IMPORTANCE OF THE HIGH ENERGY CHANNEL FOR THE GAMMA-RAY BURST DATA

A. MÉSZÁROS

Dpt. Astron., Charles Univ., Prague 5, Švédská 8, Czech Rep. and Konkoly Obs., Budapest, Box 67, H-1525, Hungary

Z. BAGOLY

Eötvös Univ., Lab. Inform. Technol., Múzeum krt.6-8, H-1088 Budapest, Hungary

L. G. BALÁZS

Konkoly Obs., Budapest, Box 67, H-1525, Hungary

AND

I. HORVÁTH, P. MÉSZÁROS

Dpt. Astron., Penn. State Univ., 525 Davey Lab., University Park, PA 16802, USA

1. Principal Component Analysis

Extensive data bases on Gamma-Ray Burst (GRB) properties such as the BATSE 3B catalog (Meegan et al. 1996) contain a wealth of statistical information. The nine entries of the 3B database for each GRB consist of two durations, T_{50}, T_{90}, which contain 50% and 90% of the burst energy, respectively; four fluences (time-integrated energy fluxes) $F1$, $F2$, $F3$, $F4$, defined over different energy channels; and three measures of the peak flux (each summed over the four energy channels), measured over three different resolution timescales (64 ms, 256 ms and 1024 ms). Thus the initial number of variables is $n = 9$. There is, of course, some incompleteness in the catalog. There are 625 GRBs having all 9 non-zero quantities, and only they are considered here.

The Principal Component Analysis (PCA) is done on these. This statistical procedure is explained in (Jolliffe 1986) and (Kendall & Stuart 1976). It can determine the number m of important quantities $(m < n)$.

2. Results

The results of this procedure are the following. The first PC is roughly given by the sum of all the quantities, with some extra weight on the first three fluences. Because of the different dimensions involved, it has only a formal meaning. The second PC, accounting for 26.5% of the variation, is clearly important, this value being far above Jolliffe's $70/n = 7.8\%$ cut-off level (Jolliffe 1972). This PC is roughly given by the formal difference of the logarithmic peak fluxes and durations. Hence, the duration, peak flux and total fluence are undoubtedly important quantities, but only two of them are independent. This means that in the roughest approximation, it is enough to consider, e.g., a total fluence and a duration, and these two represent 91.6% of the information content in the 3B catalog. The third PC, practically defined by $F4$ alone, accounts for 5.1% of the variation and is already below Jolliffe's level. Hence $m = 2$. Because it is just below Jolliffe's level, this third PC might however be of some importance. The fourth PC, with its 1.5 %, is far below Jolliffe's limit, and is unimportant.

Some of the results are unexpected. For instance, the analysis indicates that an allowable approximation would be to combine (add) the fluences in the first three channels, and consider them in conjunction with the fluence $F4$ in the fourth channel as the basis vectors for the fluence space. This singling out of $F4$ based on its information content appears to be new.

The details of this contribution will be presented elsewhere (Bagoly et al. 1997).

References

Bagoly, Z., et al. (1997) A Principal Component Analysis of the 3B Gamma-Ray Burst Data, *Astrophys. J.*, submitted for publication.
Jolliffe, I. T. (1972) Discarding Variables in a Principal Component Analysis, I: Artificial data, *Appl. Statist.*, **21**, 160-173.
Jolliffe, I. T. (1986) *Principal Component Analysis*. Springer, New York.
Kendall, M. and Stuart, A. (1976) *The Advanced Theory of Statistics*. Griffin, London.
Meegan, C. A., et al. (1996) The Third BATSE 3B Gamma-Ray Catalog, *Astrophys. J. Suppl.*, **106**, 65-110.