Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **455,** 703–711 (2016) doi:10.1093/mnras/stv2350

Investigation of redshift- and duration-dependent clustering of gamma-ray bursts

T. N. Ukwatta★† and P. R. Woźniak★

Space and Remote Sensing (ISR-2), Los Alamos National Laboratory, Los Alamos, NM 87544, USA

Accepted 2015 October 7. Received 2015 September 22; in original form 2015 July 25

ABSTRACT

Gamma-ray bursts (GRBs) are detectable out to very large distances and as such are potentially powerful cosmological probes. Historically, the angular distribution of GRBs provided important information about their origin and physical properties. As a general population, GRBs are distributed isotropically across the sky. However, there are published reports that once binned by duration or redshift, GRBs display significant clustering. We have studied the redshift- and duration-dependent clustering of GRBs using proximity measures and kernel density estimation. Utilizing bursts detected by Burst and Transient Source Experiment, *Fermi*/gamma-ray burst monitor, and *Swift*/Burst Alert Telescope, we found marginal evidence for clustering in very short duration GRBs lasting less than 100 ms. Our analysis provides little evidence for significant redshift-dependent clustering of GRBs.

Key words: gamma-ray burst: general.

1 INTRODUCTION

Gamma-ray bursts (GRBs) are often referred to as the biggest explosions in the Universe since the big bang. Their powerful prompt electromagnetic emission and afterglows make them detectable out to very high redshifts $z > 10$ (Lamb & Reichart 2000). Therefore, GRBs offer a potential probe to study inhomogeneities and anisotropies in the Universe on the largest scales.

Historically, the angular distribution of GRBs provided important information about their origin and physics. These early studies have shown that the sky distribution of GRBs is isotropic (Meegan et al. 1992; Briggs et al. 1996; Tegmark et al. 1996), providing early indications – before the discovery of afterglows – that GRBs are at cosmological distances. While the isotropic sky distribution is well established for long-duration GRBs, there have been reports of clustering for short GRBs (T90 $<$ 2 s; Balazs, Meszaros & Horvath 1998; Magliocchetti, Ghirlanda & Celotti 2003), very short GRBs (T90 $<$ 100 ms, VSGRB; Cline, Matthey & Otwinowski 1999; Cline et al. 2005; Cline 2011), and intermediate-duration GRBs (2 s $<$ T90 $<$ 8 s; Mészáros et al. 2000; Litvin et al. 2001). Other reports of clustering are based on parameters such as spectral lags (Ukwatta et al. 2010, 2012). For example, long-lag GRBs tend to cluster in the supergalactic plane (Norris 2002; Foley et al. 2008). If confirmed, these observations may point to the existence of new subpopulations of GRBs, possibly with distinct progenitors.

Thanks to the rapid localizations delivered by the *Swift* mission (Gehrels et al. 2004), we now have more than 300 GRBs with redshift measurements. This opens up the exciting possibility to explore the distant universe using high-redshift GRBs. Recently, Horváth, Hakkila & Bagoly (2014) have reported evidence of strong anisotropy in the observed sky distribution of GRBs in the redshift range between 1.6 and 2.1 based on the distribution of the angular distance to the n-th nearest neighbour and a two-dimensional Kolmogorov–Smirnov (KS) test. Detection of very large scale structures such as the one implied by the Horváth et al. (2014) study would have profound implications for the cosmological principle which states that on average the Universe is homogeneous and isotropic.

Here, we investigate the angular sky distribution of GRB subpopulations using various density and proximity estimators. Our study is divided into two parts. First, we investigate redshift-dependent GRB clustering using a sample of 311 *Swift* bursts with measured redshifts. We use full Monte Carlo simulations to generate the relevant probability distributions and evaluate the significance of potential angular structures. This analysis is limited to bursts detected by *Swift* because an overwhelming majority of GRBs with redshift measurements are in that category and also because of the possibility to correct the density estimates to account for irregularities in the exposure time allocated to various parts of the sky. In the second part of the paper, we use combined samples of GRBs detected by multiple instruments to study duration-dependent angular distributions. To examine the significance of clustering results, we construct approximate exposure maps for multi-instrument samples assuming that the intrinsic distribution of long GRBs is isotropic.

★ E-mail: tilan.ukwatta@gmail.com (TNU); wozniak@lanl.gov (PRW)
† Director's Postdoctoral Fellow.

The paper is organized as follows. In Section 2, we describe our methodology. In Sections 3 and 4, we present the analysis of the redshift- and duration-dependent GRB clustering. In Section 5, we discuss the caveats and implications of our findings, and Section 6 summarizes the conclusions.

2 DENSITY ESTIMATION

The search for clustering in the angular distribution of GRBs begins with computing all-sky maps for a number of density and proximity measures. Detailed Monte Carlo simulations are then used to establish the statistical significance of the observed clumps and investigate possible systematics.

2.1 *n*-th Nearest-Neighbour Density Estimator

The distance to the n-th nearest neighbour is a widely used proximity measure in astronomy (Ivezić et al. 2014). In the case of sky distributions, this distance is the angular distance between two distinct points on the sphere. One can obtain a two-dimensional density measure by calculating the area enclosed within some radius and then inverting it. Let θ_i be the angular distance to the i-th nearest neighbour. The area a_i enclosed by θ_i is

$$a_i = 2\pi(1 - \cos\theta_i). \tag{1}$$

For $n > 2$, we can write an unbiased n-th nearest-neighbour density estimator ρ_n and its variance σ^2 as

$$\rho_n = \frac{n-1}{a_n}, \qquad \sigma^2(\rho_n) = \frac{\rho_n^2}{n-2}. \tag{2}$$

In the uniform density case, ρ_n is a sufficient statistic, meaning that all neighbours closer than the n-th do not contribute additional information. For details, see the discussion in Woźniak & Kruszewski (2012).

2.2 Gaussian kernel density estimator

The choice of n for the nearest-neighbour density estimator is somewhat arbitrary and reflects a tradeoff between the variance and the spatial scale of the observed density fluctuations. A better approach is to use kernel density estimators, a class of non-parametric estimators with a flexible functional form that can utilize all available data. A kernel density estimator on the sphere is defined as

$$\hat{f}_h(x) = \frac{1}{N\Omega(h)} \sum_{i=1}^{N} K\left(\frac{\theta(x, x_i)}{h}\right). \tag{3}$$

Here x is the location at which the density is predicted based on the set of N data points x_i. The smoothing length h is sometimes referred to as the bandwidth. The kernel function is normalized so that $\hat{f}_h(x)$ represents a probability density over the spherical measure of area $\mathrm{d}\Omega = \sin\theta \mathrm{d}\theta\mathrm{d}\phi$:

$$\Omega(h) = \int K\left(\frac{\theta}{h}\right) \mathrm{d}\Omega(\theta). \tag{4}$$

To obtain the number density of GRBs per unit solid angle for a particular sample, equation (3) must be multiplied by N, the total number of bursts in the sample covering the entire sky. A variety of kernel functions have been explored in the literature (Klemela 2000). For this study we choose the Gaussian kernel function defined as

$$K(y) = \exp\left(-\frac{1}{2}y^2\right). \tag{5}$$

Our density estimator is therefore

$$\hat{f}_h(x) = \frac{1}{N\Omega(h)} \sum_{i=1}^{N} \exp\left(\frac{-\theta_i^2}{2h^2}\right), \tag{6}$$

where

$$\Omega(h) = 2\pi \int_0^{\pi} \exp\left(\frac{-\theta^2}{2h^2}\right) \sin\theta \mathrm{d}\theta. \tag{7}$$

Here $\theta_i = \theta(x, x_i)$ is the spherical distance between an arbitrary line of sight x and the location x_i of the i-th GRB in the sample. Each term in $\hat{f}_h$ is shown to have the same norm $\Omega(h)$ by shifting the pole of the spherical coordinate system (ϕ, θ) to the corresponding data point.

The performance of kernel density estimators does not depend strongly on the choice of the kernel function; however, it is crucial to select a good value of the smoothing length. The bandwidth h can be robustly optimized by minimizing the mean integrated squared error (MISE; e.g. Ivezić et al. 2014)

$$\int (\hat{f}_h - f)^2 \mathrm{d}\Omega$$

$$= \int \hat{f}_h^2 \mathrm{d}\Omega - \frac{2}{N} \sum_{i=1}^{N} \hat{f}_{h,-i}(\theta_i) + \int f^2 \mathrm{d}\Omega. \tag{8}$$

The last term that includes the unknown true density $f(x)$ does not depend on h and is effectively a constant offset. The middle term is proportional to the expectation value of the density estimator. It is approximated here using leave-one-out cross-validation, i.e. by taking a mean of N independent density estimates at the location of each burst i based on $N - 1$ remaining bursts. Finally, the first term is evaluated by the brute force numerical integration because the order of the summation and integration can no longer be switched.

3 REDSHIFT-DEPENDENT CLUSTERING OF GRBS

3.1 GRB sample

Clustering in the sky distribution of GRBs over a limited range of redshift indicates the presence of large-scale structures in the Universe. However, statistically significant overdensities may also occur due to observational biases that are unrelated to any physical groups. In order to construct a uniform, high-quality sample, we limit this part of our study to GRBs detected by *Swift*. Since most bursts with a known redshift are in fact detected by *Swift*, this selection does not drastically reduce the sample size. Our sample includes 311 *Swift* GRBs with redshift measurements from the third *Swift* Burst Alert Telescope Gamma-ray Burst Catalog (Lien 2015). The redshift distribution of our GRB sample is shown in Fig. 1. Both short and long GRBs are included in the sample. All GRBs are believed to be tracers of galaxies and matter, so in principle no further selection cuts are needed to search for clusters. The sample is strongly skewed towards low redshift. There are 50 bursts with $z > 3.5$ and only 25 with $z > 4$.

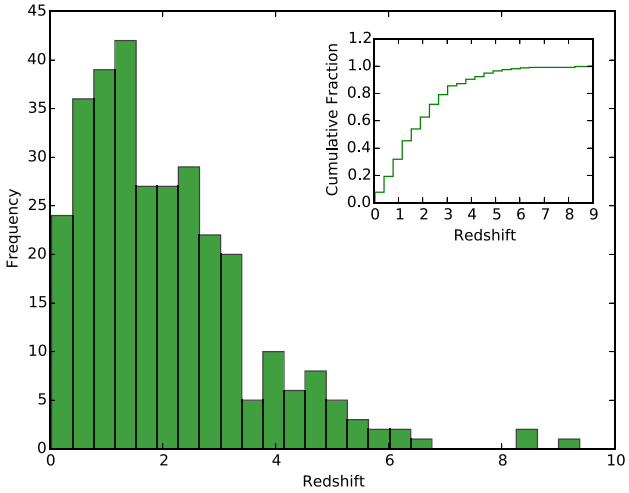

Figure 1. Redshift distribution for a sample of 311 *Swift* bursts.

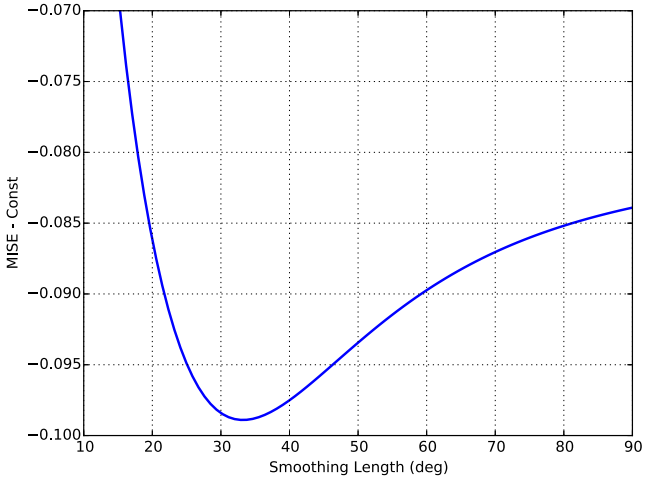

Figure 2. MISE as a function of the smoothing length for a sub-sample of 34 bursts from Fig. 3. The optimal smoothing length is ∼33° found by minimizing MISE in equation (8).

3.2 GRB density map

The first step in our search for GRB clustering is computing all-sky density maps for various samples and redshift ranges using equation (6). Throughout this paper, each map is normalized to represent a probability density function (PDF) that integrates to 1 over the entire sphere (4π solid angle). The number density of bursts per steradian depends on the sample and is easily obtained by multiplying the PDF by the total number of data points in the sample. Other quantities such as the total exposure time per line of sight are also best visualized as PDFs and scaled as needed.

Horváth et al. (2014) examined the angular distribution of GRBs in multiple redshift bins using a two-dimensional KS test and the *n*-th nearest-neighbour distance. Based on a somewhat different data set they found that the apparent anisotropy is strongest for $1.6 < z < 2.1$. In our GRB catalogue, this redshift range contains 34 bursts and also approximately maximizes the observed density contrast. To enable a direct comparison with Horváth et al. (2014), we focus our analysis of redshift-dependent clustering in the same redshift bin. The optimal smoothing length h is calculated by minimizing MISE (equation 8) as shown in Fig. 2 for *Swift* bursts with $1.6 < z < 2.1$. The best value of h depends on both the size of the sample and the angular scale of the most significant clusters. In this particular case $h \sim 33^\circ$. The resulting GRB density map is shown in Fig. 3. Taken at face value, the map may suggest that the distribution is clumpy, a rigorous estimate of the significance of the observed density fluctuations is required to draw any conclusions, especially for small samples.

Variations in the total observing time from one line of sight to another may introduce spurious density fluctuations in GRB samples. Fortunately, this information is available for *Swift* in the relevant time interval. Fig. 4 shows the *Swift* exposure map in galactic coordinates based on 104 months of observing. The partial covering fraction was set to 100 per cent over the entire Burst Alert Telescope (BAT) field of view (Baumgartner, private communication). Similar to the number density plots, the density and the corresponding colour scale were converted to a probability density for easy comparisons. Multiplying by 3.41×10^8 converts back to the total exposure time in seconds. It is intriguing that both the density map (Fig. 3) and the exposure map (Fig. 4) display similar peaks and valleys in roughly the same directions, which points to a possible observational bias. However, Fig. 3 shows a peak-to-valley density contrast around ∼6.5, while the exposure time ratio between the most and the least observed areas in Fig. 4 is only a factor of 2 or so.

Another important effect to consider is the selection bias introduced by redshift measurements. Since it is difficult to measure redshift of a GRB near the galactic plane, we expect to find GRBs with redshift measurements preferentially at high galactic latitudes. Indeed, the effect is clearly visible once all redshift bins are combined for better statistics (Fig. 5). In order to account for both the *Swift* exposure and redshift selection, we need to use a combined *Swift* probability map shown in Fig. 6 obtained by multiplying maps shown in Figs 4 and 5. This combined density map is somewhat similar to the GRB distribution in Fig. 3 with about the same peak-to-valley density contrast ∼ 6.7.

We evaluate the significance of the observed density contrast using a Monte Carlo simulation. The simulator generates 5000 synthetic samples of 34 GRBs with random locations following three different scenarios for the underlying distribution: (1) the *Swift* exposure map from Fig. 4, (2) the combined probability map from Fig. 6 and (3) the uniform PDF over the entire sphere. A density map similar to Fig. 3 and the maximum density value is then computed for each sample. The distribution of these maxima is then used to derive the probability that the maximum density observed in actual data (Fig. 3) may occur due to random fluctuations referred to as the *p*-value. The probability of getting the maximum density value seen in Fig. 3 from random fluctuations following a uniform distribution is 0.013 or 1.3 per cent. The same density peak is easier to generate by chance from the *Swift* exposure map and the corresponding *p*-value is 0.022. It is even easier to generate (*p*-value of 0.025) with the combined map which also incorporates redshift selection effects.

Another density indicator can be obtained from the cumulative distribution of the *n*-th nearest-neighbour distance as was done by Horváth et al. (2014). Fig. 7 shows the cumulative distribution of the distance to the 10th nearest neighbour in our sample of 34 GRBs with redshifts between 1.6 and 2.1 (shown in red). This is compared to the mean cumulative distribution of the 10th nearest-neighbour distance for GRBs distributed according to the *Swift* exposure map

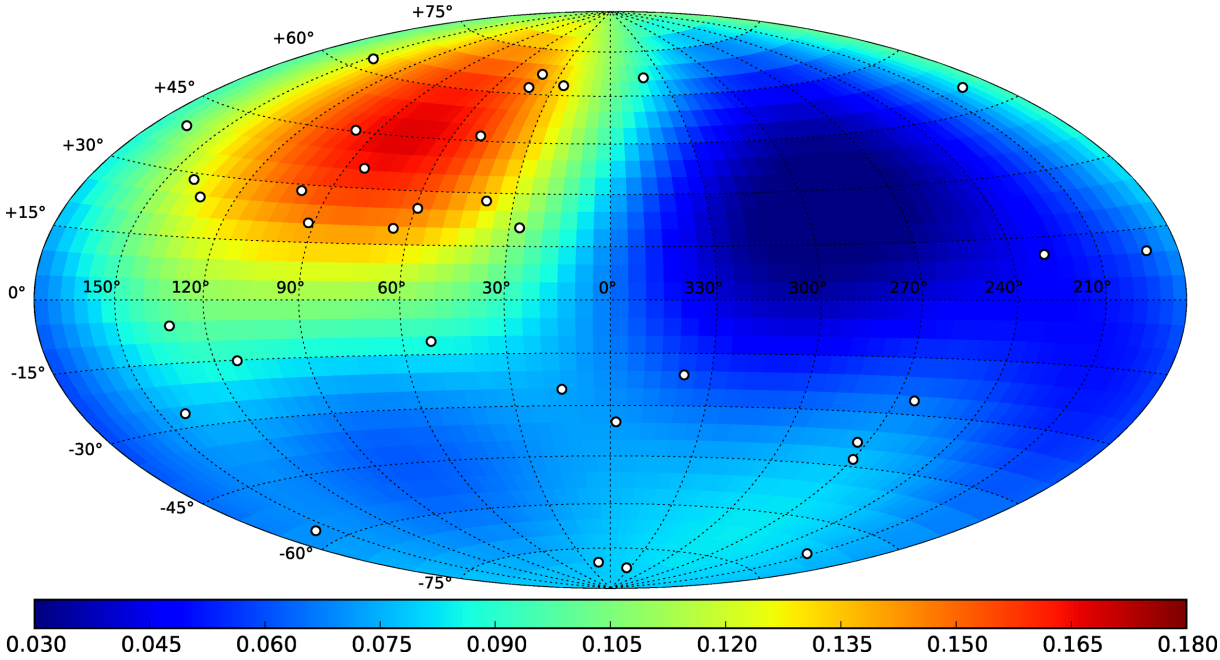

Figure 3. Density map in galactic coordinates for a sub-sample of 34 GRBs in the redshift range $1.6 < z < 2.1$. The colour-coded quantity is the PDF normalized to the full sky. The optimal smoothing length is 33°. Circles indicate the actual burst locations. The maximum and the minimum density values in this map are 0.167 and 0.026, correspondingly. The probability of generating this density contrast by chance estimated using a Monte Carlo simulation is 0.013 assuming that the true sky distribution is uniform, 0.022 assuming that burst detections follow the *Swift* exposure map, and 0.025 considering both *Swift* exposure function and redshift selection bias.

(green line), the *Swift* map conditional on redshift measurement (black line) and uniformly distributed GRBs (blue line). As before, 5000 samples were simulated with 34 GRBs per sample to calculate the reference cumulative distance distributions. The p-value was derived from the distribution of the maximum differences between the distribution of the 10th nearest-neighbour distance in each simulated sample and the reference (mean) distribution for both scenarios under consideration. The resulting p-value is 0.022 for simulated GRBs distributed according to the *Swift* exposure map, 0.127 for the combined *Swift* map and 0.016 for the uniform case. The results for other values of n nearest neighbours are shown in Fig. 8.

Probabilities from both analyses (based on density estimation and proximity) are consistent with each other. However, we note that proximity analysis based on the combined *Swift* prior seems to give a particularly large p-value. Without any additional insight all probabilities obtained so far would indicate that observing the actual sky distribution of the 34 GRBs in our sample is a somewhat unlikely event. This is not sufficient, however, to conclude with good confidence that there is a significant clustering of GRBs in the redshift range $1.6 < z < 2.1$. Moreover, the density maps in other possible redshift bins look very flat and are entirely consistent with random fluctuations. Perhaps the most serious problem is that the quoted p-values do not take into account the (unknown) number of implicit trials that occurred when the redshift range was selected from all possible slices of the original data by Horváth et al. (2014). Since there is a large overlap between our data set and the one used by Horváth et al. (2014), the corrected probabilities could easily be an order of magnitude larger. Therefore, we find no significant evidence of redshift-dependent clustering in the *Swift* GRB sample.

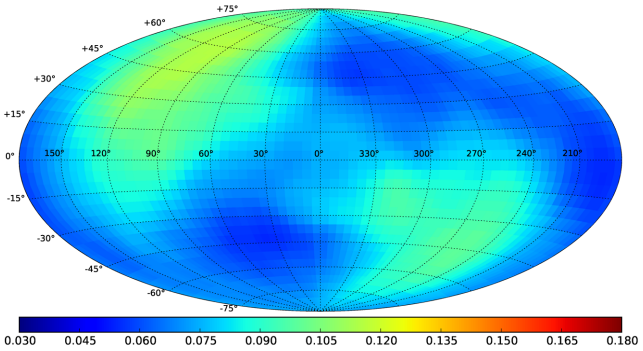

Figure 4. *Swift*/BAT exposure map in galactic coordinates from 104 months of observing. The partial covering fraction is set to 100 per cent over the entire BAT field of view (Baumgartner 2015, private communication). The colour scale indicates the PDF of observing a particular line of sight. To obtain the total exposure time in seconds the PDF should be multiplied by 3.41×10^8.

4 DURATION-DEPENDENT CLUSTERING OF GRBS

4.1 GRB sample

The two dominant populations of GRBs were originally distinguished based on the bimodal distribution of burst durations: long GRBs with T90 $\geq$ 2 s and short GRBs with T90 $<$ 2 s (Kouveliotou et al. 1993). T90 is the time interval that contains 90 per cent of the burst fluence centred on the mid-point (i.e. starting at 5 per cent). Several authors report evidence for additional GRB populations based on the distribution of durations (Balazs et al. 1998; Cline et al.

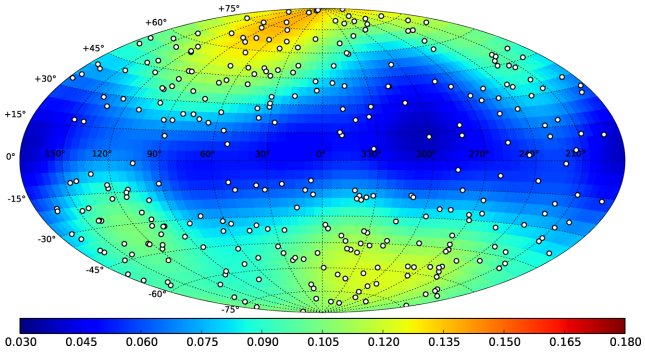

Figure 5. Density map in galactic coordinates for a sample of 311 GRBs with measured redshifts. The optimal smoothing length is 20°. Circles indicate the actual burst locations. The maximum and the minimum density values in this map are 0.139 and 0.038.

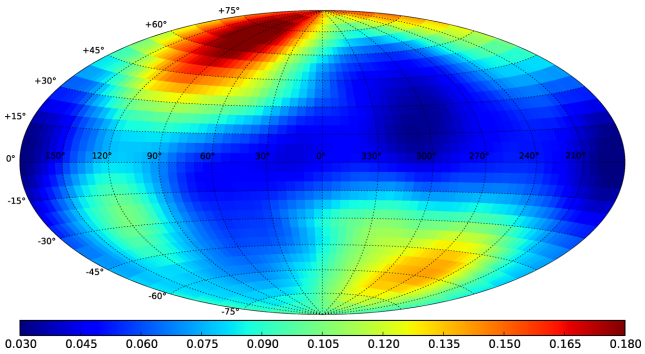

Figure 6. Combined probability density map that includes both the *Swift*/BAT exposure map and the bias due to the required redshift measurement. The maximum and the minimum density values in this map are 0.188 and 0.028.

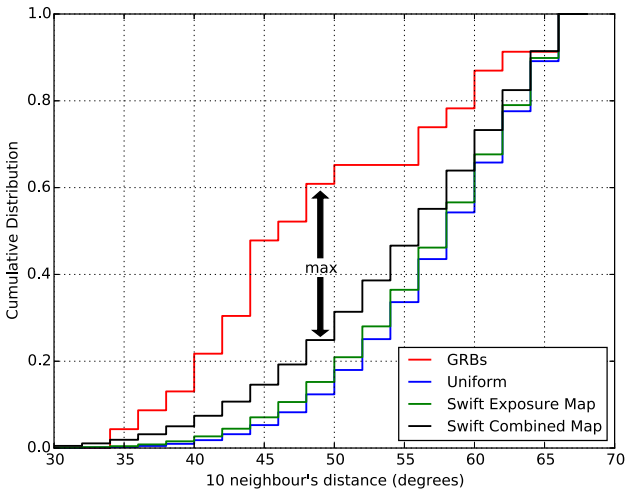

Figure 7. Observed 10th nearest-neighbour distance distribution compared to the mean of 5000 GRB samples generated from the *Swift* exposure map. The *p*-value is 0.016 for the deviation from the uniform distribution, 0.022 for the deviation from the *Swift* exposure function, and 0.127 for the deviation from the combined *Swift* probability map.

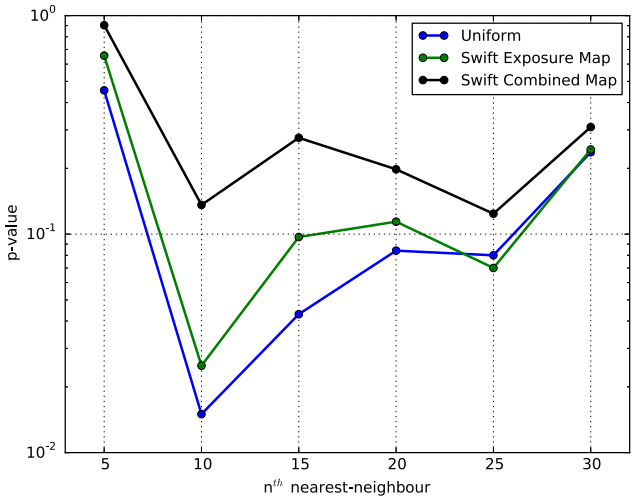

Figure 8. Probability of observing the measured difference between the expected distribution of the *n*-th nearest neighbour and the actual *Swift* data for several values of *n*.

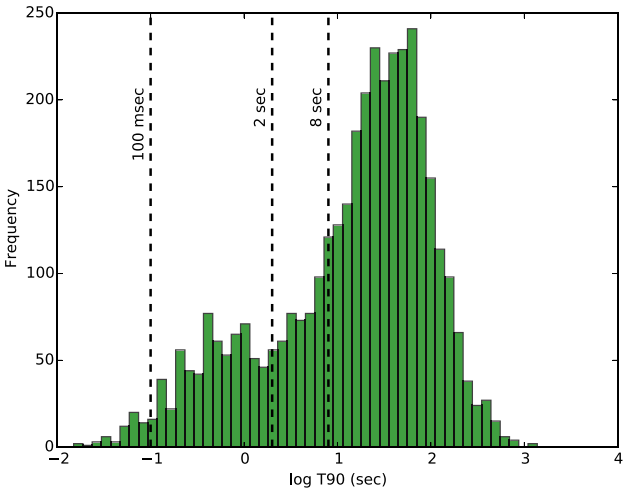

Figure 9. Distribution of durations for our sample of 3798 bursts.

1999, 2005; Mészáros et al. 2000; Litvin et al. 2001; Magliocchetti et al. 2003; Cline 2011). Duration-dependent clustering in the sky distribution of GRBs may help to identify a new GRB population with distinct physical properties. We investigated this possibility using several different samples of GRBs detected by *Swift*, *Fermi* gamma-ray burst monitor (GBM) and *Compton Gamma Ray Observatory* Burst and Transient Source Experiment (BATSE). Our data set includes 2037 BATSE GRBs[1] (Paciesas et al. 1999), 997 GBM GRBs (Gruber et al. 2014; von Kienlin et al. 2014) and 889 *Swift* GRBs (Lien 2015). Within this combined catalogue, 125 bursts were detected by both *Swift* and GBM, and for those we used the *Swift* observations because they provide much better localizations. Our final combined sample includes a grand total of 3798 GRBs. The corresponding T90 distribution is shown in Fig. 9. We used the same approach as in Section 3 to investigate the duration-dependent clustering in the GRB sky distribution.

[1] http://gammaray.msfc.nasa.gov/batse/grb/catalog/current/

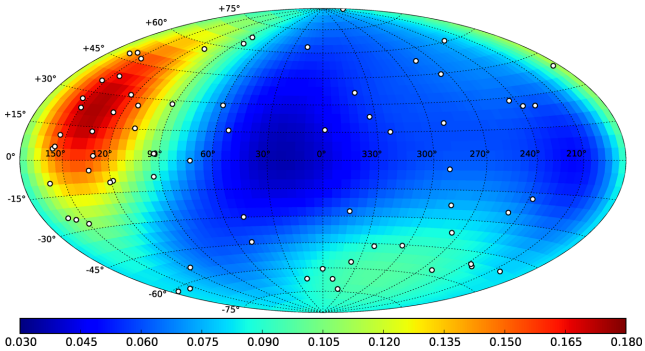

Figure 10. Density map in galactic coordinates for a sub-sample of 68 VSGRBs with T90 $\leq$ 100 ms. The optimal smoothing length is 28°. Circles indicate the actual burst locations. The maximum and the minimum density values in this map are 0.174 and 0.036, correspondingly. The probability of generating this density contrast by chance estimated using a Monte Carlo simulation is 0.0008 assuming that burst detections follow the multi-instrument exposure map.

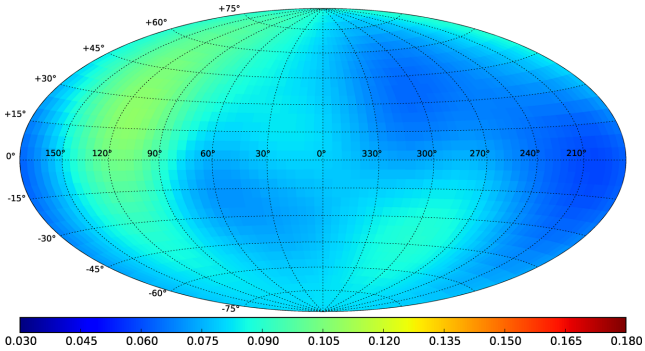

Figure 11. *Swift* exposure map derived assuming that long GRBs are uniformly distributed. The map was computed using a sample of 808 long bursts and the optimal smoothing length of 24°.

4.2 Very short duration GRB density map

There are 68 GRBs with T90 $\leq$ 100 ms in our sample. Fig. 10 shows the corresponding density map computed using the optimal smoothing length of 28°. A concentration of GRBs can be clearly seen towards the left-hand side of the map. Using a different compilation of burst localizations, Cline et al. found evidence that VSGRBs tend to cluster in an area of the sky that roughly coincides with the overdensity in Fig. 10 (Cline et al. 1999, 2005; Cline 2011). How significant is this density contrast?

Unlike for *Swift*, exposure maps are not available for *Fermi* and BATSE. However, if we had a good sample of GRBs from a population for which the true sky distribution is uniform, we could use the density map for that sample as a proxy exposure map. We verified the validity of this approach using long GRBs detected by *Swift*. The location and amplitude of the main features in the resulting density map (Fig. 11) are similar to the actual *Swift* exposure map from Fig. 4. We therefore adopt this method to derive exposure corrections for all GRB samples except those limited to bursts detected by *Swift*. The correction is actually less important for multi-instrument samples that tend to average out any local variations specific to a single instrument. Although the *p*-values change slightly when a flat exposure map is assumed, none of our conclusions depend on the correction.

As before, we use Monte Carlo simulations to estimate the significance of the anisotropy in the distribution of VSGRBs (Fig. 10).

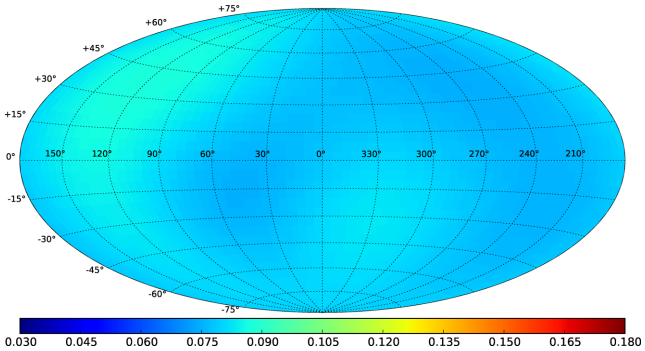

Figure 12. Multi-instrument exposure map derived assuming that long GRBs are uniformly distributed. The map was computed using a sample of 3063 long bursts detected by BATSE, *Fermi*/GBM, and *Swift*/BAT and the optimal smoothing length of 31°.

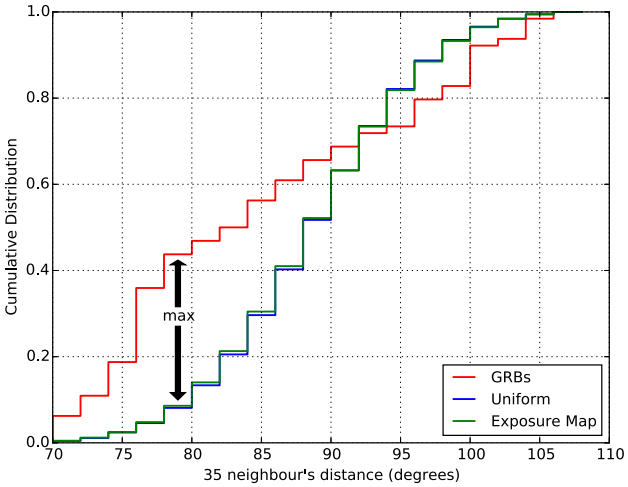

Figure 13. Cumulative distribution of the 35th nearest-neighbour distance compared to the mean of 5000 simulated samples of 68 GRBs (cf. Fig. 10). The probability of the observed deviation from the uniform distribution and the multi-instrument exposure map is, respectively, 0.0078 and 0.0084.

The effective exposure map for all three instruments computed assuming a uniform distribution of long GRBs is shown in Fig. 12. The map is based on 3063 long GRBs detected by BATSE, *Fermi*/GBM, and *Swift*/BAT. Note that the amplitude variations of this map is only ∼25 per cent, much less than the factor of two variations seen in the *Swift* exposure map. We draw 5000 samples of 68 GRBs with the probability density proportional to the exposure map, compute density maps, and compare the resulting distribution of the peak density with the observed value. The probability of generating the observed density peak (Fig. 10) from a random fluctuation is 0.0008.

Similar to Section 3, we also examined the cumulative *n*-th nearest-neighbour distance distributions for our sample of 68 VSGRBs. Fig. 13 shows the cumulative distance distribution for the 35th nearest neighbour (red) and the mean of 5000 samples drawn from a uniform distribution (blue) and from the multi-instrument exposure map (green). As expected, the two sets of simulated results are very similar because the multi-instrument exposure map is very flat. The corresponding *p*-values are 0.0078 for the uniform distribution and 0.0084 for the multi-instrument exposure map. Fig. 14 shows how the *p*-values change as a function of the *n*-th nearest neighbour. The lowest *p*-value (highest significance) is reached for $n = 35$ shown in Fig. 13. Again, there is little

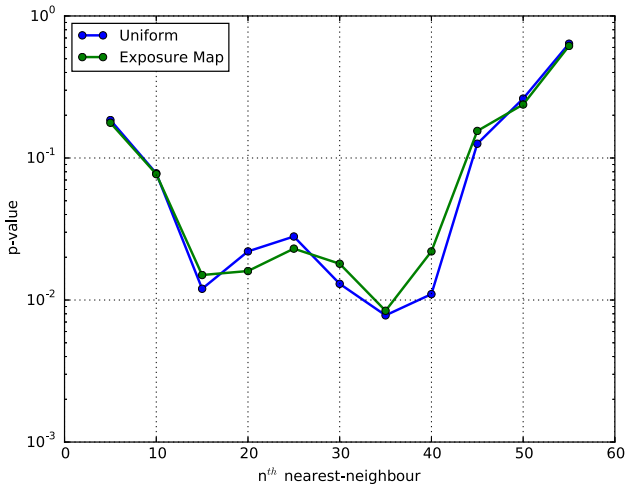

Figure 14. Probability of the maximum deviation between the observed distribution of the n-th nearest-neighbour distance and the mean of simulated samples of 68 VSGRBs. The significance peaks for $n = 35$ for both considered types of exposure maps.

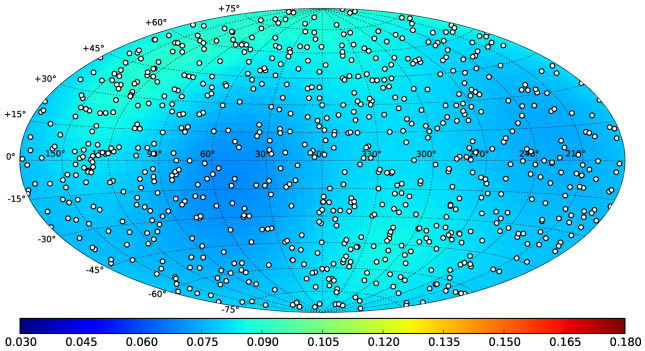

Figure 15. Density map in galactic coordinates for a sub-sample of 735 short GRBs with T90 < 2 s. The optimal smoothing length is 35°. Circles indicate the actual burst locations. The maximum and the minimum density values in this map are 0.09 and 0.07, correspondingly. The probability of generating this density contrast by chance estimated using a Monte Carlo simulation is 0.3 assuming that burst detections follow the multi-instrument exposure map.

difference between the multi-instrument exposure map and the uniform distribution.

While the significance of the observed anisotropy in the distribution of VSGRBs is not overwhelming, the probability of a chance alignment is less than 1 per cent. Unfortunately, these probabilities could become much higher when corrected for the unknown number of multiple trials incurred to select the range of GRB durations for this clustering analysis.

4.3 Short- and intermediate-duration GRB density map

The short-duration bursts – thought to originate from compact binary mergers – are expected to trace galaxies at cosmological distances. This suggests that there should not be any significant anisotropies in their sky distribution. Meanwhile Balazs et al. (1998) and Magliocchetti et al. (2003) have reported evidence that supports the existence of such anisotropies. We searched for clustering in our sample of 735 short GRBs defined here as having T90 ≤ 2 s. The resulting density map shown in Fig. 15 is very flat and allows little or no clustering. This visual impression is confirmed by a high probability (p-value $\sim$0.3) of the observed density peak from Monte Carlo simulations based on the multi-instrument exposure map.

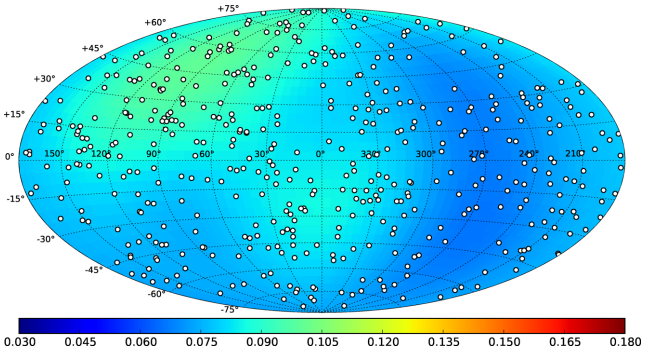

Figure 16. Density map in galactic coordinates for a sub-sample of 468 short GRBs with 2 s $<$ T90 $<$ 8 s. The optimal smoothing length is 32°. Circles indicate the actual burst locations. The maximum and the minimum density values in this map are 0.1 and 0.07, correspondingly. The probability of generating this density contrast by chance estimated using a Monte Carlo simulation is 0.05 assuming that burst detections follow the multi-instrument exposure map.

While the evidence that the intermediate-duration GRBs defined as having 2 s $<$ T90 $<$ 8 s constitute a separate physical class is much weaker, there is some evidence that they exhibit a noticeable level of anisotropy in their sky distribution (Mészáros et al. 2000; Litvin et al. 2001). Our data set includes 468 intermediate-duration GRBs. Fig. 16 shows the density map for this sample with the optimal smoothing radius of 32°. We find no indication of significant clustering in the distribution of intermediate-duration GRBs with the formal p-value of $\sim$0.05.

5 DISCUSSION

A recent analysis of GRB clustering across the redshift space by Horváth et al. (2014) revealed a potential anisotropy in the redshift range $1.6 < z < 2.1$. Using a sample of 31 GRBs detected by multiple instruments, these authors argued that there is evidence for a large-scale structure in this particular redshift bin. Our study based on 34 *Swift* GRBs in the same redshift range does not confirm the significance of features reported by Horváth et al. (2014). Although our estimated density maps show similar looking peaks and valleys (Fig. 3), a rigorous Monte Carlo simulation demonstrates that, given the small sample size, the observed density variations are consistent with random fluctuations at a few per cent probability level. This conclusion is reinforced when we consider the effect of multiple trials necessary to identify one interesting redshift bin from the entire GRB catalogue. For example, 10 trials would roughly increase the probability by an order of magnitude. Given the available data, the observed clustering is not significant.

The probability of observing a purely random density fluctuation of a certain amplitude clearly depends on the size of the sample. It is interesting to investigate how many GRBs are needed in a redshift bin to exclude a chance alignment with a reasonably good confidence (e.g. pre-trial p-value $\sim$0.001). We generated samples of varying size using the *Swift* exposure map and estimated the probability of observing a global PDF maximum of 0.167 or higher corresponding to our actual measurement from Fig. 3. The results are shown in Fig. 17. More than 60 GRBs per redshift bin are required to reach the p-value of 0.001 in this case. With the current rate of GRB discoveries and redshift measurements, it will likely

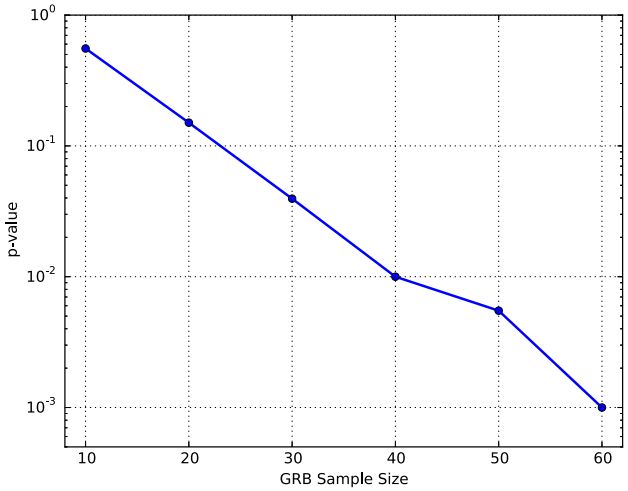

Figure 17. Probability of observing a global density peak from Fig. 3 as a function of the sample size. The simulation is based on the *Swift* exposure map with 33° smoothing radius.

take longer than 10 yr to accumulate a sufficiently large sample and make a definitive statement about the presence of clustering, assuming that the currently observed density contrast persists.

The most intriguing result from our duration-dependent clustering analysis is the relatively strong anisotropy in the distribution of VSGRBs with T90 $\leq$ 100 ms. The pre-trial p-value of the density peak seen in Fig. 10 is 0.0008. Somewhat lower confidence levels were obtained from the analysis of the cumulative n-th nearest-neighbour distance distribution. It is not clear in this case how many trials were incurred when this particular range of durations was selected. In our analysis, we examined four duration bins: T90 $\leq$ 100 ms, T90 $\leq$ 2 s, 2 s $<$ T90 $<$ 8 s and T90 $>$ 2 s. However, in order to properly account for trials, one needs to also consider the ones incurred by previous authors in their analysis. Ultimately, this unknown number of trials will degrade the significance of the clustering seen in the distribution of VSGRBs. If confirmed by future GRB detections, this clustering may potentially help to identify a new population of GRBs.

Cline (2011) proposes that VSGRBs may originate from evaporating primordial black holes (PBH) in the solar neighbourhood. The main problem with this scenario, assuming that there is in fact a significant clustering, is that the PBH sources would have to reside preferentially in a relatively confined region of the solar neighbourhood consistent with the Galactic anticentre. Another problem is that the typical time profile of these bursts exhibits relatively complex structure with multiple peaks (Czerny et al. 2011). This characteristic is inconsistent with the final stages of the black hole evaporation, when the predicted emission from the black hole is smooth as it only depends on one parameter, the mass (Carr et al. 2010; Ukwatta et al. 2015). All PBH bursts are therefore expected to look similar and have a single peak. The observed variety of VSGRB light curves argues against the PBH burst origin for most of the population.

The fact that some VSGRBs may be at cosmological distances is also incompatible with the PBH hypothesis. The initial mass of PBHs expiring today is $\sim 5.0 \times 10^{11}$ kg (Carr et al. 2010; Ukwatta et al. 2015). During the final stages of the burst, only about $\sim 10^5$ kg is left in a typical PBH. Even assuming that all mass is converted into photons in the keV/MeV energy range, the maximum possible distance for the PBH to remain detectable is less than few parsecs. In our sample, there are 13 *Swift* GRBs in the VSGRB category. Three of those coincide with host galaxies at known redshifts: GRB 050509B, GRB 060502B, and GRB 100206A. While the host galaxy redshift measurement alone does not rule out the PBH origin of some VSGRBs because of the possibility of chance associations, the presence of the long-lived lower energy emission does. One of the three bursts in question (GRB 050509B) displayed an X-ray afterglow(Castro-Tirado et al. 2005) that excludes the PBH origin for this particular burst. However, it is still possible that some VSGRBs are in fact due to PBH bursts.

Other possible progenitors for VSGRBs are mergers of binary systems. This scenario is an extension of the standard model for short GRBs. It is possible to tweak merger models to generate time-scales observed in VSGRBs (Czerny et al. 2011). However, in this case it is still difficult to account for the observed potential clustering of VSGRBs.

6 CONCLUSIONS

We analysed the redshift- and duration-dependent clustering in the sky distribution of GRBs using the Gaussian kernel density estimator and the n-th nearest-neighbour distance. Contrary to previous reports (Horváth et al. 2014), our redshift-dependent analysis did not provide evidence of significant clustering. This is especially the case after considering the number of trials incurred to find a hypothetical overdensity that requires repeated redshift binning and multiple visual searches of density maps. The duration-dependent analysis demonstrated that neither short-duration nor intermediate-duration GRBs display any significant clustering. However, very short duration GRBs appear to show some degree of clustering. This has been previously noted by Cline et al. (1999, 2005) as well as Cline (2011) and warrants further attention as it may ultimately provide evidence for the existence of a separate source population.

The apparent angular distribution of long GRBs detected by *Swift* is proportional (within the statistical error) to the total exposure time for a given line of sight. In other words, the true sky distribution of long GRBs is very uniform and can be used to construct approximate exposure maps for various instruments. This technique will be particularly useful for multi-instrument exposure maps and in situations where it is not straightforward to calculate the exposure map directly. A good example of the latter is the interplanetary network.

ACKNOWLEDGEMENTS

This work was funded by the US Department of Energy. TNU acknowledges support from the Laboratory Directed Research and Development programme at the Los Alamos National Laboratory. We thank Brenda Dingus, Pat Harding, Krista Smith and Kevin Hurley for useful conversations on the analysis. We also thank the referee Jean-Luc Atteia for comments that significantly improved the paper.

REFERENCES

Balazs L. G., Meszaros A., Horvath I., 1998, A&A, 339, 1
Briggs M. S. et al., 1996, ApJ, 459, 40
Carr B. J., Kohri K., Sendouda Y., Yokoyama J., 2010, Phys. Rev. D, 81, 104019
Castro-Tirado A. J. et al., 2005, A&A, 439, L15
Cline D. B., 2011, Int. J. Astron. Astrophys., 1, 164
Cline D. B., Matthey C., Otwinowski S., 1999, ApJ, 527, 827

Cline D. B., Czerny B., Matthey C., Janiuk A., Otwinowski S., 2005, ApJ, 633, L73
Czerny B., Janiuk A., Cline D. B., Otwinowski S., 2011, New Astron., 16, 33
Foley S., McGlynn S., Hanlon L., McBreen S., McBreen B., 2008, A&A, 484, 143
Gehrels N. et al., 2004, ApJ, 611, 1005
Gruber D. et al., 2014, ApJS, 211, 12
Horváth I., Hakkila J., Bagoly Z., 2014, A&A, 561, L12
Ivezić Ž., Connelly A. J., VanderPlas J. T., Gray A., 2014, Statistics, Data Mining, and Machine Learning in Astronomy. Princeton Univ. Press, Princeton, NJ
Klemela J., 2000, J. Multivariate Anal., 73, 18
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
Lamb D. Q., Reichart D. E., 2000, ApJ, 536, 1
Lien A., 2015, Proc. Sci., Swift: 10 Years of Discovery. La Sapienza University, Rome, Italy
Litvin V. F., Matveev S. A., Mamedov S. V., Orlov V. V., 2001, Astron. Lett., 27, 416
Magliocchetti M., Ghirlanda G., Celotti A., 2003, MNRAS, 343, 255
Meegan C. A., Fishman G. J., Wilson R. B., Horack J. M., Brock M. N., Paciesas W. S., Pendleton G. N., Kouveliotou C., 1992, Nature, 355, 143
Mészáros A., Bagoly Z., Horváth I., Balázs L. G., Vavrek R., 2000, ApJ, 539, 98
Norris J. P., 2002, ApJ, 579, 386
Paciesas W. S. et al., 1999, ApJS, 122, 465
Tegmark M., Hartmann D. H., Briggs M. S., Meegan C. A., 1996, ApJ, 468, 214
Ukwatta T. N. et al., 2010, ApJ, 711, 1073
Ukwatta T. N. et al., 2012, MNRAS, 419, 614
Ukwatta T. N., Stump D., MacGibbon J. H., Linnemann J. T., Marinelli S. S., Yapici T., Tollefson K., 2015, preprint (arXiv:1507.01648)
von Kienlin A. et al., 2014, ApJS, 211, 13
Woźniak P. R., Kruszewski A., 2012, Acta Astron., 62, 409

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