see Genest & Favre 2007, for an introduction). It appears unlikely that a standard bivariate parametric distributional form could do the job of modelling the observed distribution.

Similar considerations apply in the case of the bivariate (T_{90}, HR) *RHESSI* data studied by Řípa et al. (2009). The authors concluded that the distribution of the data can be modelled by a mixture of three bivariate Gaussian. Here we focus of the log HR component. Based on 1000 simulations, the significance level of $\Lambda_{21} = 0.0013$ is $p = 1$, implying that the HR data are unimodal. In fact, $p > 0.1$ for both the Anderson–Darling and D'Agostino–Pearson tests for normality of the data, confirming that a single Gaussian is an excellent model for the logs of the hardness ratios.

Finally, a two-dimensional KS test was applied to the Horváth et al. (2010) data to ascertain whether the three-component bivariate mixture fits the (T_{90}, HR) data adequately. The test statistic was first described by Peacock (1983), but a simplified form due to Fasano & Franceschini (1987) was used here. The significance level was determined by using the bootstrapping recipe given above: $p = 0.43$, from 1000 simulated data sets. Contrary to what was found from the marginal distributions, this suggests that the three-component mixture is a good fit to the data. However, simulation experiments

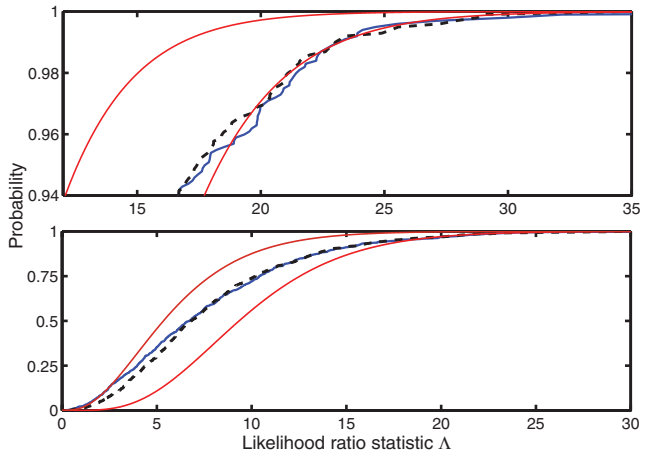

Figure 13. Bottom panel: the simulated CDFs of the LRSs Λ_{32} (broken line) and Λ_{43} (inner solid line). The outer envelopes are the CDFs of the χ^2_6 (upper curve) and χ^2_{10} (lower curve) distributions, respectively. Upper panel: an expanded view of the upper percentiles of the distributions.

Table 3. Two competing mixture models for the distribution of 571 *Swift* BAT values of log T_{90}.

Component	Mean μ	Std. dev. σ	Proportion
1	−0.61	0.44	0.069
2	1.42	0.62	0.669
3	1.98	0.29	0.262
1	−0.52	0.47	0.080
2	0.89	0.37	0.232
3	1.83	0.42	0.688

indicate that the KS statistic may not be very powerful. Further investigation of this point is required.

This section of the paper is closed with a note of caution: there is no guarantee that the components of a mixture correspond to physically distinct classes of objects. It is entirely possible that the distributions of class properties, such as log T_{90}, are non-normal: in such a case, spurious classes would be identified due to the modelling of a non-normal distribution by normal components. The point is underlined by the following analysis. Duration data for 571 *Swift* BAT GRBs are currently (May 2011) available, and various mixture models were fitted to these. The LRS which compares the two- and three-component models is $\Lambda_{32} = 18.2$; 5000 simulations based on the two-component solution assigns this a significance level of 0.36 per cent. (For values of the LRS larger than about 3, its distribution is bracketed by χ^2_4 and χ^2_6 distributions.) The Anderson–Darling statistic is $A^2 = 0.119$, with a significance level of 67 per cent, i.e. the three-component mixture provides a good representation of the data distribution.

The first half of Table 3 contains the estimated parameters of the optimal three-component mixture. The log likelihood is −619.16. The likelihood has a second local maximum, with value −619.70, corresponding to a quite distinct solution (second half of Table 3). The LRS comparing the two models is $\Lambda = 1.08$: Λ has the conventional χ^2_8 distribution in this case, and the two solutions do not differ significantly. The two mixtures are compared in Fig. 14: although the two PDFs with the lowest mean values are very similar, the other two components identified in the rival models are entirely different. The two sums of mixtures, shown by the dotted lines, are virtually indistinguishable on the scale of these plots, as may have been expected on the basis of their closely similar likelihoods. This means that there are at least two quite distinct, but statistically equivalent, descriptions of the data set.

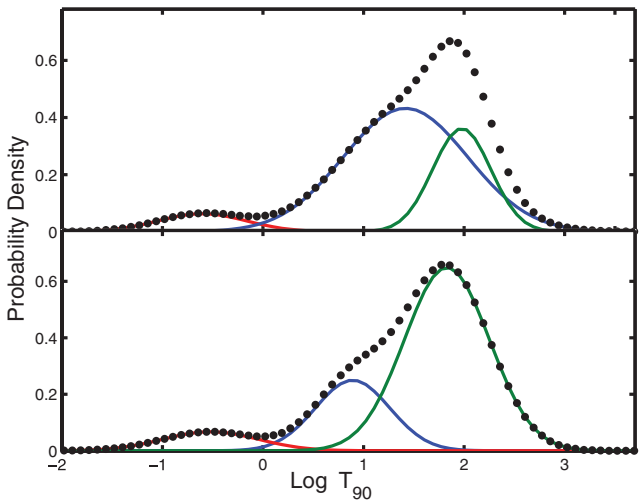

Figure 14. Two different three-component mixture models of the distribution of *Swift* BAT GRB durations. The dotted lines indicate the sums of the three component PDFs.

6 SUMMARY

(i) A slight improvement in the normalization of the GRB ACF was suggested in the form of equation (3). This is still not always adequate, as demonstrated in Fig. 3; hence extrapolation of $A(\ell)$ from larger lags to $\ell = 0$, in order to determine $A(0)$, is a useful alternative.

(ii) An alternative, more robust, measure of the ACF width was proposed.

(iii) The distribution of the ACF widths appears to be normal, i.e. there is no significant evidence for bimodality or mixtures of distributions which may have suggested different classes of GRBs.

(iv) Simulated percentage points of LRSs confirm that the bivariate Gaussian mixture with three components is the preferred model for the bivariate distribution of GRB durations and hardness ratios. The model cannot be considered a good fit though, since the marginal distribution does not fit the observed distribution of hardness ratios.

(v) Study of the two univariate data sets shows that the durations and hardness ratios are, respectively, best modelled by three- and two-component Gaussian mixtures. In the case of the *RHESSI* data, the distribution of log HR is that of a single Gaussian.

(vi) It is demonstrated that two very different mixtures, both of which are statistically acceptable, can be used to model the currently largest collection of homogenous T_{90} values. This casts doubt on the interpretation of the number of mixture components as being indicative of the number of GRB classes.

ACKNOWLEDGMENTS

CK thanks Kim Page (University of Leicester) and Scott Barthelmy (NASA) for assistance with *Swift* BAT related queries. The authors are indebted to Dr Istvan Horváth for sharing his duration/hardness ratio data; to Taka Sakamoto and Scott Barthelmy for the BAT data products; to Drs Cristiano Guidorzi and Luis Borgonovo for exchange of ideas regarding the calculation of GRB widths; and to the referee for useful comments. CK acknowledges funding from the South African National Research Foundation.

REFERENCES

Andrews D. W. K., 2001, Econometrica, 69, 683
Borgonovo L., 2004, A&A, 418, 487 (B04)
Borgonovo L., Frontera F., Guidorzi C., Montanari E., Vetere L., Soffitta P., 2007, A&A, 465, 765 (B07)
Botev Z. I., Grotowski J. F., Kroese D. P., 2010, Ann. Statistics, 38, 2916
D'Agostino R. B., Stephens M. A., eds., 1986, Goodness-of-fit Techniques. Marcel Dekker, Inc., New York
Fasano G., Franceschini A., 1987, MNRAS, 225, 155
Feng Z. D., McCulloch C. E., 1994, Biometrics, 50, 1158
Fenimore E. E., in 't Zand J. J. M., Norris J. P., Bonnell J. T., Nemiroff R. J., 1995, ApJ, 448, L101
Gehrels N., Ramirez-Ruiz E., Fox D. B., 2009, ARA&A, 47, 567
Genest C., Favre A.-C., 2007, J. Hydrologic Eng., 12, 347
Hartigan J. A., Hartigan P. M., 1985, Ann. Statistics, 13, 70
Horváth I., Bagoly Z., Balázs L. G., de Ugarte Postigo A., Veres P., Mészáros A., 2010, ApJ, 713, 552

Hurn M. M., Justel A., Robert C. P., 2003, J. Comput. Graphical Statistics, 12, 55
Koen C., Lombard F., 2003, MNRAS, 343, 241
McLachlan G. J., 1987, J. R. Statistical Soc. C, 36, 318
Miloslavsky M., van der Laan M. J., 2003, Comput. Statistics Data Analysis, 41, 413
Peacock J. A., 1983, MNRAS, 202, 615
Protassov R., van Dyk D. A., Connors A., Kashyap V. L., Siemiginowska A., 2002, ApJ, 571, 545
Řípa J., Mészáros A., Wigger C., Huja D., Hudec R., Hajdas W., 2009, A&A, 498, 399
Silverman B. W., 1986, Density Estimation for Statistics and Data Analysis. Chapman & Hall, London
Stute W., González-Manteiga W., Presedo Quindimil M., 1993, Metrika, 40, 243
Turner T. R., 2000, J. R. Statistical Soc. C, 49, 371
Vasquez N., Kawai N., 2011, Phys. E, 43, 689
Wand M. P., Jones M. C., 1995, Kernel Smoothing. Chapman & Hall, London

This paper has been typeset from a TEX/LATEX file prepared by the author.