RAA 2019 Vol. **19** No. **2**, 28(8pp) doi: 10.1088/1674–4527/19/2/28

http://www.raa-journal.org *http://iopscience.iop.org/raa*

Research in Astronomy and Astrophysics

Cosmology using long gamma-ray bursts: statistical analysis of errors in calibrated data

Meghendra Singh[1], Shashikant Gupta[2], Amit Sharma[3] and Satendra Sharma[4]

[1] Dr. A.P.J. Abdul Kalam Technical University, Lucknow 226021, Uttar Pradesh, India; *meghendrasingh_db@yahoo.co.in*
[2] G.D. Goenka University, Gurgaon, Haryana 122103, India
[3] Amity University Haryana, Gurgaon (Manesar), Haryana 122413, India
[4] Yobe State University, Damaturu, Yobe State, Nigeria

Received 2018 March 18; accepted 2018 August 29

Abstract We investigate direction dependence and non-Gaussian features in high-z cosmological data using Δ_{χ^2} and Δ_{χ} statistics and the Kolmogorov-Smirnov test. These techniques are applied on a set of calibrated long gamma-ray bursts (GRBs) and its combination with recent Type Ia supernovae data (Union2). Our statistical analysis shows a weak but consistent direction dependence in both the data sets. The analysis also indicates a non-Gaussian nature of errors in both data sets.

Key words: gamma-ray bursts — observational cosmology — cosmological parameters — Gaussian distribution — supernovae

1 INTRODUCTION

Gamma-ray bursts (GRBs) are among the most energetic explosions in the Universe. Their high energy photons in the gamma-ray band are almost unaffected by dust extinction, and as a result they are detectable up to very high redshifts (Salvaterra et al. 2009; Cucchiara et al. 2011). Long GRBs (for which pulse duration is greater than 2 s) can be used to investigate the Universe at high redshifts, which is difficult to access by Type Ia supernovae (SNe Ia); thus long GRBs have been proposed as a complementary probe to SNe Ia. However, due to an extensive variety of isotropic equivalent luminosities and energy outputs (Ghirlanda et al. 2006), GRBs cannot be treated as standard candles. Efforts have been made to calibrate long GRBs using the following empirical relations:

- Isotropic equivalent radiated energy ($E_{\rm iso}$) and peak energy (E_p) correlation for GRBs (Amati et al. 2008).
- Peak energy (E_p) in vFv spectra and collimated emission (E_γ) correlation for GRBs (Ghirlanda et al. 2004).

Unfortunately, physical interpretation of the above correlations is not well understood. Thus, standardization of these correlations is subject to availability of low redshift GRBs. Since not many low redshift GRBs are available (Kodama et al. 2008; Li et al. 2008), the calibration is obtained by assuming a specific cosmology (specific values of cosmological parameters). As a result, the GRB data obtained from this calibration depend on cosmology, and hence cannot be further used for estimation of cosmological parameters. This problem is known as the "circularity problem" in GRB cosmology. It should be noted that while deriving distances to SNe Ia, sometimes the numerical value of the Hubble constant is inserted by hand (Amanullah et al. 2010) which makes it cosmology-dependent. However, one can easily overcome this problem by making use of nearby SNe Ia.

In order to use GRBs as cosmological probes, one needs specific statistical techniques to avoid the circularity problem. Many cosmology-independent methods have been proposed to achieve this, e.g., the collimation-corrected energy method (Ghirlanda et al. 2004), the luminosity indicator method (Liang & Zhang 2005) derived from 157 SNe Ia, the Bayesian method (Firmani et al. 2005) and the Markov Chain Monte Carlo (MCMC) method (Li et al. 2008). Recently Liu & Wei (2015) have used the Padé approximation to calibrate high-redshift

GRBs which can be used to constrain cosmological models.

The cosmological principle (hereafter CP) has a significant role in modern cosmology; it states that the Universe is homogeneous and isotropic at large scales (Weinberg 2008). However, it is essential and important to check if the CP is consistent with the latest observational data. As more high-redshift GRB data are now available (Liang et al. 2008; Wei 2010), it becomes possible to search for any possible anisotropy in the high redshift Universe. Earlier, Gupta et al. (2008), Gupta & Saini (2010) and Gupta & Singh (2014) (hereafter GS14) investigated the direction dependence using techniques based on extreme value statistics through different sets of SNe Ia. As a biproduct of the technique, they also obtained information about non-Gaussian features in the data. In the present work, we investigate the direction dependence and non-Gaussianity in the latest long GRB data.

The paper is structured in the following way. The data analyzed for the current work are briefly discussed in Section 2, while the methodology, i.e., the Δ_{χ^2} statistic, Δ_χ statistic and Kolmogorov-Smirnov (KS) test are explained in Section 3. Results and conclusions are presented in Section 4 and Section 5 respectively.

2 DATA SET

In this paper we consider a sample of 79 calibrated long GRBs from Liu & Wei (2015) at high-redshift ($z > 1.4$), which is claimed to be free from the circularity problem and thus can be used to test cosmological models. For our analysis, we use Union2 data from Amanullah et al. (2010), which contain 557 SNe Ia with redshift up to ($z = 1.4$) and their combination (hereafter UNGRB) having 636 data points.

3 METHODOLOGY

3.1 The Δ_{χ^2} Statistic

For a given GRB, the difference between apparent magnitude ($m(z)$) and absolute magnitude (M) is called distance modulus (μ)

$$\mu(z) = m(z) - M\,, \tag{1}$$

where $m(z)$ depends on the intrinsic luminosity of a GRB and z is its redshift. We can also express distance modulus (measured in Mpc) as

$$\mu(z) = 5\log\left(D_L(z)\right) + 25\,. \tag{2}$$

For a ΛCDM Universe, D_L can be written as

$$D_L(z) = \frac{c(1+z)}{H_0}\int_0^z \frac{dx}{h(x)}\,, \tag{3}$$

where $h(z;\Omega_{\rm M},\Omega_{\rm X}) = H(z;\Omega_{\rm M},\Omega_{\rm X})/H_0$, and thus depends only on the cosmological parameters matter density $\Omega_{\rm M}$ and dark energy density $\Omega_{\rm X}$. For a flat ΛCDM Universe, we calculate best fit values for a complete data set by χ^2 minimization, which we express as

$$\chi^2 = \Sigma_{i=1}^{N}\left[(\mu^i_{\rm obs} - \mu^i_{\rm th})/\sigma_i\right]^2\,, \tag{4}$$

where the total number of data points is given by N. $\mu^i_{\rm obs}$ is the measured distance modulus from the data and $\mu^i_{\rm th}$ is the calculated theoretical distance modulus. Observed standard error in μ^i is denoted by σ_i. We obtain χ_i for each GRB as

$$\chi_i = [\mu^i_{\rm obs} - \mu^i_{\rm th}(z_i;\Omega_{\rm M})]/\sigma_i\,, \tag{5}$$

where $\mu^i_{\rm th}(z_i;\Omega_{\rm M})$ is calculated by using the best fit values of Ω_M and H_0. We divide the data into two hemispheres labeled by the direction vector $\hat{n}$, and take the difference of the χ^2 computed for the two hemispheres separately ($\Delta\chi^2_{\hat{n}}$). Now to calculate the maximum absolute difference, we rotate the direction vector $\hat{n}$ and calculate the Δ_{χ^2} several times (for further details see GS14)

$$\Delta_{\chi^2} = \max\{|\Delta\chi^2_{\hat{n}}|\}\,. \tag{6}$$

We generate the distribution of Δ_{χ^2} by shuffling the data values z_i, μ^i and σ_i over the SN positions and refer to it as a bootstrap distribution. The distribution when plotted looks skewed as expected from extreme value theory (Haan & Ferreira 2006). A greater distance of the original Δ_{χ^2} (without shuffling the data values) from the peak of the bootstrap distribution would indicate that the SN positions in the data are important, i.e., it indicates direction dependence in the data. Since the analytic expression for distribution of Δ_{χ^2} is not available, one can obtain the theoretical distribution numerically and compare it with the bootstrap distribution. If the errors (σ_is) in the data follow a normal distribution, then the χ_is in Equation (5) would also obey a standard normal distribution. To obtain the theoretical distribution, one can generate χ_is using a standard normal distribution instead of Equation (5). If the bootstrap and theoretical distributions do not match, this would indicate that χ_is in Equation (5) do not follow a standard normal distribution and hence the errors (σ_is) have a non-Gaussian nature.

3.2 The Delta-Chi ($\Delta\chi$) Statistic

χ_i defined in Equation (5) could be positive or negative, but this information is lost when we square it. Thus, the Δ_{χ^2} statistic does not contain information about the SN being above or below the fit. We consider another method

without squaring the χ_is that does contain information regarding whether the SN at a given redshift is closer or farther from us. An additional advantage of this method is that the distribution of maxima can be calculated analytically as discussed below.

We have two subsets of data represented by two hemispheres and labeled by the direction vector $\hat{n}$. We also have $N_{\rm north}$ and $N_{\rm south}$ SNe, where N is the total number of SNe such that $N = N_{\rm north} + N_{\rm south}$, and define the quantity

$$\Delta\chi_{\hat{n}} = \frac{1}{\sqrt{N}} \left(\sum_{i=1}^{N_{\rm north}} \chi_i - \sum_{j=1}^{N_{\rm south}} \chi_j \right) . \quad (7)$$

From Equation (7), it is clear that $\langle \Delta\chi_{\hat{n}} \rangle = 0$ and $\langle (\Delta\chi_{\hat{n}})^2 \rangle = 1$. It follows from the central limit theorem that if $N \gg 1$, the quantity $\Delta\chi_{\hat{n}}$ obeys a normal distribution with zero mean and unit standard deviation. To obtain the maximum absolute difference, we maximize $\Delta\chi_{\hat{n}}$ by varying the direction $\hat{n}$ across the sky

$$\Delta_\chi = \max\{|\Delta\chi_{\hat{n}}|\} . \quad (8)$$

Extreme value theory shows that the distribution of $\Delta\chi$ follows a two parameter distribution known as a Gumbel distribution and is given by (Haan & Ferreira 2006)

$$\begin{aligned} P(\Delta) =& \frac{1}{s} \exp\left[-\frac{\Delta - m}{s} \right] \\ & \times \exp\left[-\exp\left(-\frac{\Delta - m}{s} \right) \right] , \end{aligned} \quad (9)$$

where the position parameter m and the shape parameter s can be determined analytically in the theoretical limit of the Δ_χ statistic. In the limit ($N_{\rm dir} \gg 1$, where $N_{\rm dir}$ is number of independent directions) the parameters are given by

$$\begin{aligned} m &= \sqrt{2 \log N_{\rm dir} - \log\log N_{\rm dir} - \log 4\pi} , \\ s &= \frac{1}{m} . \end{aligned} \quad (10)$$

Here it is assumed that the total number of SNe is large, i.e., $N \gg 1$, since the distribution for $\Delta\chi_{\hat{n}}$ becomes Gaussian only in this limit.

3.3 Kolmogorov-Smirnov Test

Finally we apply the KS test which is described below. Originally, this approach was presented in Singh et al. (2016a) and in Singh et al. (2016b) to investigate non-Gaussianity in the HST Key Project Data and in SNe data.

If $\mu^i_{\rm th}(z)$ is the theoretical value of the distance modulus for the i^{th} GRB at redshift z, then the observed value $\mu^i_{\rm obs}$ will be

$$\mu^i_{\rm obs} = \mu^i_{\rm th}(z) \pm \sigma_i . \quad (11)$$

As discussed earlier, χ_i should follow a standard Gaussian distribution. To check this, we use the KS test (Press et al. 2002). We define our null hypothesis as "The errors in GRB data are drawn from a Gaussian distribution". If the null hypothesis is true, the χ_is in Equation (5) would follow a standard Gaussian distribution. To test this we use Matlab function *kstest[h,p,k,cv]* where p is the probability of the data errors being drawn from a Gaussian distribution and k represents the maximum distance between the cumulative distribution function (CDF) of χ_is with that of a standard Gaussian. cv is the critical value which is decided by the significance level (α). Different values of α indicate different tolerance levels for false rejection of the null hypothesis. cv is the critical value of the probability to obtain the data set in question given the null hypothesis and can be compared with p. A value $h = 1$ is returned by the test if $p < cv$ and the null hypothesis is rejected. On the other hand if $p > cv$, h remains 0 and the null hypothesis is not rejected.

4 RESULTS

First, we calculate the best fit values of matter density ($\Omega_{\rm M}$) and Hubble constant (H_0) for both GRB and SN data. It is clear from Table 1 that GRB data favor a smaller value of $\Omega_{\rm M}$ and a larger value of H_0 compared to Union2. A smaller value of χ^2 in the case of GRBs indicates that the errors may have been overestimated. The best fit values of $\Omega_{\rm M}$ and H_0 are the same for Union2 and UNGRB data, reflecting the fact that SNe dominate in the combined data set. The directions of maximum discrepancy and the values of (Δ_{χ^2}) and (Δ_χ) are shown in Tables 2 and 3 respectively. The values of p, k and cv for various data sets are presented in Table 4.

For reference, the results of simulations for each data set are plotted in panel (a) of Figures 1, 2, 3, 4 and 5. A specific bias related to the bootstrap distribution was discussed in GS14. The theoretical distribution is obtained by assuming that the χ_is follow a normal distribution with zero mean and unit standard deviation, thus theoretical χ_is are unbounded. On the other hand, the bootstrap distribution is prepared by shuffling through a *specific realization* of χ_i, which has a maximum value for some SNe/GRBs. Thus the χ_is in the bootstrap distribution have an upper bound, causing the bootstrap distribution to be shifted slightly to the left of the theoretical distribution. This can also be seen in the figures displaying the results of simulations in Figures 1, 2, 3, 4 and 5.

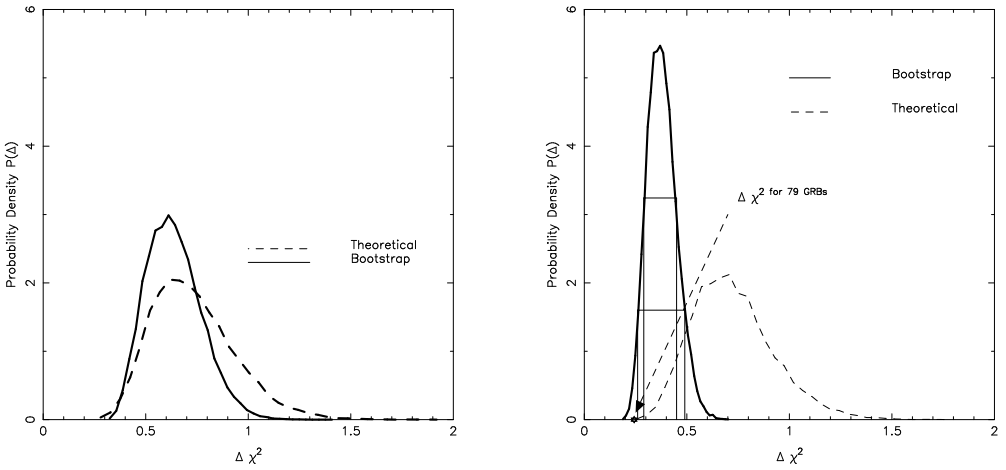


Fig. 1 (a) Result of the simulation for 79 GRBs, which shows a comparison between bootstrap and theoretical probability distributions. Positions of GRBs were generated randomly on the sky. (b) The bootstrap and theoretical probability distributions of 79 GRBs for the Δ_{χ^2} statistic. The shape of the bootstrap distribution is different from that in Figure 1(a). A shift in the bootstrap on the left indicates non-Gaussianity.

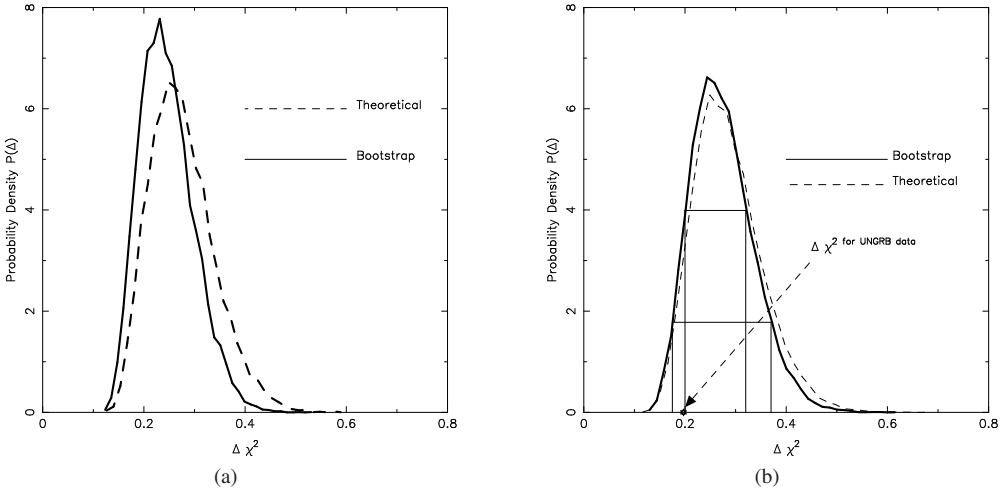


Fig. 2 (a) Result of the simulation for UNGRB, which shows a comparison between bootstrap and theoretical probability distributions. (b) The bootstrap and theoretical probability distributions of UNGRB for the Δ_{χ^2} statistic. In comparison with Figure 2(a), the bootstrap distribution that appears on the right side shows an indication of non-Gaussianity.

Table 1 Best fit values of cosmological parameters (ΛCDM) for different data sets.

Set	Data	$\Omega_{\rm M}$	H_0	χ^2/dof
GRB	79	0.24	79.5	0.42
Union2	557	0.27	70.0	0.97
UNGRB	636	0.27	70.0	0.90

Table 2 Direction of maximum Δ_{χ^2} for different data sets. Long and Lat stand for longitude and latitude respectively in the Galactic coordinate system.

Set	Data	Δ_{χ^2}	Long	Lat
GRB	79	0.243	215.7	26.6
Union2	557	0.220	65.5	55.8
UNGRB	636	0.197	45.8	49.2

4.1 Results for Δ_{χ^2} Statistic

- **GRBs**: The theoretical and bootstrap distributions for the set of 79 GRBs are shown in panel (b) of Figure 1. Comparison with Figure 1(a) demonstrates that the shift in bootstrap is much larger than expected, suggesting the non-Gaussian nature of the errors. The lo-

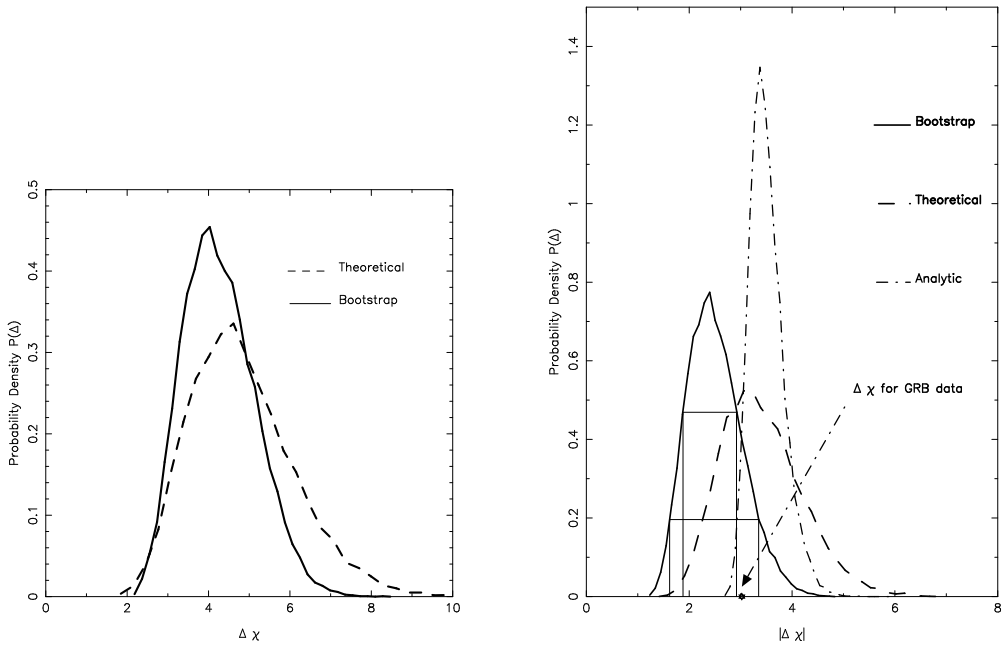


Fig. 3 (a) Result of the simulation for 79 GRBs, which shows a comparison between bootstrap and theoretical probability distributions. (b) The bootstrap, theoretical and analytic probability distributions for the Δ_χ statistic for GRBs. The analytic distribution obtained from Equation (9) with parameters defined in Equation (10). The bootstrap distribution shows anisotropy at the 1σ level.

Table 3 Direction of maximum Δ_χ for different data sets. Long and Lat stand for longitude and latitude respectively in the Galactic coordinate system.

Set	Data	Δ_χ	Long	Lat
GRB	79	3.02	317.1	37.4
Union2	557	4.32	129.5	17.8
UNGRB	636	4.45	129.4	17.7

Table 4 Results of KS test for different data sets at 1% significance level (α).

Set	p value	k	Cv
GRB	0.5519	0.0875	0.1806
Union2	0.7328	0.0288	0.0572
UNGRB	0.4953	0.0327	0.0536

cation of Δ_{χ^2} for GRBs is nearly 2σ away from the maximum of the bootstrap distribution, which suggests direction dependence in the data.

- **Union2**: The results of the Δ_{χ^2} statistic for Union2 data have been discussed in GS14. The bootstrap and theoretical distributions are shown in figure 5 of GS14.
- **UNGRB**: Figure 2(b) presents the two distributions for UNGRB. Comparison with Figure 2(a) shows that the leftward shift of the bootstrap distribution is much smaller than expected, indicating slight non-Gaussianity in the residuals. Δ_{χ^2} for UNGRB is about 1σ away from the peak which indicates weak direction dependence in UNGRB data. It should be noted that the direction dependence is small as compared to GRB data.

4.2 Results for Δ_χ Statistic

- **GRB**: The Δ_χ statistic has an advantage since the distribution can be calculated analytically. The theoretical, bootstrap and analytic distributions for the set of 79 GRBs are plotted in Figure 3(b). The Δ_χ for GRBs is more than 1σ away from the mode of the bootstrap distribution which shows a slight direction dependence in GRBs. The analytic distribution is obtained using Equation (9) with parameters defined in Equation (10). The shape of the analytic distribution is quite different. The mismatch could arise if all the directions are not independent. In this case, the number of actual independent directions would be less than N_d and hence the actual value of m would be smaller than what is calculated in Equation (10). This also makes the spread larger in bootstrap and theoretical distributions compared to the analytic ones.
- **Union2**: Theoretical, bootstrap and analytic distributions for Union2 are plotted in Figure 4(b). Since the theoretical χ_is are unbounded while the bootstrap

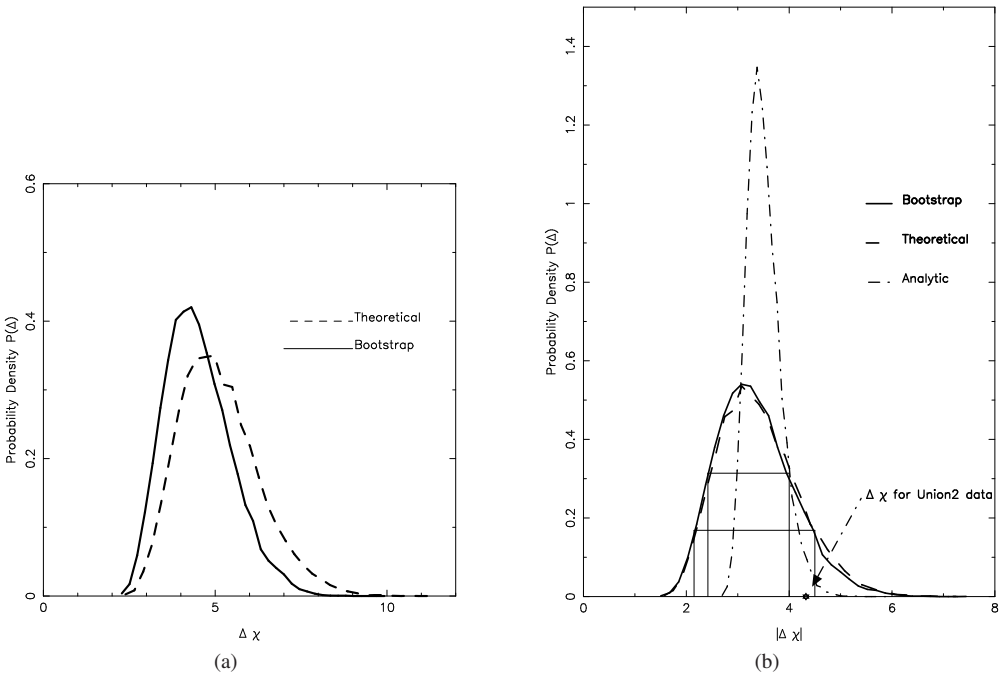


Fig. 4 (a) Result of the simulation for 557 SNe, which shows a comparison between bootstrap and theoretical probability distributions. (b) The bootstrap, theoretical and analytic probability distributions for the Δ_χ statistic for Union2 data. The bootstrap distribution indicates anisotropy away from the 1σ level.

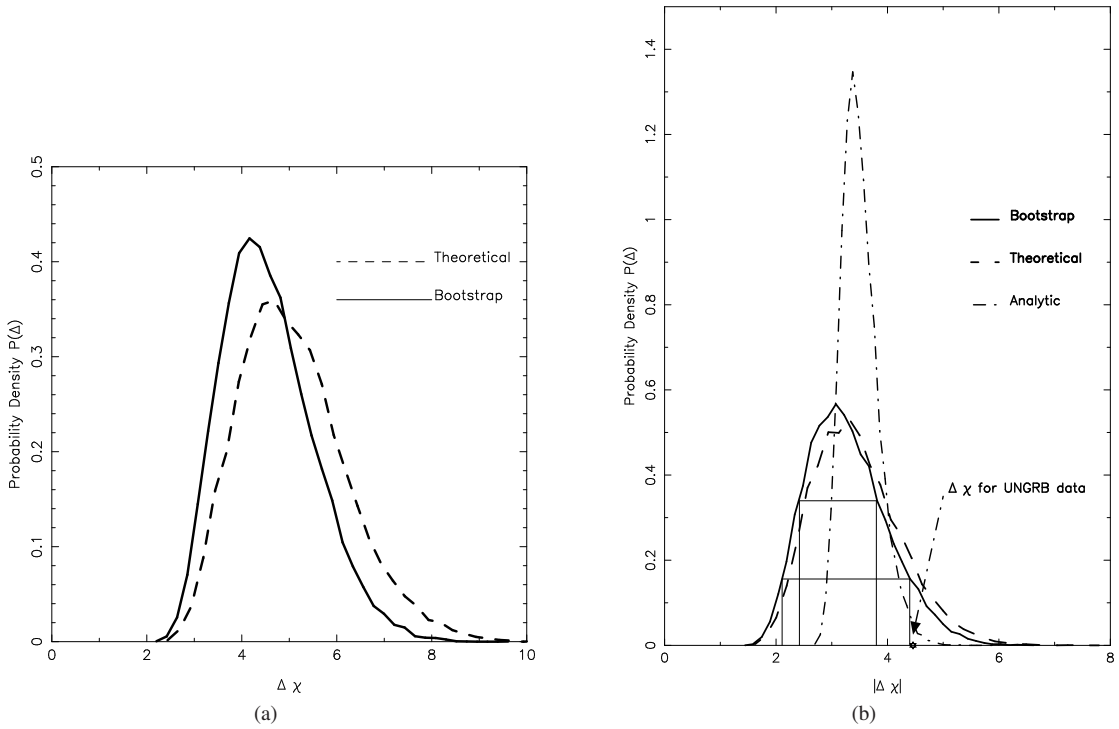
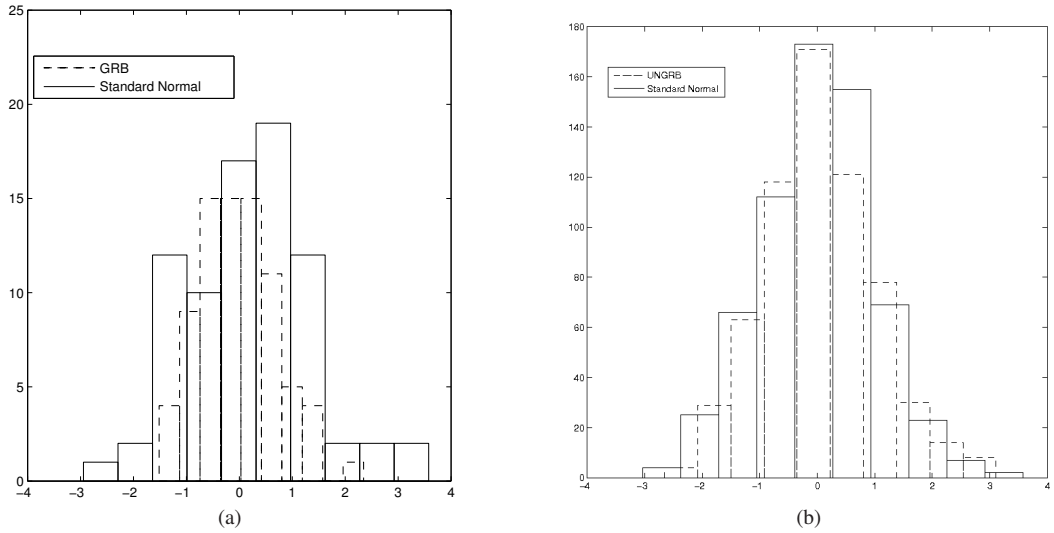
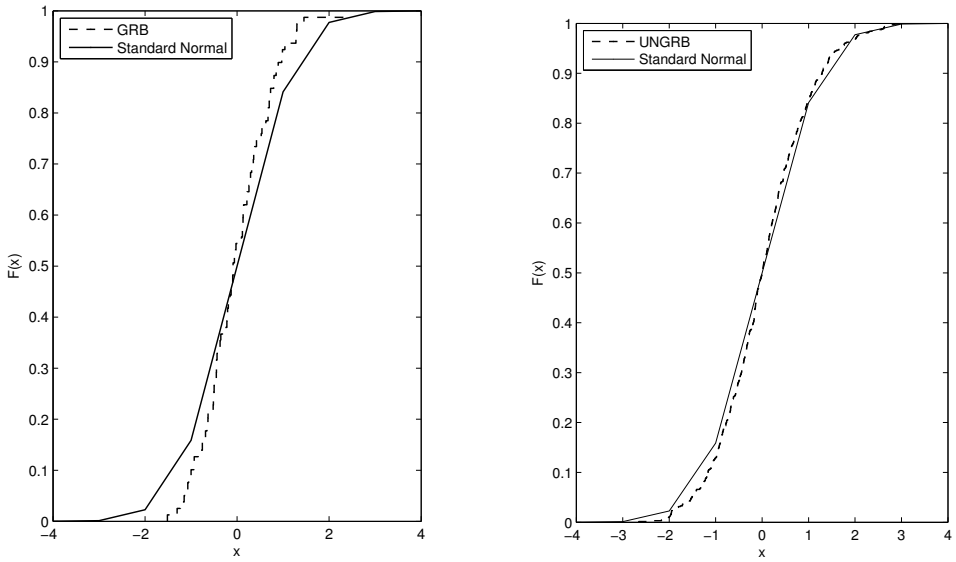


Fig. 5 (a) Result of the simulation for 636 data points, which shows a comparison between bootstrap and theoretical probability distributions. (b) The bootstrap, theoretical and analytic probability distributions for the Δ_χ statistic for UNGRB data. The bootstrap distribution indicates anisotropy at the 2σ level.



(a) (b)

Fig. 6 Histograms of χ_i for GRB and UNGRB data are compared with those of a standard normal distribution.



(a) (b)

Fig. 7 Comparison of the CDFs of χ_i for GRB and UNGRB data with their Gaussian CDFs. *Solid curves* represent the Gaussian CDFs.

χ_is are bounded, the bootstrap distribution should lay slightly on the left of the theoretical distribution. As shown in Figure 4(a) in contrast to the above expectation, Figure 4(b) demonstrates that the bootstrap distribution matches quite well with the theoretical one, indicating slight non-Gaussian features in the errors.

- **UNGRB**: The theoretical, bootstrap and analytic distributions for UNGRB data have been plotted in Figure 5(b). Although the analytic and bootstrap distributions appear quite different in the figure, the bootstrap and theoretical distributions match quite well. Comparison with Figure 5(a) indicates a weak signature of non-Gaussianity. Δ_χ is nearly 2σ away from the mode of the bootstrap distribution.

4.3 Results for KS Test

We calculate χ_is as defined in Equation (5) for GRB data using the best-fit values presented in Table 1. Further, we generate a set of 79 random numbers following a Gaussian distribution with zero mean and unit standard deviation

which represent the theoretical χ_is. Figure 6(a) shows the histogram of both the associated χ_is. It is clear from the figure that the spread in χ_is from GRB data is smaller compared to the theoretical χ_is. Figure 6(b) shows a similar graph for the UNGRB data comprising 636 data points.

Figure 7 along with Table 4 presents the main result of the KS test. Figure 7(a) shows the CDF of χ_is from GRB and theoretical χ_is, while Figure 7(b) depicts the same in the case of UNGRB. The second, third and fourth columns in Table 4 list values of p, k and cv respectively. Since $p > cv$ in all cases gives $h = 0$, we cannot reject the null hypothesis that the errors have been drawn from a Gaussian distribution. However, Figure 7 shows that the CDF of GRB χ_is has a large deviation from the standard normal compared to the UNGRB counterpart.

5 CONCLUSIONS

In this paper we have presented our results for the GRB/UNGRB data using the statistics discussed in Section 3.1, Section 3.2 and Section 3.3. As discussed earlier, Δ_{χ^2} and Δ_χ statistics are based on extreme value theory. The latter has the advantage over the former since its distribution can be calculated analytically. Both Δ_{χ^2} and Δ_χ quantify the direction dependence in the data. They also qualitatively indicate the presence/absence of non-Gaussian features in the data. The KS test (Sect. 3.3) on the other hand quantitatively measures non-Gaussianity in the data. Our main conclusions for this work are:

- A set of 79 GRBs shows evidence for non-Gaussian behavior of errors, and UNGRB data also show non-Gaussian nature of errors, but weaker compared to GRBs.
- The set of 79 GRBs is different from the peak of the bootstrap distribution by about 2σ, while UNGRB is different from the peak by about 1σ. These results signify a weak but consistent direction dependence in both data sets. This consistency indicates a preferred direction for both data sets.
- The maximum anisotropies from different measurements are presented in Tables 2 and 3. The directions for Union2 and UNGRB are similar for the Δ_χ statistic.
- In all cases, we see that results of UNGRB are very similar to those of Union2. This could be due to the fact that the SNe are greater in number and hence dominate the UNGRB data.

Acknowledgements Meghendra Singh thanks colleagues from DMRC for long-term support, and Shashikant Gupta thanks Saini Tarun Deep for discussion.

References

Amanullah, R., Lidman, C., Rubin, D., et al. 2010, ApJ, 716, 712
Amati, L., Guidorzi, C., Frontera, F., et al. 2008, MNRAS, 391, 577
Cucchiara, A., Cenko, S. B., Bloom, J. S., et al. 2011, ApJ, 743, 154
Firmani, C., Ghisellini, G., Ghirlanda, G., & Avila-Reese, V. 2005, MNRAS, 360, L1
Ghirlanda, G., Ghisellini, G., Lazzati, D., & Firmani, C. 2004, ApJ, 613, L13
Ghirlanda, G., Ghisellini, G., & Firmani, C. 2006, New Journal of Physics, 8, 123
Gupta, S., Saini, T. D., & Laskar, T. 2008, MNRAS, 388, 242
Gupta, S., & Saini, T. D. 2010, MNRAS, 407, 651
Gupta, S., & Singh, M. 2014, MNRAS, 440, 3257
Haan, L. D., & Ferreira, A. 2006, Extreme Value Theory: an Introduction (Springer)
Kodama, Y., Yonetoku, D., Murakami, T., et al. 2008, MNRAS, 391, L1
Li, H., Xia, J.-Q., Liu, J., et al. 2008, ApJ, 680, 92
Liang, E., & Zhang, B. 2005, ApJ, 633, 611
Liang, N., Xiao, W. K., Liu, Y., & Zhang, S. N. 2008, ApJ, 685, 354
Liu, J., & Wei, H. 2015, General Relativity and Gravitation, 47, 141
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2002, Numerical Recipes 3rd Edition: The Art of Scientific Computing (Cambridge University Press; 3 edition)
Salvaterra, R., Della Valle, M., Campana, S., et al. 2009, Nature, 461, 1258
Singh, M., Gupta, S., Pandey, A., & Sharma, S. 2016a, J. Cosmol. Astropart. Phys., 8, 026
Singh, M., Pandey, A., Sharma, A., Gupta, S., & Sharma, S. 2016b, RAA (Research in Astronomy and Astrophysics), 16, 171
Wei, H. 2010, J. Cosmol. Astropart. Phys., 8, 020
Weinberg, S. 2008, Cosmology (Oxford: Oxford University Press)