Mon. Not. R. Astron. Soc. **328,** 283–290 (2001)

Reclassification of gamma-ray bursts

Andreu Balastegui,[1★] Pilar Ruiz-Lapuente[1,2★] and Ramon Canal[1,3★]
[1]*Departament d'Astronomia i Meteorologia, Universitat de Barcelona, Martí i Franqués 1, Barcelona 08028, Spain*
[2]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, 85740 Garching bei München, Germany*
[3]*Institut d'Estudis Espacials de Catalunya, Nexus Building, 2-4 Gran Capità, Barcelona 08034, Spain*

Accepted 2001 August 3. Received 2001 July 31; in original form 2001 May 21

ABSTRACT

We have applied two different automatic classifier algorithms to the BATSE Current GRB Catalog data and we obtain three different classes of gamma-ray bursts (GRBs). Our results confirm the existence of a third, intermediate class of GRBs, with mean duration ∼25–50 s, as deduced from a cluster analysis and from a neural network algorithm. Our analyses imply longer durations than those found by Mukherjee et al. and Horváth, whose intermediate class had durations ∼2–10 s. From the neural network analysis no difference in hardness between the two longest classes is found, and from both methods we find that the intermediate-duration class constitutes the most homogeneous sample of GRBs in its space distribution, while the longest-duration class constitutes the most inhomogeneous one with $\langle V/V_{\max}\rangle \sim 0.1$, being thus the deepest population of GRBs with $z_{\max} \sim 10$. The trend previously found in long bursts, of spatial inhomogeneity increasing with hardness, only holds for this new longest-duration class.

Key words: methods: data analysis – methods: statistical – gamma-rays: bursts.

1 INTRODUCTION

Since their discovery in the late 1960s (Klebesadel, Strong & Olson 1973), gamma-ray bursts (GRBs) have been a long-remaining puzzle (see Piran 2000, Mészáros 2001 and Castro-Tirado 2001 for recent reviews). Models involving a short distance scale implied an emitted energy $\sim 10^{42}$ erg, whereas a cosmological origin required 10^{51} erg at least. The increasing degree of isotropy found as the sample of GRBs grew, plus the lack of faint GRBs (Meegan, Fishman & Wilson 1985), favoured a cosmological scenario. Moreover, after the launch, in 1991, of the BATSE instrument on board the *CGRO*, a very high degree of isotropy was found in the new, much larger sample. Finally, the measurement in 1997 of the first GRB redshift (Metzger et al. 1997) and the subsequent ones have confirmed that GRBs are at cosmological distances.

Concerning the physical mechanism of GRBs, the fireball shock model (Rees & Mészáros 1992; Mészáros & Rees 1993; Daigne & Mochkovitch 1998) is a progenitor-independent model for radiation emission that succeeds in explaining both the burst itself and its afterglow. There are, however, a variety of proposed objects that are capable of generating the GRBs (Nemiroff 1994): from mergings of neutron stars with neutron stars (Paczyński 1990) or with black holes (Narayan, Paczyński & Piran 1992) to collapsars (Woosley 1993; MacFadyen & Woosley 1999) and hypernovae (Paczyński 1998). Magnetic instability in a neutron star being spun up by accretion in an X-ray binary could also produce them (Spruit 1999). There are more exotic models, involving quark stars (Ma & Xie 1996), mirror stars (Blinnikov 1999) or cosmic strings (Berezinsky, Hnatyk & Vilenkin 2001). In the face of such a boiling-pot of theoretical ideas, to know the positions of the GRBs with respect to their host galaxies and especially their redshifts are key issues (Bloom, Kulkarni & Djorgovski 2000). Once a fair sample of redshifts becomes available, which should happen soon with the up-to-date technology of new missions like *HETE 2* and *Swift*, the distances will be known and with that the released energy and the luminosity function of GRBs, together with their distribution across the Universe. That should certainly discriminate among existing models, and it should also give unprecedented information on the very structure of the Universe up to redshifts far higher than 5 (Lamb & Reichart 2000, hereafter LR00), on the cosmic star formation history (Totani 1999), and on the evolution of galaxies (Totani 1997).

It is likely that more than strictly one progenitor could give rise to GRBs, since it has been shown that different objects can produce a burst of gamma-rays with the observed characteristics. Therefore the catalogue of GRBs may reflect the manifestations of various phenomena. To uncover them, various attempts to separate different classes of GRBs have been made.

The two most generally accepted classes of GRBs are those arising from the bimodal distribution of their durations (Kouveliotou et al. 1993, hereafter K93), which separates long (lasting for more than 2 s) and short (less than 2 s) GRBs, the short

★E-mail: abalaste@am.ub.es (AB); pilar@am.ub.es (PR-L); ramon@am.ub.es (RC)

Table 1. Principal component analysis for the standardized logarithms of fluencies, peak fluxes and durations. There are shown, in each row, the components of each principal axis in the base of our original variables (columns), together with the percentage of the variance carried by each of the new axes (first column). For instance, the first principal component is $-0.39\log F_{\rm Ch1} - 0.40\log F_{\rm Ch2} - 0.40\log F_{\rm Ch3} - 0.33\log F_{\rm Ch4} - 0.22\log P_{64} - 0.28\log P_{256} - 0.36\log P_{1024} - 0.29\log T_{50} - 0.30\log T_{90}$.

%	$\log F_{\rm Ch1}$	$\log F_{\rm Ch2}$	$\log F_{\rm Ch3}$	$\log F_{\rm Ch4}$	$\log P_{64}$	$\log P_{256}$	$\log P_{1024}$	$\log T_{50}$	$\log T_{90}$
64.3	−0.39	−0.40	−0.40	−0.33	−0.22	−0.28	−0.36	−0.29	−0.30
27.0	+0.15	+0.12	+0.04	−0.05	−0.53	−0.47	−0.30	+0.44	+0.41
4.9	−0.22	−0.19	+0.06	+0.92	−0.10	−0.13	−0.17	−0.06	−0.08
1.7	+0.48	+0.41	+0.21	+0.03	−0.22	−0.25	−0.13	−0.47	−0.46
0.8	+0.56	−0.05	−0.77	+0.19	+0.16	+0.06	−0.09	−0.01	+0.12
0.6	+0.01	+0.11	+0.19	−0.06	+0.60	−0.03	−0.75	+0.13	−0.00
0.4	−0.02	−0.05	+0.11	−0.02	+0.04	−0.04	−0.07	−0.69	+0.71
0.2	+0.49	−0.78	+0.38	−0.07	−0.04	+0.05	−0.03	+0.04	−0.06
0.1	−0.01	+0.08	+0.01	+0.00	−0.46	+0.78	−0.40	−0.03	+0.01

bursts being at the same time spectrally harder than the long bursts. The different spatial distributions of the two classes have also been shown (Katz & Canel 1996), and it is consistent with isotropy in both. However, the longer GRBs appear to be more inhomogeneous in their space distribution than the shorter ones, as deduced from the higher value of $\langle V/V_{\rm max}\rangle$ for the short GRBs (this quantity measures the deviation of the space distribution from a homogeneous Euclidean distribution). Separating the long-bursts class into two groups of hardness H_{32} (the fluence ratio of spectral channel 3 to spectral channel 2), respectively higher than and lower than 3, Tavani (1998, hereafter T98) found that the long/hard bursts are more inhomogeneously distributed than the long/soft ones. We will show here that there exists, indeed, a trend, among the long bursts, of harder bursts being more inhomogeneous in their space distribution, which holds for hardnesses up to ∼4, with a slight rise of $\langle V/V_{\rm max}\rangle$ beyond that point. In the three-class classification that we propose here this trend, however, exists only for the longest class.

In the following we first describe the two methods (cluster analysis and neural network) that are used to classify GRBs in the present work. The results from the two different methods are then discussed and compared. We find that both methods point to a classification in three classes, that resulting from splitting the 'old' long-burst class into two, and we compare the characteristics of the three new groups. Finally, the possible physical meaning of the new classification is discussed.

2 METHODOLOGY

The usual approach in GRB classification has been based on the study of bivariate distributions. However, as noted by Bagoly et al. (1998, hereafter B98), the BATSE catalogue provides up to nine quantities intrinsic to the burst (seven related to energy and two related to duration), plus other quantities corresponding to spatial distribution and to errors in the magnitudes. New composite quantities can also be defined, such as the different measures of spectral hardness (from the fluence ratios in different spectral channels), and also $V/V_{\rm max}$. That involves a large number of variables which is difficult to handle, complex relationships among them (including non-linear ones) probably being present. Such relationships can hardly be directly visualized, and thus multivariate analysis is needed.

Starting from the BATSE Current GRB Catalog (available at http://www.batse.msfc.nasa.gov/batse/grb/catalog/current) in its version of 2000 September, 1599 bursts have been selected: those for which non-zero values of all nine magnitudes are given. These magnitudes are the four time-integrated fluences $F_{\rm Ch1}$–$F_{\rm Ch4}$, respectively corresponding to the 20–50, 50–100, 100–300 and 300+ keV spectral channels; the three peak fluxes P_{64}, P_{256} and P_{1024}, measured in 64-, 256- and 1024-ms bins, respectively; and the two measures of burst duration T_{50} and T_{90}, the times within which 50 and 90 per cent of the flux arrives. Then a principal component analysis (PCA) of the standardized logarithms (zero mean and unity variance) of those quantities has been performed, obtaining results (Table 1) that are very similar to those of B98. As is well known, PCA is a statistical method used in multivariate data analysis to obtain new variables, linear combinations of the original ones, which carry most of the variance of the system. Based on the correlations among the original variables, some of the new variables can be disregarded if they carry very little information. For further details on PCA, the reader is referred to B98 and to Murtagh & Heck (1987, hereafter M87).

As seen in Table 1, with only three variables, linear combinations of the original ones, ∼96 per cent of the system information can be accounted for. The first row shows that 64 per cent of the information is carried by a variable that is a weighted sum of all the original variables with nearly the same weight for each of them. The second principal component in importance is approximately the difference between the weighted sum of the logarithms of the three peak fluxes and that of the logarithms of the two durations, all again with similar weights. With 5 per cent of the total variance of the system, the logarithm of the fluence in the fourth channel is found.

Our current goal is to achieve an automatic classification based on the nine original variables, and for that two different methods are used: a cluster analysis applied to the results of the PCA, and a neural network algorithm.

2.1 Cluster analysis

For the cluster analysis the MIDAS statistical package has been used. As stated above, a PCA is first performed. In this way are obtained new variables into which the problem becomes easier to separate. This result provides the starting point for the cluster analysis, where Ward's criterion of minimum variance (Ward 1963; see also M87) is used. The analysis follows an agglomerative hierarchical clustering procedure, which starts from n points spread over the nine-dimensional space and groups them until it ends up with a single cluster. The algorithm searches for clusters with minimum variance among objects belonging to the same cluster and with maximum variance between clusters, and works with the

'centre of gravity' of each cluster. That gives clusters as compact and as detached from each other as possible.

A dendrogram is obtained, which shows the way in which groups are clustering, as well as the inner variance of the resulting groups. Thus detecting a large rise in the variance by the union of two clusters means that two groups with remarkably different characteristics have been merged.

It is important to notice that the PCA looks for combinations of variables to obtain axes that have the maximum possible variance. As seen above, with just three of those axes more than 95 per cent of the total variance is accounted for. The cluster analysis could, therefore, have been applied to those three new variables only. This has been checked and it is found that the results are barely affected by this reduction in the number of variables, since most of the system information is conserved. Here, however, the results of the complete nine-dimensional analysis will be presented, since those same nine variables will later be used for the neural network analysis.

Concerning the previous work of Mukherjee et al. (1998, hereafter M98), it must be noted that they chose six variables for their analysis, three of them being the peak flux $\log P_{256}$ plus the two durations ($\log T_{50}$ and $\log T_{90}$), the other three being the total fluence $\log F_{\rm total}$ and two hardnesses, $\log H_{321}$ and $\log H_{32}$. It has been learned, from the PCA, that three variables are necessary which together carry more than 95 per cent of the variance, one of them being approximately the logarithm of the fluence in the fourth channel, $F_{\rm Ch4}$ (see above). So it seems that M98 do lose information by not taking into account the fluence in the fourth channel separately, and also by not considering any peak flux later on. It must equally be noticed, when comparing results, that our GRB sample is twice as large as theirs.

The main weakness of the cluster analysis is that it only deals with linear combinations of the variables. Such a weakness can be overcome by means of a neural network analysis, which also detects non-linear relationships.

2.2 Neural network

Neural networks are artificial intelligence algorithms that can be used for an automatic and objective classification. We do not want to start from any prior classification. Therefore a non-supervised algorithm is used. As we do not wish to introduce any tracer object either, the net is initialized at random. The 'Self-Organizing Map' algorithm (Kohonen 1990), implemented in the SOM_PAK package from the Laboratory of Computer and Information Science of the University of Helsinki, is used.

As in the cluster analysis, the entrance parameters are the logarithms of the same nine variables.

The dimension of the output space must be specified, and based on the results of the cluster analysis the network is run twice, asking first for a two- and then for a three-dimensional output space, thus grouping either two or three classes of GRBs. The net is trained in two steps before looking for results.

3 RESULTS

3.1 Cluster analysis

In Fig. 1 the dendrogram with the last six levels of clustering is shown. It can be seen that the first important increase of the variance occurs when joining group 3 with group 2, which tells us that two groups with somewhat different characteristics have been merged, but the most significant rise in variance occurs when merging cluster 2 with cluster 1. From that it is concluded that there are two well-separated classes plus an emergent third class.

Fig. 2 shows what happens when adding, to the nine starting variables, the two extra variables H_{32} and $\langle V/V_{\rm max}\rangle$. In that case the sample is reduced to 757 bursts only (instead of 1599), for which all 11 quantities are known. It can be seen that the three-class classification is the most favoured one.

Next, in Table 2, the main characteristics of each GRB class are shown, 2-I and 2-II corresponding to the two-class classification, and 3-I, 3-II and 3-III corresponding to the three-class classification. The deviations correspond to $\sigma/\sqrt{N-1}$. The results of the 11-dimensional cluster analysis are not shown here. They are

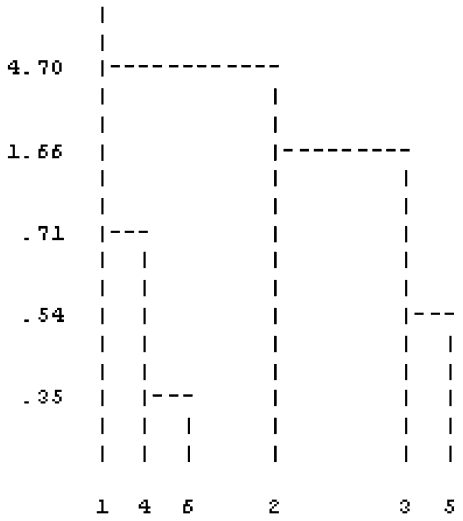

Figure 1. Dendrogram of the nine-dimensional analysis. The numbers at the bottom of the diagram are identifiers of the groups, and those at the left show the inner variance of the groups. For instance, when merging group 6 with group 4 the variance of the cluster is 0.35. The largest increment in the variance occurs when merging group 2 with group 1, with a variance increase of 3.04.

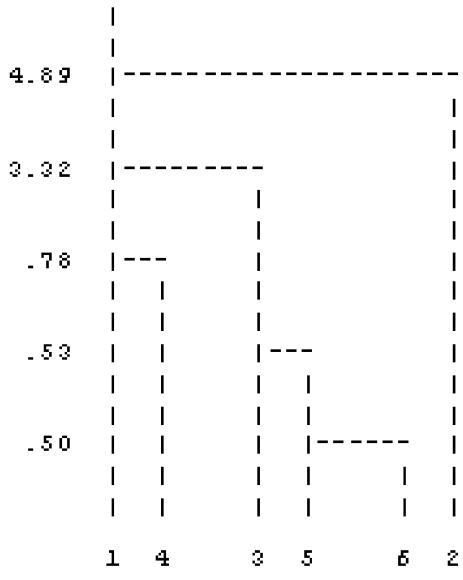

Figure 2. Dendrogram of the 11-dimensional analysis. Here the largest increase in the cluster variance occurs when joining groups 3 and 1, suggesting three different classes of GRBs.

Table 2. Characteristics of the classification from the nine-dimensional cluster analysis. T_{90} is in units of s, P_{1024} in units of photon $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, and F_{total} in units of $10^{-6}\,\mathrm{erg\,cm}^{-2}$.

Class	N	$\langle T_{90}\rangle$	$\langle H_{32}\rangle$	$\langle V/V_{\max}\rangle$	$\langle P_{1024}\rangle$	$\langle F_{\mathrm{total}}\rangle$	$\langle\cos\theta\rangle$	$\langle\sin^2 b - 1/3\rangle$
2-I	580	2.65 ± 0.17	5.96 ± 0.20	0.265 ± 0.017	1.29 ± 0.08	1.75 ± 0.13	-0.031 ± 0.026	-0.006 ± 0.013
2-II	1019	59.7 ± 2.1	3.11 ± 0.05	0.184 ± 0.008	3.33 ± 0.20	22.5 ± 1.8	-0.004 ± 0.019	$+0.001 \pm 0.010$
3-I	580	2.65 ± 0.17	5.96 ± 0.20	0.265 ± 0.017	1.29 ± 0.08	1.75 ± 0.13	-0.031 ± 0.026	-0.006 ± 0.013
3-II	570	51.3 ± 2.3	2.85 ± 0.07	0.296 ± 0.012	0.88 ± 0.02	4.58 ± 0.21	$+0.021 \pm 0.025$	-0.000 ± 0.013
3-III	449	70.3 ± 3.8	3.43 ± 0.06	0.051 ± 0.004	6.44 ± 0.41	45.3 ± 3.9	-0.035 ± 0.030	$+0.002 \pm 0.015$

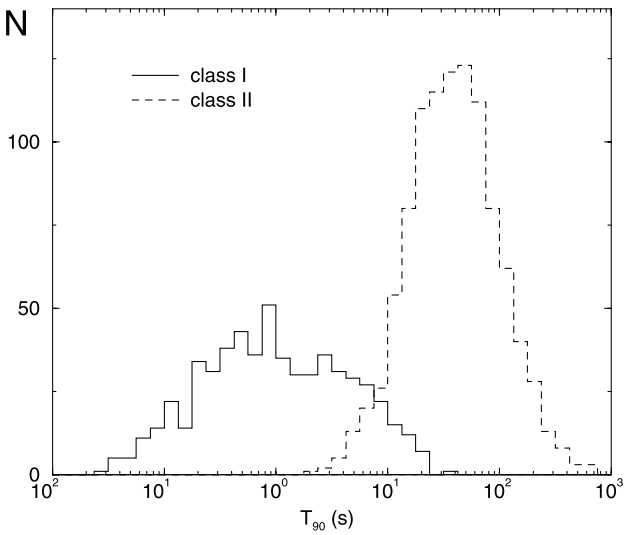

Figure 3. Duration distributions of classes 2-I and 2-II.

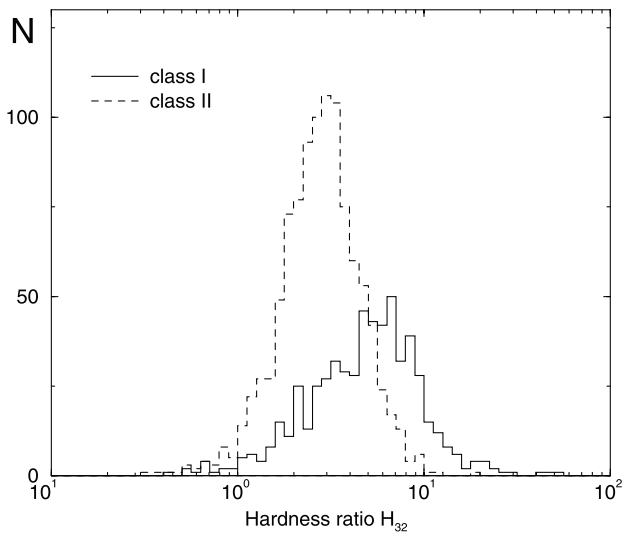

Figure 4. Hardness distributions of classes 2-I and 2-II.

very similar to those of the nine-dimensional one but are less significant because the sample is reduced to one-half. We only comment that since the hardness has been added there, its weight has thus been enforced, and then class I becomes slightly harder and shorter than when obtained from nine variables.

It must be noted that, by just looking at the values of the dispersions in the four variables T_{90}, H_{32}, P_{1024} and F_{total} given in Table 2, it might seem that the variance would increase when shifting from the two-class to the three-class classification, but that is just an effect of projecting the groups on to these particular variables (three of them composites); the full nine-dimensional analysis shows the opposite, as stated above.

Adopting the same expected values for isotropy as for the 4B catalogue (Fishman et al. 1999), that is for the Galactic dipole moment $\langle\cos\theta\rangle = -0.009$ and for the quadrupole Galactic moment $\langle\sin^2 b - 1/3\rangle = -0.004$, it can be seen in Table 2 that only one of the corresponding values for classes 3-I–3-III lies beyond 1σ of the expected value, and that is the dipole for the 3-II class, which is $+1.2\sigma$ above. Just this value being above 1σ appears not to be significant and it is concluded that all three classes are isotropically distributed.

In calculating the $\langle V/V_{\max}\rangle$ parameter, not all the 1599 bursts could be used, but only those for which that value could be derived, and in a similar way when calculating the dipole and quadrupole moments the GRBs that were overwrites (Fishman et al. 1999) were not taken into account.

In Table 2 the class numbers are given in order of increasing durations $\langle T_{90}\rangle$. With the two-class classification the 'classical' GRB types are recovered: short/hard, which are fainter (taking as brightness the peak flux $\langle P_{1024}\rangle$); and long/soft, which are brighter and more non-Euclidean in their space distribution. As seen from Fig. 3, two classes with an overlapping distribution of durations have been obtained, in contrast with the classical definition of short ($T_{90} < 2$ s) and long ($T_{90} > 2$ s) GRBs. Now the short class has durations up to $\sim$20 s while the long-duration class starts at $\sim$2 s. This overlapping of the two classes was obviously supposed to exist, but based on the distribution of durations alone it could not be decided whether, in the overlapping region, a given GRB belonged to either of the two classes. Now the algorithm handles all the available magnitudes and assigns each GRB to the cluster to the characteristics of which it is closer. The hardness distribution (Fig. 4) does not differ significantly from that in K93.

A first look at the three-class classification reveals that class I is exactly the same as in the two-class grouping: this is because the clustering method is agglomerative, which means that new groups are formed by merging former ones, so the passage from three to two classes happens when merging class II and class III GRBs.

Let us concentrate on the new three-class classification. As stated already, 'old' class II has been divided into class II and class III. Class II is not properly an intermediate class: it has intermediate duration but still of the same order of magnitude as class III, and with an almost coincident distribution, as seen in Fig. 5. Class II is the softest and faintest class and the one most homogeneously distributed in space. Despite its duration being of the same order as that of class III, the fluence is one order of magnitude lower.

The most striking result of this new classification is the extremely low value of $\langle V/V_{\max}\rangle$ in class III, which means that we are dealing with an extremely deep population that extends up to

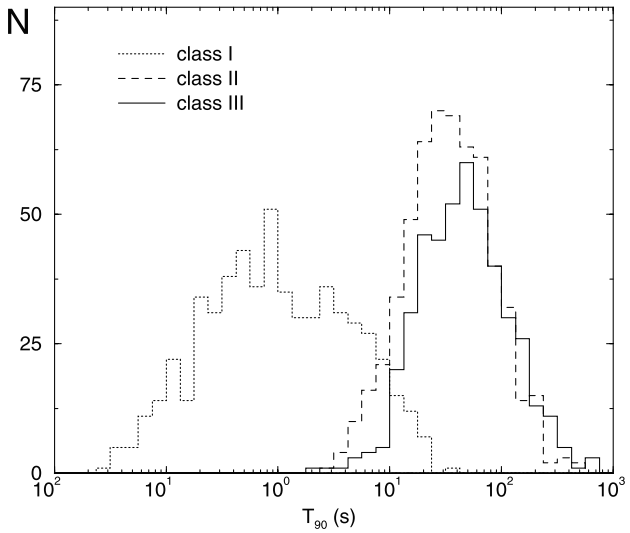

Figure 5. Duration distributions of the three-class classification from cluster analysis.

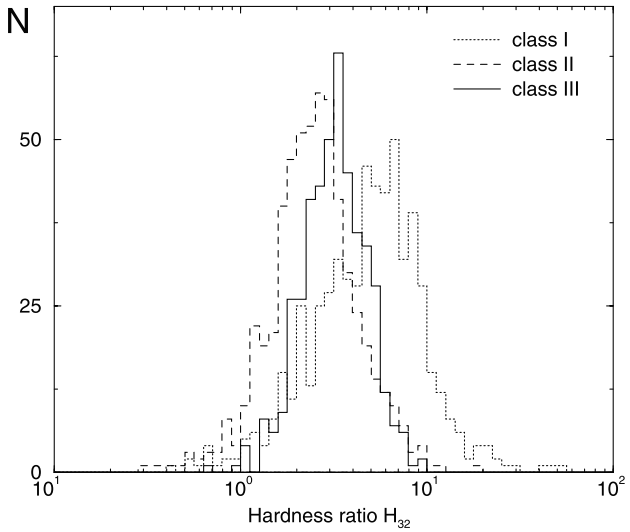

Figure 6. Hardness distributions of the three-class classification from cluster analysis.

very high redshifts. Following the same procedure as in Mao & Paczyński (1992), that is to calculate the theoretical value of $\langle (F_{\min}/F)^{3/2} \rangle$ as a function of $z_{\max}$, and taking $\langle V/V_{\max} \rangle$ in Table 2 as an empirical value for that quantity, a value of $z_{\max}$ for the distribution of the GRBs can be derived. Adopting a model universe with $\Omega_{\rm M} = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 65\,{\rm km\,s^{-1}\,Mpc^{-1}}$, assuming the GRBs to be standard candles with a spectral slope $\alpha = 1$ (Mallozzi Pendleton & Paciesas 1996), and assuming constant comoving GRB rate, one obtains $z_{\max} = 4.06^{+0.66}_{-0.57}$ for class I, $3.08^{+0.35}_{-0.32}$ for class II and $45.24^{+4.23}_{-3.55}$ for class III. The latter is an exceedingly high value, but, as will be seen in Section 4, class III can have very massive stars as progenitors, and in that case the GRB rate should be proportional to the star formation rate (SFR) rather than being constant. Taking as SFR(z) that of Madau & Pozzetti (2000), $z_{\max} = 11.30^{+0.56}_{-0.43}$ is obtained for class III, which is a more reasonable value.

It was already known that separating long-class GRBs into two hardness groups results in two very different degrees of inhomogeneity (T98). The $\langle V/V_{\max} \rangle$ values for the whole Current GRB Catalog have been calculated and the results are $\langle V/V_{\max} \rangle = 0.268 \pm 0.011$ for GRBs with $H_{32} < 3$, and $\langle V/V_{\max} \rangle = 0.182 \pm 0.012$ for GRBs with $H_{32} > 3$. Now, with the three-class classification, very different degrees of inhomogeneity within the same interval of hardness are found, as can be seen by comparing Table 2 with Fig. 6, and with little difference in mean hardness: those of class II and class III differ here with 4×10^{-3} significance in a Student test, but in the neural network classification the significance will only be 0.42 while the $\langle V/V_{\max} \rangle$ difference will remain. We leave a discussion of the value of $\langle V/V_{\max} \rangle$ as related to hardness for Section 4.

It can also be pointed out that, in contrast with what happened in the binary long/short classification, where shorter bursts were harder, now class II is shorter than class III but it is slightly softer at the same time.

3.2 Neural network

In the neural network case, how many classes are to be obtained must be decided beforehand: knowing the dendrograms that result from the cluster analysis, we ask for either two or three classes. Their main characteristics are summarized in Table 3, using the same units as in Table 2.

As should be expected, there are some differences in the composition of the classes as compared with those obtained from the clustering method, since the neural network method is not agglomerative. So, for instance, class I is no longer identical, in the two-group classification, to class I in the three-group scheme. Also, the 'short' GRBs which make up this class now have longer average durations than in the cluster analysis.

There is also some change from the results of the cluster analysis in the three-group classification. Classes II and III now become more widely separated in duration (Fig. 7), basically because of the decrease in duration of class II. The difference in hardness between classes II and III, in contrast, has decreased (Fig. 8).

As in the cluster analysis, all three classes are highly isotropic, with no value of the moments above 0.8σ of the values expected for isotropy.

Class II is now the intermediate class in peak flux. From both methods, cluster analysis and neural network, it is seen that, despite the difference of one order of magnitude between the durations of classes I and II, their respective total fluences remain of the same order.

The high inhomogeneity in the space distribution of class III is seen once more, and also how class II is again the most homogeneous one. Now the sample depths are, for constant comoving GRB rate, $z_{\max} = 3.34^{+0.55}_{-0.48}$, $2.78^{+0.53}_{-0.46}$ and $15.45^{+1.50}_{-1.33}$ for classes I, II and III respectively. For GRB rate proportional to SFR, $z_{\max} = 3.22^{+0.22}_{-0.20}$, $2.98^{+0.23}_{-0.20}$ and $6.67^{+0.30}_{-0.30}$. The value of $z_{\max} \sim 11.3$, obtained from the cluster analysis, corresponds to a universe with an age of 4.3×10^8 yr, and is in good agreement with the expectation of GRBs occurring out to at least $z \approx 10$. The value obtained from the neural network analysis of $z_{\max} \sim 6.7$ corresponds to a universe with an age of 8.8×10^8 yr. Both values for $z_{\max}$ are below the redshift limit, $z \approx 15$–20, given for Population III stars by LR00.

We conclude that the three classes respectively obtained from the cluster analysis and from the neural network algorithm show similar characteristics, and thus both treatments are mutually consistent.

Table 3. Characteristics of the classification with the neural network. T_{90} is in units of s, P_{1024} in units of photon cm^{-2} s^{-1}, and $F_{\rm total}$ in units of 10^{-6} erg cm^{-2}.

Class	N	$\langle T_{90}\rangle$	$\langle H_{32}\rangle$	$\langle V/V_{\rm max}\rangle$	$\langle P_{1024}\rangle$	$\langle F_{\rm total}\rangle$	$\langle\cos\theta\rangle$	$\langle\sin^2 b - 1/3\rangle$
2-I	685	6.24 ± 0.50	5.50 ± 0.18	0.288 ± 0.015	0.94 ± 0.04	1.44 ± 0.09	$+0.002 \pm 0.024$	-0.005 ± 0.012
2-II	914	63.5 ± 2.3	3.12 ± 0.05	0.159 ± 0.008	3.82 ± 0.22	25.1 ± 2.0	-0.024 ± 0.021	$+0.001 \pm 0.010$
3-I	531	3.05 ± 0.34	6.20 ± 0.22	0.287 ± 0.017	0.81 ± 0.04	1.13 ± 0.07	-0.003 ± 0.027	-0.014 ± 0.014
3-II	341	25.0 ± 1.4	3.05 ± 0.10	0.307 ± 0.019	1.25 ± 0.08	2.82 ± 0.16	-0.012 ± 0.033	$+0.009 \pm 0.016$
3-III	727	71.8 ± 2.8	3.15 ± 0.05	0.123 ± 0.008	4.51 ± 0.28	30.8 ± 2.5	-0.022 ± 0.023	$+0.003 \pm 0.012$

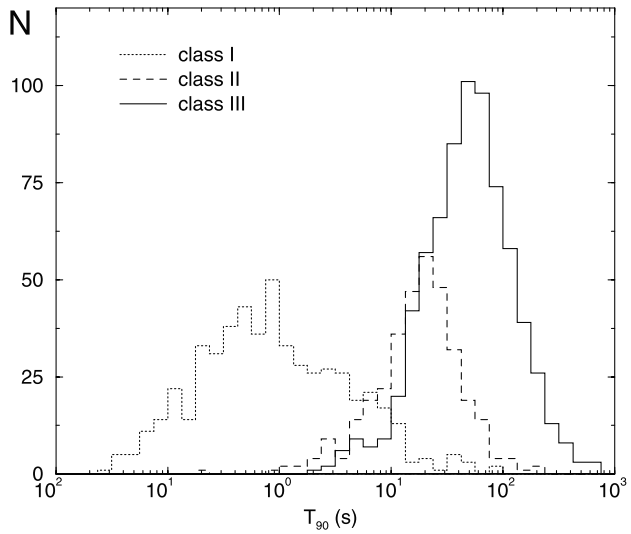

Figure 7. Duration distributions of the three-class classification from neural network analysis.

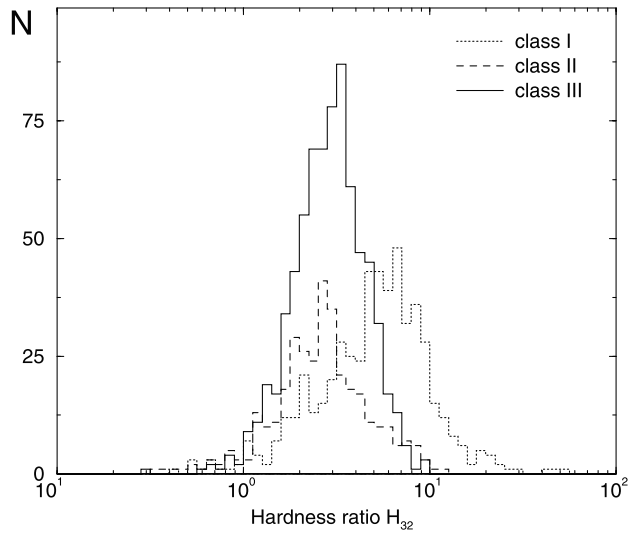

Figure 8. Hardness distributions of the three-class classification from neural network analysis.

4 DISCUSSION

The two different automatic classifier methods above suggest the existence of three different groups of GRBs with different properties. We will now examine whether such a classification does make physical sense and we are actually dealing with three classes of GRBs.

Let us first discuss the proposal by T98, of taking the long-burst class and dividing it into two groups with H_{32} higher and lower than 3 respectively. As seen in Section 3.1, long/hard bursts are more inhomogeneously distributed than long/soft bursts. This might seem to be in contradiction with the cosmological scenario, in which more distant bursts are expected to be softer because of the spectrum redshift, and it leads us to conclude that class III bursts are intrinsically much harder than those of classes I and II.

Fig. 9 shows that there is indeed evolution in hardness: the value of $\langle V/V_{\rm max}\rangle$ decreases with increasing H_{32}. The hardness bins are taken so as to include similar numbers of bursts ($\sim$60) in each of them, in order to have comparable error bars. The value of $\langle V/V_{\rm max}\rangle$ is displayed in the position of the mean of the hardness for each bin, and no error bars for the hardness are shown because the deviation is less than the symbol size, except for the last bin for which it is about 0.2.

When $\langle V/V_{\rm max}\rangle$ decreases one is dealing with a more distant sample of GRBs, and then Fig. 9 tells us that, when sampling to higher distances, GRBs tend to be harder, and taking into account the hardness–intensity correlation (Dezalay et al. 1997) they should also be more luminous. This effect has to be interpreted, in a cosmological scenario, as a source evolution. There is a possible explanation: it is generally admitted that the upper limit of the stellar initial mass function (IMF) depends on metallicity, and that lower metallicity allows more massive stars to form. When sampling GRBs farther away, one looks to a younger universe, with lower metallicity, and thus with more massive stars. Therefore, if GRBs come from very massive stars, those ancient GRBs had sources with higher power and they were brighter and harder.

Next are displayed, in Figs 10 and 11, $\langle V/V_{\rm max}\rangle$ versus hardness for classes II and III, from the results of the cluster and of the neural network analyses, respectively.

Here can be appreciated, with particular clarity in the figure corresponding to the neural network result, how in the three-class scheme the 'new' class II no longer shows any trend of $\langle V/V_{\rm max}\rangle$ decreasing with increasing H_{32}. Such a trend in the 'old' class II was due to the fusion into it of the 'new' classes II and III, and now it is seen that class III is the only one to uphold the trend. To evaluate the correlation numerically, a Spearman rank test (see, for instance, its implementation in Press et al. 1992) has been applied, obtaining for class III a Spearman rank correlation coefficient $r_{\rm s} = -0.344$, with a significance of 2×10^{-8}, for the class resulting from the cluster analysis, and $r_{\rm s} = -0.354$, with significance 4×10^{-13}, for that resulting from the neural network analysis. In contrast, for class II from the cluster analysis $r_{\rm s} = 0.066$ is obtained, with a significance of 0.26, and for class II from the neural network $r_{\rm s} = -0.051$ is obtained, with a 0.51 significance level. It is concluded, therefore, that class III really shows clues of cosmological source evolution, which could be due

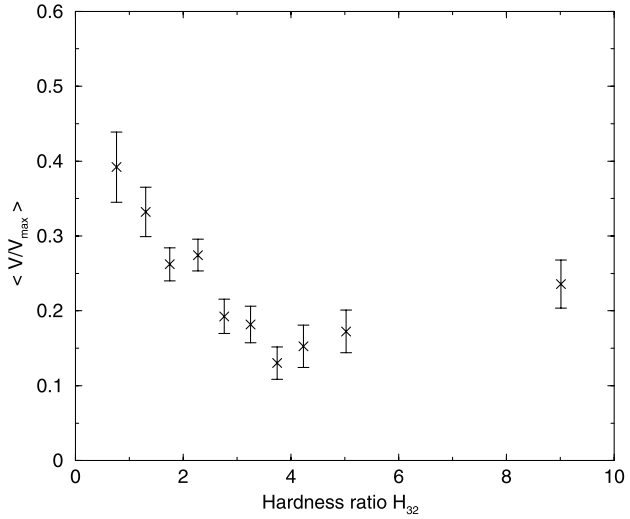

Figure 9. $\langle V/V_{\rm max}\rangle$ versus hardness for GRBs with $T_{90} > 2$ s. The correlation between these two variables is clearly seen. Hardness bins have been taken so as to include similar numbers of bursts in all of them. Each hardness bin contains ~ 60 GRBs.

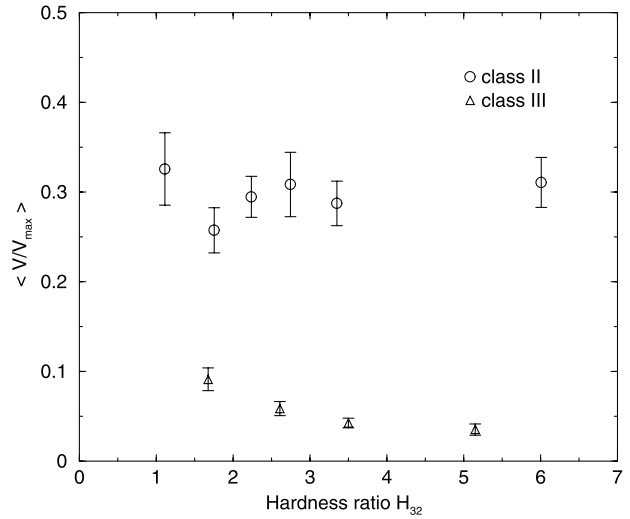

Figure 10. $\langle V/V_{\rm max}\rangle$ versus hardness for classes II and III from the cluster analysis. Here the nearly constant value of $\langle V/V_{\rm max}\rangle$ over the interval of hardnesses covered by class II is seen. The trend for class III of lower $\langle V/V_{\rm max}\rangle$ with higher H_{32} can also be seen. As in Fig. 9, the hardness bins have been taken so as to include similar numbers of bursts in all of them.

to its being composed of GRBs produced by very massive stars, probably through collapsars.

Out of all the bursts with known redshifts, eight entered into our classification. In the case of the neural network classification, seven out of those eight bursts are classified as class III, and only one of them as class II. With the clustering method classification, six belong to class III and two to class II. One of these last two bursts is GRB 980425, presumably related to SN 1998bw (Galama et al. 1999), which is thought to be a peculiar GRB. In both classifications GRB 970508 (Metzger et al. 1997) was assigned to class II. The assignation of any individual GRB to a given class by our methods is not entirely reliable, however, and has an uncertainty that is most important in the limiting region of each cluster. By looking at the scatter graphs, it has been checked that GRB 970508 in fact lies in the region near class III. It must be

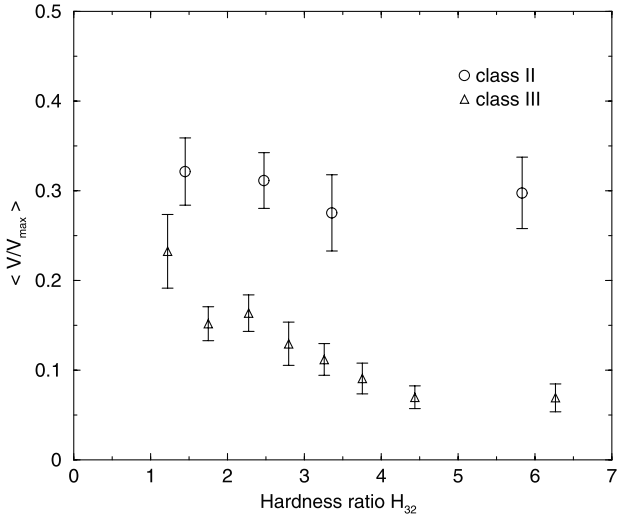

Figure 11. $\langle V/V_{\rm max}\rangle$ versus hardness for classes II and III from neural network analysis (same criterion as in Fig. 10 for the hardness bins). In this graph, the trend for class III of lower $\langle V/V_{\rm max}\rangle$ with higher H_{32} is more evident than in Fig. 10.

stressed that, given the high isotropy of all three classes found here, there is no evidence of Galactic structure for any of them. It could be argued that, in the same way as no redshift has yet been measured for any GRB of class I, because the fact that they are short makes them difficult to detect with *BeppoSAX* (which is mainly sensitive to bursts longer than about 5–10 s), no redshifts of GRBs of class II have been measured either, because they are faint and their detection is equally hard. As an alternative explanation, the fact that no afterglows from GRBs of class I or II have been seen is due to their being produced by neutron star–neutron star or neutron star–black hole mergers, which are expected to happen mostly outside galaxies where the interstellar medium is too tenuous to produce any detectable afterglow (see Panaitescu, Kumar & Narayan 2001, for instance). That would also be consistent with the $\langle V/V_{\rm max}\rangle$ values found: while collapsars should appear first and be more frequent in the distant, early Universe, neutron star–neutron star and neutron star–neutron black hole mergers should start later and be more homogeneously distributed down to low redshifts. In this context we can even speculate whether the differences between class I and class II GRBs might be due to one of them corresponding to black hole–neutron star (or black hole–white dwarf) mergers, the other class being produced by neutron star–neutron star mergers. The whole question should be solved with the new generation of GRB detectors aboard the *HETE 2* and *Swift* satellites.

One should be concerned whether the structure of the GRB data may partially reflect instrumental biases. Hakkila et al. (2000) have suggested that the three-class classification obtained by M98 might arise from a bias in measuring some burst properties, such as duration and fluence, which would make some bursts in 'classical' class II take on 'new' class II characteristics (by lowering their duration and fluence). The fluence–duration bias, however, in spite of being qualitatively understood, is not well quantified. Hakkila et al. based their analysis on M98's classes: their intermediate-duration class has durations T_{90} of between 2 and 10 s, while the class II deduced here extends up to T_{90} longer than 100 s. Moreover, such bias acts on the farthest bursts, while what we find is that our 'new' class II GRBs are the closest ones. In addition, any bias that would make some bursts in the 'old' class II appear

shorter and with lower fluence could hardly separate at the same time the evolutionary effects that we see in Fig. 9 into two groups: one with evolution (class III) and the other one without it ('new' class II).

5 CONCLUSION

There are reasons to think that there exists more than one type of possible progenitor for GRBs, and each type may give rise to groups of burst with different properties. We have searched for these groups in the current BATSE catalogue, with the aid of two automated classification algorithms, and confirmed that there exist two clearly separate classes of GRBs corresponding to the 'classical' classification of long/short GRBs. In addition, we have also obtained clear hints that there exists a third class, different from those previously reported. An oversimplified way of looking at this would be to say that the third class arises from splitting the original long class into two groups with high and low peak fluxes, similar to the way in which the whole sample of GRBs has been divided, in previous studies, into pairs of groups according to duration (K93), hardness (T98), brightness (Nemiroff et al. 1994), or other characteristics (Pendleton et al. 1997). The present work, however, goes beyond that, since nine quantities related to the bursts are used for the classification instead of taking a single parameter and then finding a value separating the bursts into two classes: there are overlapping zones in every original variable. What our procedures do is trace a surface in the nine-dimensional space, separating classes from the way in which each variable relates to all others. Two different classes may well have the same duration or show nearly the same distribution for a given variable, but by taking into account the other variables as well, these procedures still detect their existence. In contrast, univariate distributions would overlook them.

Apart from the power of the method, the new grouping of the bursts thus obtained has to be examined for its possible physical meaning and its correspondence with separate classes of GRB progenitors and/or mechanisms. Classes I, II and III here defined correspond to different observational depths ($z_{\max}$), and may result from varying geometries of the observer with respect to the emitter, different parameters of the explosion, or different progenitors having different spatial distributions. Thus every class has to be compared with several possible models. The physical separation of classes II and III is strongly supported by the fact (which can hardly be due to chance alone) that when considering both classes together they show evolution of hardness and intensity with the maximum distance sampled, while when separated such evolution exists only in class III. We conclude, therefore, that class III, which probably has collapsars as progenitors, is the one that can be detected up to very large redshifts, and it should thus be the most suitable one to learn about the history of the Universe at high z. We also suggest that classes I and II could correspond to neutron star–neutron star or neutron star–black hole mergers instead.

ACKNOWLEDGMENTS

We thank J. Bloom for suggestions and comments. AB thanks R. Cabezón for helpful discussions and comments.

REFERENCES

Bagoly Z., Mészáros A., Horváth I., Balázs L. G., Mészáros P., 1998, ApJ, 498, 342 (B98)
Berezinsky V., Hnatyk B., Vilenkin A., 2001, Phys. Rev. D., 64, 043004
Blinnikov S., 1999, astro-ph/9902305
Bloom J. S., Kulkarni S. R., Djorgovski S. G., 2000, astro-ph/0010176
Castro-Tirado A., 2001, Exploring the Gamma-ray Universe. in press, astro-ph/0102122
Daigne F., Mochkovitch R., 1998, MNRAS, 296, 275
Dezalay J. P. et al., 1997, ApJ, 490, L17
Fishman G. J. et al., 1999, ApJS, 122, 465
Galama T. J. et al., 1999, A&AS, 138, 465
Hakkila J., Haglin D. J., Pendleton G. N., Mallozzi R. S., Meegan C. A., Roiger R. J., 2000, ApJ, 538, 165
Horváth I., 1998, ApJ, 508, 757
Katz J. I., Canel L. M., 1996, ApJ, 471, 915
Klebesadel R. W., Strong I. B., Olson R. A., 1973, ApJ, 182, L85
Kohonen T., 1990, IEEC Proc., 78, 1464
Kouveliotou C., Meegan C. A., Fishman G. J., Bhat N. P., Briggs M. S., Koshut T. M., Paciesas W. S., Pendleton G. N., 1993, ApJ, 413, L101 (K93)
Lamb D. Q., Reichart D. E., 2000, ApJ, 536, 1 (LR00)
Ma F., Xie B., 1996, ApJ, 462, L63
MacFadyen A. I., Woosley S. E., 1999, ApJ, 524, 262
Madau P., Pozzetti L., 2000, MNRAS, 312, L9
Mallozzi R. S., Pendleton G. N., Paciesas W. S., 1996, ApJ, 471, 636
Mao S., Paczyński B., 1992, ApJ, 388, L45
Meegan C. A., Fishman G. J., Wilson R. B., 1985, ApJ, 291, 479
Mészáros P., 2001, Sci, 291, 79
Mészáros P., Rees M. J., 1993, ApJ, 405, 278
Metzger M. R. et al., 1997, IAU Circ. 6655
Mukherjee S., Feigelson E. D., Babu G. J., Murtagh F., Fraley C., Raftery A., 1998, ApJ, 508, 314, (M98)
Murtagh F., Heck A., 1987, Multivariate Data Analysis, Astrophysics and Space Science Library. Reidel, Dordrecht (M87)
Narayan R., Paczyński B., Piran T., 1992, ApJ, 395, L83
Nemiroff R. J., 1994, astro-ph/9402012
Nemiroff R. J., Norris J. P., Bonnell J. T., Wickramasinghe W. A. D. T., Kouveliotou C., Paciesas W. S., Fishman G. J., Meegan C. A., 1994, ApJ, 435, L133
Paczyński B., 1990, ApJ, 363, 218
Paczyński B., 1998, in Meegan C. A., Preece R. D., Koshut T. M., eds, 4th Huntsville Symp., Gamma Ray Bursts. Am. Inst. Phys., New York, p. 783
Panaitescu A., Kumar P., Narayan R., 2001, astro-ph/0108132
Pendleton G. N. et al., 1997, ApJ, 489, 175
Piran T., 2000, Phys. Rep., 333, 529
Press W. H. et al., 1992, Numerical Recipes in Fortran, 2nd edn. Cambridge Univ. Press, Cambridge
Rees M. J., Mészáros P., 1992, MNRAS, 258, 41P
Spruit H. C., 1999, A&A, 341, L1
Tavani M., 1998, ApJ, 497, L21 (T98)
Totani T., 1997, ApJ, 486, L71
Totani T., 1999, ApJ, 511, 41
Ward J. H., 1963, in Babu G. J., Feigelson E. D., eds, Statistical Challenges of Modern Astronomy. Springer, New York, p. 135
Woosley S. E., 1993, ApJ, 405, 273

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