Mon. Not. R. Astron. Soc. **391,** 1741–1748 (2008) doi:10.1111/j.1365-2966.2008.13635.x

Testing the randomness in the sky-distribution of gamma-ray bursts

R. Vavrek,[1⋆] L. G. Balázs,[2] A. Mészáros,[3] I. Horváth[4] and Z. Bagoly[5]

[1]*ESA/ESAC PO Box 50727 Villafranca del Castillo, 28080 Madrid, Spain*
[2]*Konkoly Observatory, PO Box 67, H-1525 Budapest, Hungary*
[3]*Astronomical Institute of the Charles University, V Holešovičkách 2, 180 00 Prague 8, Czech Republic*
[4]*Department of Physics, Bolyai Military University, PO Box 15, H-1581 Budapest, Hungary*
[5]*Laboratory for Information Technology, Eötvös University, Pázmány Péter sétány 1/A, H-1518 Budapest, Hungary*

Accepted 2008 June 24. Received 2008 June 21; in original form 2006 December 29

ABSTRACT

We have studied the complete randomness of the angular distribution of gamma-ray bursts (GRBs) detected by the Burst and Transient Source Experiment (BATSE). Because GRBs seem to be a mixture of objects of different physical nature, we divided the BATSE sample into five subsamples (short1, short2, intermediate, long1, long2) based on their durations and peak fluxes, and we studied the angular distributions separately. We used three methods, Voronoi tesselation, minimal spanning tree and multifractal spectra, to search for non-randomness in the subsamples. To investigate the eventual non-randomness in the subsamples, we defined 13 test variables (nine from the Voronoi tesselation, three from the minimal spanning tree and one from the multifractal spectrum). Assuming that the point patterns obtained from the BATSE subsamples are fully random, we made Monte Carlo simulations taking into account the BATSE's sky-exposure function. The Monte Carlo simulations enabled us to test the null hypothesis (i.e. that the angular distributions are fully random). We tested the randomness using a binomial test and by introducing squared Euclidean distances in the parameter space of the test variables. We concluded that the short1 and short2 groups deviate significantly (99.90 and 99.98 per cent, respectively) from the full randomness in the distribution of the squared Euclidean distances; however, this is not the case for the long samples. For the intermediate group, the squared Euclidean distances also give a significant deviation (98.51 per cent).

Key words: methods: statistical – cosmology: observations – gamma-rays: bursts.

1 INTRODUCTION

Currently, there is no doubt about the cosmological origin of gamma-ray bursts (GRBs; Zhang & Mészáros 2004; Fox et al. 2005; Mészáros 2006). Thus, assuming a large-scale isotropy for the Universe, the same property is also expected for GRBs. Another property that is also expected to occur is that GRBs should appear fully randomly (i.e. if a burst is observed, it does not give any information about the place of the next burst). If both properties are fulfilled, then the distribution is called completely random (for the astronomical context of spatial point processes, see Pásztor & Tóth 1995). There are several tests for checking the complete randomness of point patterns; however, these procedures do not always give information for both properties simultaneously.

There is increasing evidence that all GRBs do not represent a physically homogeneous group (Kouveliotou et al. 1993; Horváth 1998, 2002; Mukherjee et al. 1998; Hakkila et al. 2000, 2003; Balázs et al. 2003; Horváth et al. 2006). Hence, it is worth investigating whether the physically different subgroups are also different in their angular distributions. In the last few years, Balázs, Mészáros & Horváth (1998), Balázs et al. (1999), Mészáros, Bagoly & Vavrek (2000a) and Mészáros et al. (2000b) have provided several different tests for probing the intrinsic isotropy in the angular sky-distribution of GRBs collected in the Burst And Transient Source Experiment (BATSE) catalogue (Meegan et al. 2000). Briefly summarizing the results of these studies, we can conclude the following. (i) The long subgroup ($T_{90} > 10$ s) seems to be distributed isotropically. (ii) The intermediate subgroup ($2 \leqq T_{90} \leqq 10$ s) is distributed anisotropically on the $\simeq$(96–97) per cent significance level. (iii) For the short subgroup ($2\text{ s} > T_{90}$) the assumption of isotropy is rejected only on the 92 per cent significance level. (iv) The long and the short subclasses, respectively, are distributed differently on the 99.3 per cent significance level. (For details about the definition of subclasses, see Horváth 1998. T_{90} is the duration of a GRB, during which time 90 per cent of the radiated energy is received; Meegan et al. 2000.)

Independently and using different tests, Litvin et al. (2001) confirmed the results (i)–(iii) with one essential difference: for the intermediate subclass, there is a much higher (99.89 per cent)

⋆E-mail: Roland.Vavrek@sciops.esa.int

significance level of anisotropy. Again, the short subgroup is found to be 'suspicious', but only the $\simeq$(85–95) per cent significance level is reached. The long subclass seems to be distributed isotropically (but see Mészáros & Štoček 2003). Magliocchetti, Ghirlanda & Celotti (2003) found significant angular correlation on the 2°–5° scale for GRBs with $T_{90} < 2$ s durations. Tanvir et al. (2005) reported a correlation between the locations of previously observed short bursts and the positions of galaxies in the local Universe, indicating that between 10 and 25 per cent of short GRBs originate at low redshifts ($z < 0.025$).

It is a reasonable requirement to continue these tests using more sophisticated procedures in order to see whether the angular distribution of GRBs is completely random or has some sort of regularity. This is the subject of this paper. New tests are presented here. In particular, clarification of the behaviour of the short subgroup is expected from these tests. In this paper, similar to previous studies, the intrinsic randomness is tested; this means that the non-uniform sky-exposure function of the BATSE instrument is eliminated.

The paper is organized as follows. In Section 2, we describe three new tests; there are no new results presented but, because the methods are not widely familiar, this minimal survey may be useful. In Section 3, we give the statistical tests on the data. In Section 4, we summarize the results of the statistical tests, and in Section 5 we present the main conclusions of the paper.

2 MATHEMATICAL SUMMARY

2.1 Voronoi tesselation

The Voronoi diagram (Fig. 1), also known as Dirichlet tesselation or Thiessen polygons, is a fundamental structure in computational geometry and arises naturally in many different applications (Voronoi 1908; Stoyan & Stoyan 1994). Generally, this diagram provides a partition of a point pattern ('point field', also 'point process') according to its spatial structure, which can be used for analysing the underlying point process.

Assume that there are N points ($N \gg 1$) scattered on a sphere's surface with a unit radius. It is stated that a point field is given on the sphere. The Voronoi cell (Stoyan & Stoyan 1994) of a point is the region of the sphere's surface consisting of points that are closer to this given point than to any others of the sphere. This cell forms a polygon on this sphere. Every such cell has its area (A) given in steradians, a perimeter (P) given by the length of the boundary (one large circle of the boundary curve is also called the 'chord'), a number of vertices (N_v) given by an integer positive number, and by the inner angles ($\alpha_i; i = 1, \ldots, N_v$). This method is completely non-parametric, and therefore may be sensitive for various point pattern structures in the different subclasses of GRBs.

Note that the behaviour of this tesselation method on the sphere's surface is different from that on the infinite plane. This follows from the fact that the areas of the polygons will not be independent from each other, because the total surface of the sphere is fixed in 4π steradian. Hence, the spherical Voronoi tesselation (VT) is not effected by border effects, and the Voronoi diagram becomes a closed set of convex polygons.

The points on the sphere may be distributed completely randomly or non-randomly; the non-random distribution may have different characteristics (clustering, filaments, etc.; for a survey of these non-random behaviours, see, for example, Diggle 1983).

Every random and some regular patterns have a distribution with one characteristic maximum (unimodal) but with different variances. Multimodality means different characteristic maxima indicating a hierarchical (cluster) structure, with the number of modes determined by the number of scales in the sample. The VT method is able both to detect the non-randomness and to describe its form. For more details, see Stoyan & Stoyan (1994), and for the astronomical context, see Coles & Barrow (1990), Coles (1991), Icke & van de Weygaert (1991), Ikeuchi & Turner (1991), Subba Rao & Szalay (1992), van de Weygaert (1994), Zaninetti (1995), Doroshkevich,

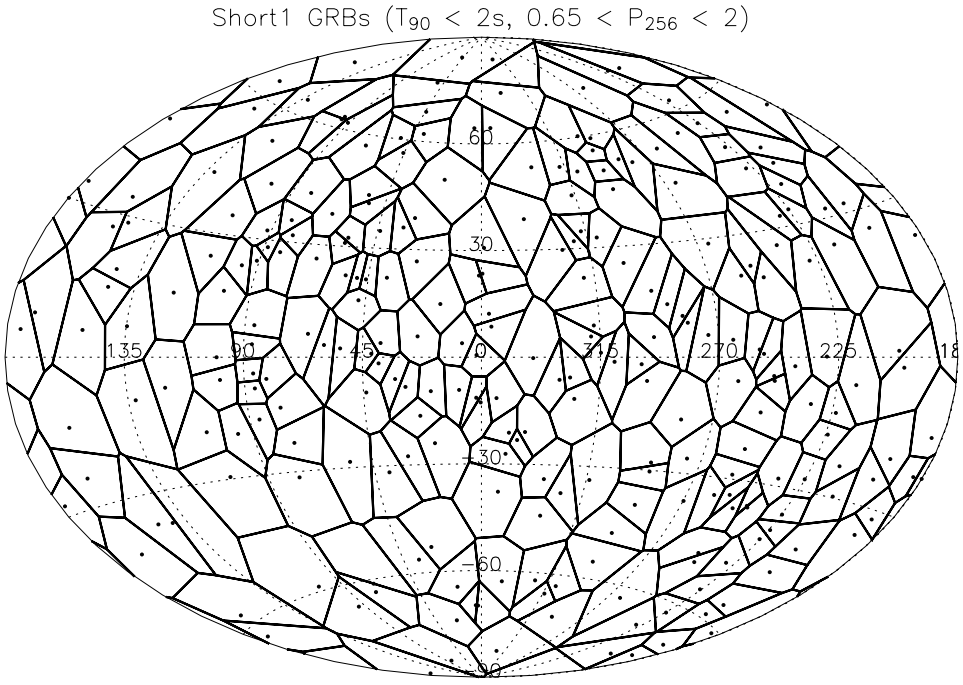

Figure 1. VT of the short GRBs (short1 sample) in the $0.65 < P_{256} < 2.00$ peak-flux range in Galactic coordinates. The peak flux is given in dimension photons $(\mathrm{cm}^2\ \mathrm{s})^{-1}$.

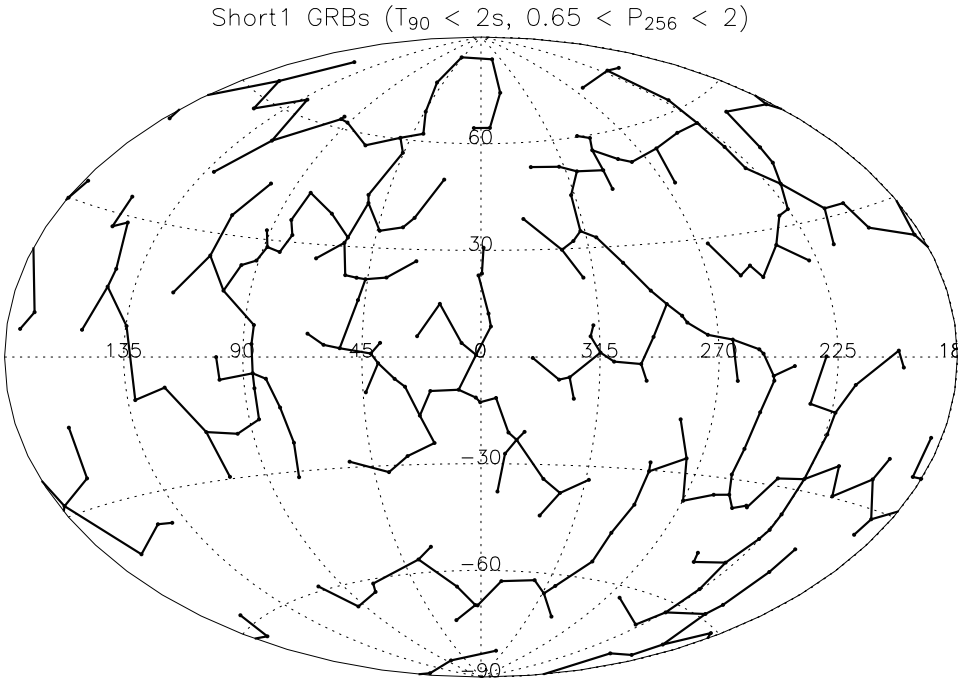

Figure 2. The MST for the sample in Fig. 1.

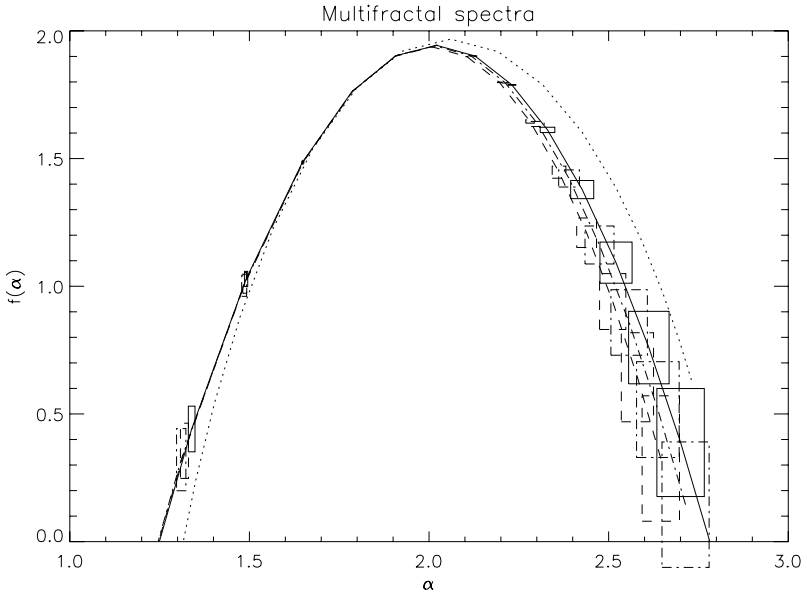

Figure 3. MFR spectra of simulated (dot-dashed), long1 (dashed), short1 (dotted) and short2 (three-dot-dashed) samples. Boxes represent the error of spectrum points derived from Monte Carlo simulations. Note the shift of the maximum of the spectrum of the short1 sample towards higher values in comparison to $\alpha = 2$, corresponding to the completely random two-dimensional Euclidean case.

Gottlöber & Madsen (1997), Yahagi, Mori & Yoshii (1999) and Ramella et al. (2001).

2.2 Minimal spanning tree

Contrary to VT, this method considers the distances (edges) among the points (vertices) (Fig. 2). Clearly, there are $N(N-1)/2$ distances among N points. A spanning tree is a system of lines connecting all the points without any loops. The minimal spanning tree (MST) is a system of connecting lines, where the sum of the lengths is minimal among all the possible connections between the points (Kruskal 1956; Prim 1957). In this paper, the spherical version of the MSF is used, following the original paper by Prim.

The $N-1$ separate connecting lines (edges) together define the MST. The statistics of the lengths and the $\alpha_{\rm MST}$ angles between the edges at the vertices can be used for testing the randomness of the point pattern. The MST is widely used in cosmology for studying the statistical properties of galaxy samples (Barrow, Bhavsar & Sonoda 1985; Bhavsar & Lauer 1996; Krzewina & Saslaw 1996; Bhavsar & Splinter 1996; Adami & Mazure 1999; Doroshkevich & Turchaninov 2001).

2.3 Multifractal spectrum

Let $P(\varepsilon)$ denote the probability for finding a point in an area of ε radius. If $P(\varepsilon)$ scales as ε^{α} [i.e. $P(\varepsilon) \propto \varepsilon^{\alpha}$], then α is called the local fractal dimension (e.g. $\alpha = 2$ for a completely random process on the plane). In the case of a monofractal, α is independent of the position. A multifractal (MFR) on a point process can be defined as unification of the subsets of different (fractal) dimensions (Paladin & Vulpiani 1987). $f(\alpha)$ usually denotes the fractal dimension of the subset of points at which the local fractal dimension is in the interval of $\alpha, \alpha + \mathrm{d}\alpha$. The contribution of these subsets to the whole pattern is not necessarily equally weighted; practically it depends on the relative abundances of subsets. The $f(\alpha)$ functional relationship between the fractal dimension of subsets and the corresponding local fractal dimension is called the MFR or Hausdorff spectrum (Fig. 3).

In the vicinity of the ith point ($i = 1, 2, \ldots, N$) we can measure from the neighbourhood structure a local dimension α_i ('Rényi dimension'). This measure approximates the dimension of the embedding subset, making it possible to construct the MFR spectrum that characterizes the whole pattern (for more details, see Paladin & Vulpiani 1987). If the maximum of this convex spectrum is equal to the Euclidean dimension of the space, then in the classical sense the pattern is not a fractal, but the spectrum remains sensitive to the non-randomness of the point set.

There is a wide variety of astronomical phenomena, where the concept of fractal and/or MFR can be successfully applied (Giraud 2000; Irwin, Widrow & English 2000; Kawaguchi et al. 2000; Pan & Coles 2000, 2002; Selman & Melnick 2000; Bottorff & Ferland 2001; Célérier & Thieberger 2001; Chappell & Scalo 2001; Tatekawa & Maeda 2001; Vavrek, Balázs & Epchtein 2001; Aschwanden & Parnell 2002; Casuso & Beckman 2002; Elmegreen 2002; Gaite & Manrubia 2002; Semelin & Combes 2002; Tikhonov 2002; Datta 2003).

3 STATISTICAL TESTS ON THE DATA

The three procedures outlined in Section 2 enable us to derive several stochastic quantities well suited for testing the non-randomness of the underlying point patterns.

3.1 Input data and the definition of samples

Up to the present, the most comprehensive all-sky survey of GRBs was carried out by the BATSE experiment on board the *Compton Gamma-ray Observatory* (*CGRO*) satellite in the period 1991–2000. In this period, the experiment collected 2704 well-justified burst events and the data are available in the current BATSE catalogue (Meegan et al. 2000).

Because there is increasing evidence (Horváth et al. 2006, and references therein) that the GRB population is actually a mixture of astrophysically different phenomena, we have divided the GRBs into three groups: short ($T_{90} < 2$ s), intermediate ($2 \leq T_{90} \leq 10$ s) and long ($T_{90} > 10$ s). To avoid problems with the changing detection threshold, we omitted GRBs having a peak flux $P_{256} \leq 0.65$ photons cm^{-2} s^{-1}. This truncation was proposed by Pendleton et al. (1997). The bursts may emerge at very different distances in the line of sight and it may happen that the stochastic structure of the angular distribution depends on it. Therefore, we also carried out tests on the bursts with $P_{256} < 2$ photons cm^{-2} s^{-1} in the short and long populations, separately. Table 1 defines the five samples studied here.

Table 1. Tested samples of BATSE GRBs.

Sample	Duration (s)	Peak flux (photons cm^{-2} s^{-1})	Number of GRBs
Short1	$T_{90} < 2$	$0.65 < P_{256} < 2$	261
Short2	$T_{90} < 2$	$0.65 < P_{256}$	406
Intermediate	$2 \leqq T_{90} \leqq 10$	$0.65 < P_{256}$	253
Long1	$T_{90} > 2$	$0.65 < P_{256} < 2$	676
Long2	$T_{90} > 10$	$0.65 < P_{256}$	966

3.2 Definition of the test variables

The randomness of the point field on the sphere can be tested with respect to different criteria. As different non-random behaviours are sensitive for different types of criteria of non-randomness, it is not necessary that all possible tests using different measures reject the assumption of randomness. In the following, we defined several test variables that are sensitive to different stochastic properties of the underlying point pattern, as proposed by Wallet & Dussert (1998).

3.2.1 Voronoi tesselation

Any of the four quantities characterizing the Voronoi cell (i.e. the area, the perimeter, the number of vertices and the inner angles) can be used as test variables or even some of their combinations as well. We defined the following test variables: cell area A; cell vertex (edge) N_{v}; cell chords C; inner angle α_i; round factor (RF) average $\mathrm{RF_{av}} = \overline{4\pi A/P}$; round factor (RF) homogeneity $1 - [\sigma(\mathrm{RF}_{\mathrm{av}})/\mathrm{RF}_{av}]$; shape factor A/P^2; modal factor $\sigma\ (\alpha_i)/N_v$; and the so-called 'AD factor' defined as $\mathrm{AD} = 1 - [1 - \sigma(A)/\langle A\rangle]^{-1}$, where $\sigma(A)$ is the dispersion and $\langle A\rangle$ is the average of A.

3.2.2 Minimal spanning tree

To characterize the stochastic properties of a point pattern, we use three quantities obtained from a MST: the variance of the MST edge-length $\sigma(L_{\mathrm{MST}})$; the mean MST edge-length L_{MST}; and the mean angle between edges α_{MST}.

3.2.3 Multifractal spectrum

Here, the only variable used is the $f(\alpha)$ MFR spectrum, which is a sensitive tool for testing the non-randomness of a point pattern.

An important problem is to study the sensitivity (discriminant power) of the different parameters to the different types of regularity inherent in the point pattern. In the case of a fully regular mesh, for example, A is constant, and so $AD = 0, \sigma(\alpha_i) = 0$, and both are increasing towards a fully random distribution. In the case of a patchy pattern, the distribution of the area of the Voronoi cells and the edge distribution of the MST become bimodal, reflecting the average area and the edge-length within and between the clusters, in comparison to the fully random case. In a filamentary distribution, the shape of the areas becomes strongly distorted, reflecting an increase of the modal factor in comparison to the case of patches.

Wallet & Dussert (1997) investigated the power of the VT and MST by discriminating between distributions having large and small clusters, full randomness and hard cores (random distributions but the mutual distances of the points are constrained by the size of a hard core), respectively. They concluded that the Voronoi roundness

factor did not separate small clusters and hardcore distributions, and the roundness factor homogeneity did not distinguish between small clusters and random distributions, or between random and hardcore distributions. The MST has very good discriminant power even in the case of hardcore distributions with close minimal interpoint distances.

Because the sensitivities of the variables are different on changing the regularity properties of the underlying point patterns, we can measure significant differences in one parameter but not in the other, even when these are correlated otherwise. This is not a trivial issue. In most cases, extended numerical simulations are necessary to study the statistical significance of the different parameters.

3.3 Estimation of the significance

Let ξ denote one of the 13 test variables defined in Section 3.2. The probability that $\xi < x$ occurs is given by $P(\xi < x) = F(x)$, where $F(x)$ is the probability distribution function. We approximated $F(x)$ numerically by the $F_n(x)$ empirical probability distribution function, which can be calculated by $F_n(x) = k/n$, where n is the number of simulations and k is the number of cases for which the simulated $\xi < x$ holds.

Similarly, the probability that ξ is within the interval $[x_1;x_2]$ can be obtained by $F_n(x_1) - F_n(x_2) = (x_2 - x_1)/n$. Then, the β probability that ξ is outside this region is given by $\beta = 1 - (x_2 - x_1)/n$. In the following, we suppose that the $[x_1;x_2]$ interval is symmetric to the $\overline{x}$ sample mean. To obtain the empirical distributions of the test variables, we carried out 200 simulations for each of the five samples. The numbers of simulated points were identical to those of the samples defined in Section 3.1.

We generated the fully random catalogues using Monte Carlo (MC) simulations of fully random GRB celestial positions, and taking into account the BATSE sky-exposure function (Fishman et al. 1994; Meegan et al. 2000).

Assuming that the point patterns obtained from the five samples, defined in Table 1, are fully random, we calculated the probabilities for all 13 test variables selected in Section 3.2. Based on the simulated distributions, we computed the level of significance for all 13 test variables and in all samples defined.

4 DISCUSSION OF THE STATISTICAL PROPERTIES

4.1 Significance of independent multiple tests

In Section 3.3, we calculated the numerical significance for the tests, assuming they were performed individually. The calculated significance levels are given in Table 2. In reality, however, these figures would mean significance at a certain level if we performed only that single test. Assuming that all the single tests were independent, the $P_n(m)$ probability that among n trials at least m will result in significance only by chance at a certain level is given by the following equality:

$$P_n(m) = \sum_{k=m}^{n} P_k^n. \tag{1}$$

Here, P_k^n is the binomial distribution giving the probability of k successes among n trials:

$$P_k^n = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}. \tag{2}$$

In equation (2), p denotes the probability that a single test has given a significant result only by chance. It is easy to see that this equation results in $P_n(1) = 1 - (1-p)^n \approx np$, which gives a significance of $1 - np$ approximately instead of $1 - p$. This means, for example, that a single test resulting in $1 - p = 0.95$ significance is reduced to $1 - 0.95^{13} = 0.49$ if we perform $n = 13$ independent tests but only one results in $1 - p = 0.95$ significance.

Inspecting Table 2, which lists the calculated numerical significance of single tests, we can infer that the short1 sample has four tests with a significance of $1 - p = 0.95$ or more. Taking into account the calculations at the end of the previous paragraph, however, we have to emphasize that the calculated numerical significance, based on the individual probability distribution of the test variables separately, does not have its original meaning. Significance refers to the certainty of rejecting the null hypothesis on the basis of the series of tests as a whole. Applying equation (1) with $m = 4$ and $n = 13$, we obtain a significance of 99.69 per cent. Applying the same sequence of arguments to the short2 sample, we obtain a figure of only 86.46 per cent ($m = 2$ and $n = 13$). For the intermediate, long1

Table 2. Calculated significance levels for the 13 test variables and the five samples. A calculated numerical significance greater than 95 per cent is given in bold face.

Name	Var	Short1	Short2	Intermediate	Long1	Long2
Cell area	A	36.82	29.85	94.53	79.60	82.59
Cell vertex (edge)	N_v	36.82	87.06	2.99	26.87	7.96
Cell chords	C	47.26	52.24	18.91	84.58	54.23
Inner angle	α_i	**96.52**	21.39	87.56	37.81	63.18
RF average	$\overline{4\pi A/P}$	65.17	**99.98**	33.83	10.95	86.07
RF homogeneity	$1 - [\sigma\ (\mathrm{RF}_{av})/\mathrm{RF}_{av}]$	19.90	24.38	58.71	55.72	32.84
Shape factor	A/P^2	91.04	94.03	90.05	55.22	63.68
Modal factor	$\sigma\ (\alpha_i)/N_v$	**97.51**	1.99	7.46	56.22	8.96
AD factor	$1 - \{1 - [\sigma(A)/\langle A\rangle]\}^{-1}$	32.84	25.37	11.44	**95.52**	52.74
MST variance	$\sigma(L_{\mathrm{MST}})$	52.74	38.31	22.39	13.93	59.70
MST average	L_{MST}	**97.51**	7.46	89.05	56.72	8.96
MST angle	α_{MST}	85.07	14.43	36.82	73.63	60.70
MFR spectra	$f(\alpha)$	**95.52**	**96.02**	**98.01**	73.63	36.32
Binomial test	(equation 1 with $p = 0.05$)	**99.69**	86.46	51.33	51.33	–
Squared Euclidean	distance	**99.90**	**99.98**	**98.51**	93.03	36.81

and long2 samples, we cannot obtain figures above the 95 per cent significance level.

There may be a serious concern, however, with the results obtained above. Namely, the basic requirement of the independence of the single tests is not fulfilled in our case. On the contrary, there are strong correlations between the test variables in Table 2. In Section 4.2, we try to outline an approach that takes into account the correlations between the test variables and overcomes this difficulty.

4.2 Evaluation of the joint significance levels

We assigned to every MC simulated sample 13 values of the test variables and, consequently, a point in the 13-dimensional (13D) parameter space. Completing 200 simulations in all of the subsamples, in this way we obtain a 13D sample representing the joint probability distribution of the 13 test variables. Using a suitable chosen measure of distance of the points from the sample mean, we can obtain a stochastic variable characterizing the deviation of the simulated points from the mean only by chance. An obvious choice would be the squared Euclidean distance.

In the case of a Gaussian distribution with unit variances and without correlations, this would result in a χ^2 distribution of 13 degrees of freedom. The test variables in our case are correlated and have different scales. Before computing the squared Euclidean distances, we transformed the test variables into non-correlated variables with unit variances. Because of the strong correlation between some of the test variables, we can assume that the observed quantities can be represented with non-correlated variables fewer in number. Factor analysis (FA) is a suitable way to represent the correlated observed variables with non-correlated variables fewer in number.

As our test variables are stochastically dependent following Wallet & Dussert (1998), we attempted to represent these by fewer non-correlated hidden variables, supposing that

$$X_i = \sum_{j=1}^{k} a_{ij} f_j + s_i \qquad i = 1, \ldots, p; \qquad k < p. \tag{3}$$

In equation (3), X_i, f_j and s_i denote the test variables ($p = 13$ in our case), the hidden variables and a noise term, respectively. Equation (3) represents the basic model of FA. After making some reasonable assumptions (Kendall & Stuart 1973), k can be constrained by the following inequality

$$k \leq (2p + 1 - \sqrt{8p + 1})/2, \tag{4}$$

which gives $k \leq 8.377$ in our case.

Factor analysis is a common ingredient of professional statistical software packages (BMDP,® SAS,® S-plus,® SPSS®, etc.). The default solution for identifying the factor model is to perform principal component analysis (PCA). We kept as many factors as were meaningful with respect to equation (4). Taking into account the constraint imposed by equation (4), we retained eight factors. In this way, we projected the joint distribution of the test variables in the 13D parameter space into an eight-dimensional parameter space defined by the non-correlated f_i hidden variables.

The f_j hidden variables in equation (3) are non-correlated and have zero means and unit standard deviations. Using these variables, we defined the following squared Euclidean distance from the sample mean:

$$d^2 = f_1^2 + f_2^2 + \ldots + f_k^2 \qquad (k = 8 \text{ in our case}). \tag{5}$$

If the f_j variables had strictly Gaussian distributions equation (5) would define a χ^2 variable of k degrees of freedom.

4.3 Statistical results and their interpretations

In addition to the significance obtained by the binomial test in Section 4.1 using the distribution of the squared Euclidean distances, defined by equation (5), we can obtain further information about whether a BATSE sample represented by a point in the parameter space of the test variables deviates only by chance or whether it significantly differs from the fully random distribution.

In all categories (short1, short2, intermediate, long1 and long2) we carried out 200, altogether 1000, simulations. We calculated the d^2 squared distances for all simulations and compared these with those of the BATSE samples in Table 1. Fig. 4 shows a histogram of the simulated squared distances along with those of the BATSE samples. The full line represents a χ^2 distribution of $k = 8$ degrees of freedom. Fig. 4 clearly shows that the departures of samples short1 and short2 exceed all those of the simulated points. The probabilities that these deviations are non-random are 99.9 and 99.98 per cent, respectively.

The full randomness of the angular distribution of the long GRBs, in contrast to the regularity of the short and to some extent the intermediate GRBs, points towards the differences in the angular distribution of their progenitors. The recent discovery of the afterglow in some short GRBs indicates that these events are associated with the old stellar population (Fox et al. 2005) accounting probably for the mergers of compact binaries. This is in contrast to the long bursts resulting from the collapse of very massive stellar objects in young star-forming regions. The differences in progenitors also reflects the differences between the energy released by the short and long GRBs.

Unfortunately, little can be said about the physical nature of the intermediate class. The statistical studies (Horváth et al. 2006, and references therein) suggest the existence of this subgroup, at least from a purely statistical point of view. The non-random sky-distribution also occurs here, but its physical origin is not fully solved yet (Horváth et al. 2006).

5 SUMMARY AND CONCLUSIONS

We carried out additional studies on the degree of the randomness in the angular distribution of samples selected from the BATSE catalogue. According to the T_{90} durations and P_{256} peak fluxes of the GRBs in the catalogue, we defined five groups: short1 ($T_{90} < 2$ s and $0.65 < P_{256} < 2$); short2 ($T_{90} < 2$ s and $0.65 < P_{256}$); intermediate ($2 \leqq T_{90} \leqq 10$ s and $0.65 < P_{256}$); long1 ($T_{90} > 2$ s and $0.65 < P_{256} < 2$); long2 ($T_{90} > 10$ s and $0.65 < P_{256}$).

To characterize the statistical properties of the point patterns, given by the samples, we defined 13 test variables based on the VT, the MST and MFR spectra. For all five GRB samples defined, we carried out 200 numerical simulations assuming fully random angular distribution and taking into account the BATSE exposure function. The numerical simulations enabled us to define empirical probabilities for testing the null hypothesis (i.e. the assumption that the angular distributions of the BATSE samples are fully random).

Because we performed 13 single tests simultaneously on each subsample, the significance obtained by calculating it separately for each test cannot be treated as a true indication for deviating from the fully random case. At first, we supposed that the test variables were independent and, making use the binomial distribution, we computed the probability of obtaining a significant deviation in at least one of the variables only by chance. In fact, some of the test variables are strongly correlated. To concentrate the information on the non-randomness experienced by the test variables, we assumed that

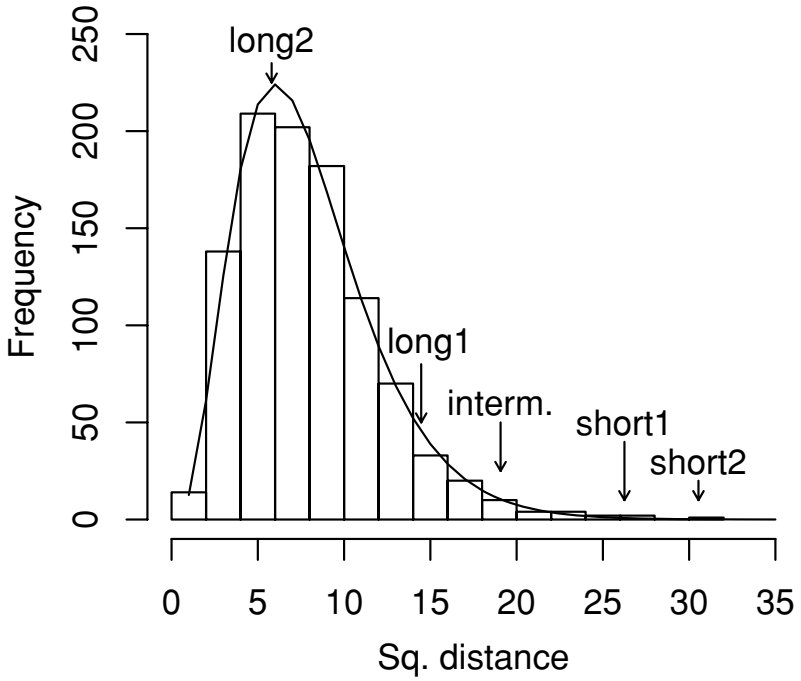

Figure 4. Distribution of the squared Euclidean distances of the simulated samples from the stochastic mean of the f_i hidden variables (factors) in the eight-dimensional parameter space. There are altogether 1000 simulated points. The full line marks a χ^2 distribution of eight degrees of freedom, normalized to the sample size. The distances of the BATSE samples are also indicated. The departures of samples short1 and short2 exceed all those of the simulated points. The probabilities that these deviations are non-random equal 99.9 and 99.98 per cent, respectively.

they can be represented as a linear combination of non-correlated hidden factors fewer in number. Actually, we estimated $k = 8$ as the number of hidden factors. Making use of the hidden factors, we computed the distribution of the squared Euclidean distances from the mean of the simulated variables. Comparing the distribution of the squared Euclidean distances of the simulated variables with the BATSE samples, we concluded that the short1 and short2 groups deviate significantly (99.90 and 99.98 per cent, respectively) from the full randomness, but this is not the case for the long samples. For the intermediate group, the squared Euclidean distances also give significant deviation (98.51 per cent).

ACKNOWLEDGMENTS

This study was supported by OTKA grant No. T048870, by a Bolyai Scholarship (IH), by a Research Program MSM0021620860 of the Ministry of Education of the Czech Republic, and by a GAUK grant No. 46307 (AM). We are indebted to an anonymous referee for his/her valuable comments and suggestions.

REFERENCES

Adami C., Mazure A., 1999, A&AS, 134, 393
Aschwanden M. J., Parnell C. E., 2002, ApJ, 572, 1048
Balázs L. G., Mészáros A., Horváth I., 1998, A&A, 339, 1
Balázs L. G., Mészáros A., Horváth I., Vavrek R., 1999, A&AS, 138, 417
Balázs L. G., Bagoly Z., Horváth I., Mészáros A., Mészáros P., 2003, A&A, 401, 129
Barrow J. D., Bhavsar S. P., Sonoda D. H., 1985, MNRAS, 216, 17
Bhavsar S. P., Lauer D. A., 1996, in Kafatos M. C., Kondo Y., eds, Proc. IAU Symp. 168, Examining the Big Bang and Diffuse Background Radiations. Kluwer, Dordrecht, p. 517
Bhavsar S. P., Splinter R. J., 1996, MNRAS, 282, 1461
Bottorff M., Ferland G., 2001, ApJ, 549, 118
Casuso E., Beckman J., 2002, PASJ, 54, 405
Célérier M. N., Thieberger R., 2001, A&A, 367, 449
Chappell D., Scalo J., 2001, ApJ, 551, 71
Coles P., 1991, Nat, 349, 288
Coles P., Barrow J. D., 1990, MNRAS, 244, 557
Datta S., 2003, A&A, 401, 193
Diggle P. J., 1983, Statistical Analysis of Spatial Point Patterns. Academic Press, London
Doroshkevich A. G., Turchaninov V., 2001, in Banday A. J., Zaroubi S., Bartelmann M., eds, Proc. MPA/ESO/MPE Workshop, Mining the Sky. Springer-Verlag, Berlin, p. 283
Doroshkevich A. G., Gottlöber S., Madsen S., 1997, A&AS, 123, 495
Elmegreen B. G., 2002, ApJ, 564, 773
Fishman G. J. et al., 1994, ApJS, 92, 229
Fox D. B. et al., 2005, Nat, 437, 845
Gaite J., Manrubia S. C., 2002, MNRAS, 335, 977
Giraud E., 2000, ApJ, 544, L41
Hakkila J., Haglin D. J., Pendleton G. N., Mallozzi R. S., Meegan C. A., Roiger R. J., 2000, ApJ, 538, 165
Hakkila J., Giblin T. W., Roiger R. J., Haglin D. J., Paciesas W. S., Meegan C. A., 2003, ApJ, 582, 320
Horváth I., 1998, ApJ, 508, 757
Horváth I., 2002, A&A, 392, 791
Horváth I., Balázs L. G., Bagoly Z., Ryde F., Mészáros A., 2006, A&A, 447, 23
Icke V., van de Weygaert R., 1991, QJRAS, 32, 85
Ikeuchi S., Turner E. L., 1991, MNRAS, 250, 519
Irwin J. A., Widrow L. M., English J., 2000, ApJ, 529, 77
Kawaguchi T., Mineshige S., Machida M., Matsumoto R., Shibata K., 2000, PASJ, 52, L1
Kendall M. G., Stuart A., 1973, The Advanced Theory of Statistics. Charles Griffin, London
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
Kruskal J. B., 1956, Proc. Am. Math. Soc., 7, 48

Krzewina L. G., Saslaw W. C., 1996, MNRAS, 278, 869
Litvin V. F., Matveev S. A., Mamedov S. V., Orlov V. V., 2001, Pis'ma v Astron. Zhurnal, 27, 489
Magliocchetti M., Ghirlanda G., Celotti A., 2003, MNRAS, 343, 255
Meegan C. A. et al., 2000, BATSE Gamma-Ray Burst Catalogue, http://gammaray.msfc.nasa.gov/batse/grb/catalog
Mészáros P., 2006, Rep. Prog. Phys., 69, 2259
Mészáros A., Štoček J., 2003, A&A, 403, 443
Mészáros A., Bagoly Z., Vavrek R., 2000a, A&A, 354, 1
Mészáros A., Bagoly Z., Horváth I., Balázs L. G., Vavrek R., 2000b, ApJ, 539, 98
Mukherjee S., Feigelson E. D., Jogesh Babu G., Murtagh F., Fraley C., Raftery A., 1998, ApJ, 508, 314
Paladin G., Vulpiani A., 1987, Phys. Rep., 156, 1
Pan J., Coles P., 2000, MNRAS, 318, L51
Pan J., Coles P., 2002, MNRAS, 330, 719
Pásztor L., Tóth L. V., 1995, in Shaw R. A., Payne H. E., Hayes J. J. E., eds, ASP Conf. Ser. Vol. 77, ADASS IV. Astron. Soc. Pac., San Francisco, p. 319
Pendleton C. N. et al., 1997, ApJ, 489, 175
Prim R. C., 1957, Bell Syst. Techn. Journ., 36, 1389
Ramella M., Boschin W., Fadda D., Nonino M., 2001, A&A, 368, 776
Selman F. J., Melnick J., 2000, ApJ, 534, 703
Semelin B., Combes F., 2002, A&A, 387, 98
Stoyan D., Stoyan H., 1994, Fractals, Random Shapes and Point Fields. Wiley, New York
Subba Rao M. U., Szalay A. S., 1992, ApJ, 391, 483
Tanvir N. R., Chapman R., Levan A. J., Priddey R. S., 2005, Nat, 438, 991
Tatekawa T., Maeda K., 2001, ApJ, 547, 531
Tikhonov A. V., 2002, Astrophys., 45, 79
Vavrek R., Balázs L. G., Epchtein N., 2001, in Montmerle T., André P., eds, ASP Conf. Ser. Vol. 243, From Darkness to Light. Astron. Soc. Pac., San Francisco, p. 149
Voronoi G., 1908, J. Reine Angew. Math., 134, 198
Wallet F., Dussert C., 1997, J. Theor. Biol., 187, 437
Wallet F., Dussert C., 1998, Europhys. Lett., 42, 493
van de Weygaert R., 1994, A&A, 283, 361
Yahagi H., Mori M., Yoshii Y., 1999, ApJS, 124, 1
Zaninetti L., 1995, A&AS, 109, 71
Zhang B., Mészáros P., 2004, IJMPA, 19, 2385

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