LABORATORI NAZIONALI DI FRASCATI
SIS – Pubblicazioni

LNF–97/038 (IR)
6 Ottobre 1997

Algorithms for Correlation Analysis Between Gamma-Bursts and Gravitational Wave Data

G. Modestino, G. Pizzella

Laboratori Nazionali INFN, Frascati
Universita' di Roma "Tor Vergata"

Abstract

The problem of searching for possible correlation between the gamma bursts of still unknown origin and the data recorded with gravitational wave antennas is studied. The accuracy that can be obtained by averaging the gravitational wave data obtained in correspondence with gamma-bursts is evaluated.

1. INTRODUCTION

One of the most important astrophysical phenomena still lacking an explanation, although it is well known to the scientific community since many years, is the occurrence of powerful gamma-ray bursts, lasting several seconds, observed near the Earth with spacecrafts [1]. The intriguing fact is their angular distribution, that is about uniform all over the sky [2]. This suggests that the sources be located very far from Earth, outside the Galaxy, but then the sources must be extremely powerful, and no convincing model has been so far proposed. The idea that the sources could be very far has been recently supported by the observations with the Beppo SAX satellite [3,4, 5], which observed the X-ray counterparts of some gamma bursts, confirming the importance to have experimental observations with different instrumentation. Other data analysis [6] support, instead, the idea that the sources be within our Galaxy.

It is plausible that the phenomena responsible for this gamma emission be due to collapsed objects, perhaps to the coalescence of compact binary systems. If it is so, then the gamma bursts should be associated with the emission of gravitational waves (GW). Possible scenarios have been conceived [7], most of them suggesting GW fluxes below the sensitivity of the presently operating GW detectors. As matter of fact, if the source is assumed to be at a distance of 1 Gpc, the GW burst-flux associated with a total conversion into GW of 1 solar mass has amplitude of the order of $h \approx 3\ 10^{-22}$, whilst the present sensitivity of the best GW antennas is $h \approx 6\ 10^{-19}$ [8] . However, due the complete novelty of this phenomenon we consider certainly worthwhile to explore whether a correlation between the gamma bursts and the data collected with the GW detectors exists.

From the point of view of a GW hunter, to search for correlation with gamma bursts is very attractive for the following two reasons.

The first reason is that when looking for coincidences with another GW detector, one is using two event files, one for each GW detector. Each file consists essentially in events due to noise, either well behaved noise like that due to the brownian motion of the antenna or to the electronic apparatus, or to non stationary noise of unknown origin (seismic, electric.....not always correlated with auxiliary sensors). Thus the coincidences due to GW, if any, are buried in a sea of accidental coincidences due to these various noises. When comparing the GW data obtained with one detector with the gamma bursts one can indeed rely on the fact that at least one file, the gamma file, includes only real signals due to the gamma bursts, as the noise has been completely eliminated by the gamma experimenters. The drawback point is that it is not sure that the gamma bursts be associated with GW emission, but certainly the statistical scenario in this case is much more appealing.

The second very important reason is that the gamma-bursts occur with a rate of about one per day. Thus one can make an investigation on a certain period of time (say, 100 days), and then, if he finds a result, he can repeat the experiment with the next 100 days. This is not possible, in general, when looking for a rare event, like a Supernova, as the GW hunters mostly do with two or more GW antennas.

In this note we describe the algorithm we shall use for correlating the gamma bursts of the BATSE detector with the data of the GW detectors Explorer and Nautilus of the Rome group and we compare it with other possible algorithms.

2. THE GW DATA

The GW raw data obtained with the antennas Explorer and Nautilus of the Rome group have been filtered with a Wiener-Kolmogoroff algorithm [9] for obtaining the signal to noise ratio (SNR) of delta-like signals as large as possible. At the output of the filter we have a filtered signal, consisting in "samples" expressed in kelvin units with a sampling time of 0.2908 seconds. The frequency bandwidth of both Explorer and Nautilus in 1997 is of the order of 1 Hz, which means that the correlation time of the filtered data is of the order of one second. It can be shown that the probability for a sample to have energy equal or greater than E, in presence of well behaved noise originated from Brownian and electronic noises both of gaussian nature, indicating with T_{eff} the effective noise temperature of the GW detector after optimum filtering, is

$$P(>E) = \exp\left(-\frac{E}{T_{eff}}\right) \qquad (1)$$

This shows that T_{eff} is the average value of the sample energy in absence of signals.

This distribution is derived from the more general χ^2 distribution valid when we consider the case of the variable

$$y = \sum_1^n x_i^2 \qquad (2)$$

where x_i has gaussian distribution with zero mean and standard deviation σ:

$$f(y) = \frac{1}{2^{n/2}\sigma^n \Gamma(n/2)} y^{(n-2)/2} e^{-y/2\sigma^2} U(y) \qquad (3)$$

In the case of (1) we have n=2 (the p and q components of the signal), and $T_{eff} = 2\sigma^2$, because the energy E of a sample is given, in absence of a real signal, by the sum of the squares of the two gaussian variables p and q with zero mean and standard deviation σ.

3. THE SUM ALGORITHM

An unexpected result obtained in studying the effect of the noise on the signal is that this effect can be very large [10]. In presence of signals of given amplitude s (representing the response of the detector to a short burst in absence of noise), the corresponding estimate $r^2 = p^2+q^2$ is not equal to s^2 but, due to effect of the noise, is a random variable with a non central χ^2 distribution probability with two degrees of freedom

$$f(r^2, s^2) = \frac{1}{T_{eff}} \exp\left(-\frac{r^2+s^2}{T_{eff}}\right) I_o\left(\frac{2\sqrt{r^2 s^2}}{T_{eff}}\right) \qquad (4)$$

where I_o is the modified Bessel function of zero order.

This distribution is very wide and allows the possibility to have values of the sample energy much different from s^2. In order to give an idea on how wide it can be, we have plotted in fig. 1 calculations made using formula (4) where $SNR=s^2/T_{eff}$.

We go now to make simulations for studying the problem of a possible correlation between the gamma bursts and measurements made with GW detectors. We consider a number M of gamma bursts whose beginning time we indicate with τ_i (i=1, 2,..., M).

Searching through the filtered files of the GW data, we extract M sequences of data, each sequence lasting one hour centered at each τ_i time (± one half hour). Each sequence will include 3600 / Δt samples, where Δt = 0.29082 is the sampling time. Thus there are 12378 samples for each sequence. We assign to each of these samples a time t relative to the corresponding gamma-burst, t = UT of the GW sample - τ_i (also in UT).

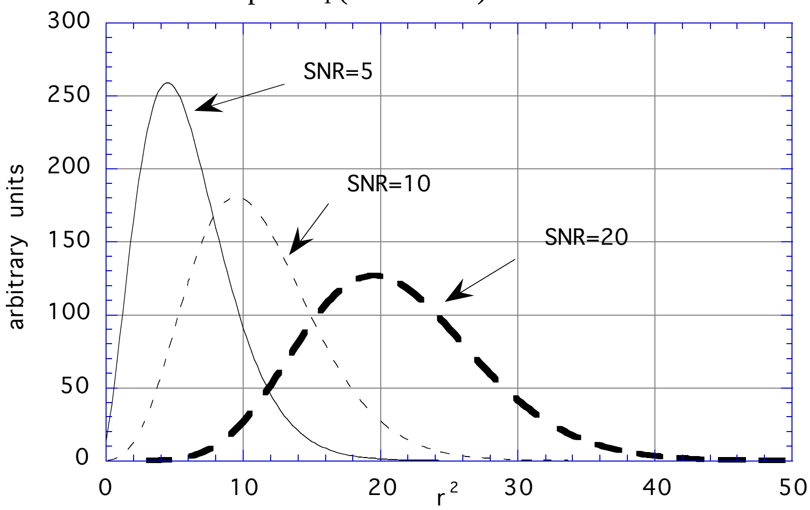

Fig. 1 – Eq. (4) for SNR = 5, 10, 20. We notice that, due to the noise, r^2 covers rather wide intervals in spite of the relatively large SNR.

The obvious algorithm to use is the combination of the M sequences by summing up the data occurring at the same relative time t, for all the values of t (or taking the average of them). The sum of the energies $E(t_k)$ of the samples occurring at the same relative time t_k is given by

$$SUM(t_k) = \sum_{i=1}^{M} E_i(t_k) \qquad (5)$$

As an example we have simulated M=100 gamma-bursts (τ_i , i=1, 2,..., 100) and for each one we have taken 100,000 random and independent data (at times t_j, j=1, ...,100,000). For each one of the M sequences we have simulated the noise samples by considering the energy $E=p^2+q^2$, where p and q have zero mean and gaussian distribution, each one with variance equal to one.

We expect the χ^2 distribution as given by (3), considering that n=2M. This distribution is shown in Fig. 2. The agreement with the expected distribution shows that the simulation is correct.

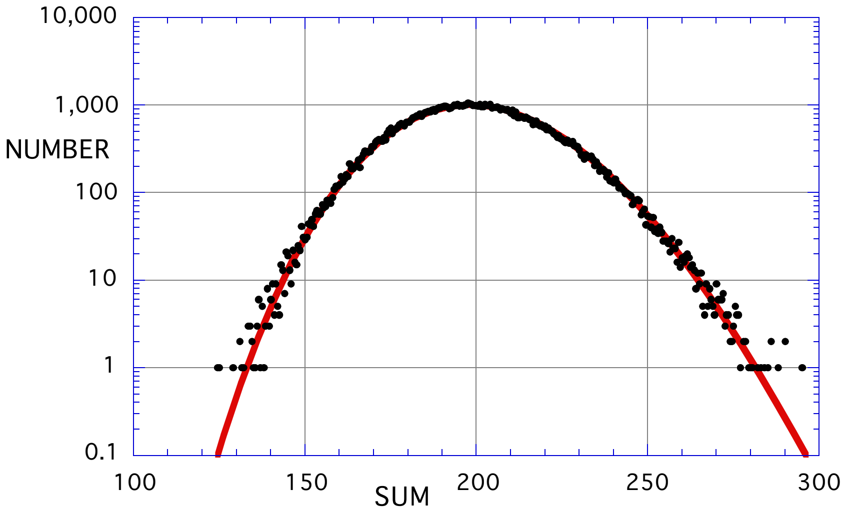

Fig. 2 – The noise distribution for the SUM algorithm. The continuos line is the χ^2 distribution as given by (3), with n=2M=200 and variance equal to 1.

The average value for the SUM is, as expected, equal to 200.

We consider useful to give an example of the procedure we plan to use. For simplicity we have taken five gamma bursts and simulated one hundred GW samples: $E_i(j)$ (i=1,..., 5; j=1,...,100) with noise T_{eff}=2. In addition we have superimposed to the noise at time j=10, for each gamma, a signal s^2=2, thus SNR=1. This is shown in fig. 3. We notice the strong fluctuations due to the noise for each gamma and that the signal at time 10 tends to emerge from the noise on the average.

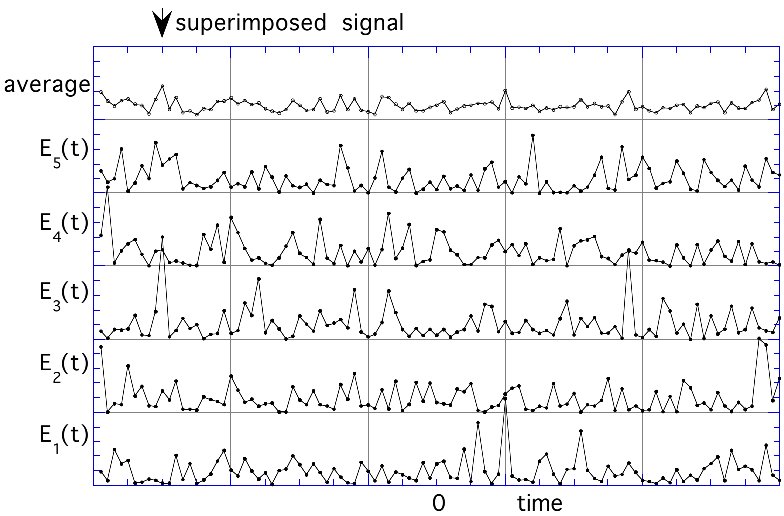

Fig. 3 – An example of the procedure for M=5 gammas. The time is in arbitrary units. The top graph is obtained by averaging, for each time, the GW energy samples relative to the five gammas.

The problem here, however, is not only to study the behavior of the SNR but essentially that of studying the "tail" effects arising from the fact that the noise affects the signals at a great extend as shown in fig. 1. That is, we must answer to the question: how many signals, due to the noise alone, are greater than those due to the real signals plus noise?

For answering to this question we have simulated GW signals and added them to the noise as follows.

For each of the M sequences (each one centered at the time of a simulated gamma burst) we have considered 100,000 independent noise-samples $A(t_j)$ at the times t_j (t_j , $j=1,\ldots,$ 100,000). We have then added to the noise a signal with a given SNR at the time t_1. We have compared the value $A(t_1)$ with the 100,000 values $A(t_j)$ and count how many times (N) $A(t_j)$ was equal or larger than $A(t_1)$. If $A(t_1)$ was the largest one we have N=1. The number N-1 divided by 100,000 is an estimation of the probability that a given value $A(t_j) \geq A(t_1)$ occurred by chance. The result of this simulation, for M=20, 100 and 200, is shown in fig.4, where the number N is plotted versus the SNR.

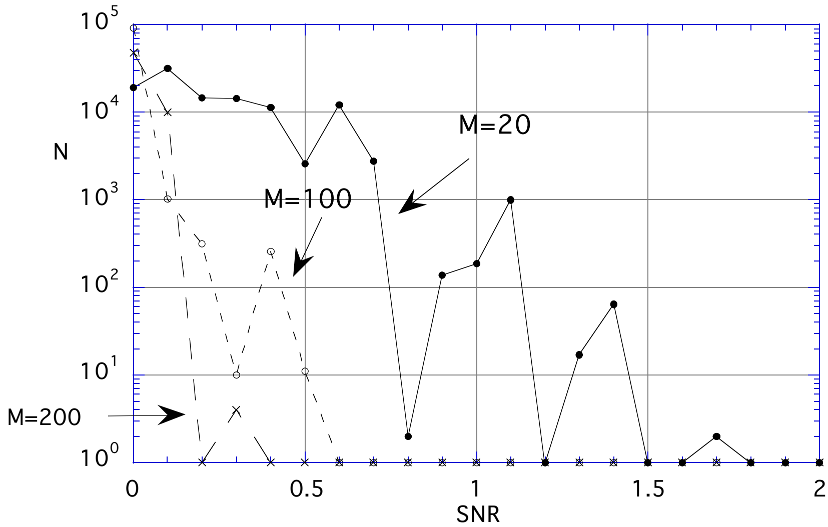

Fig. 4 – SUM algorithm . N-1 is the number of samples, out of 100,000 trials, with the given SNR, generated by the noise. N=1 means that no sample with the given SNR was simulated by the noise, and only the added signal was observed.

We notice that with M=100 gamma samples we can detect signals having a SNR $\approx$ 0.5, with a probability of the order of 10^{-4}–10^{-5} that they might be due to noise.

It is possible to calculate the SNR simulated with a given probability by the noise, as function of the number M of used gamma bursts. This is done making use of Eq. 3 with 2M degrees of freedom . The result of such a calculation is shown in Fig. 5 for probabilities of 10^{-5} and 10^{-3}.

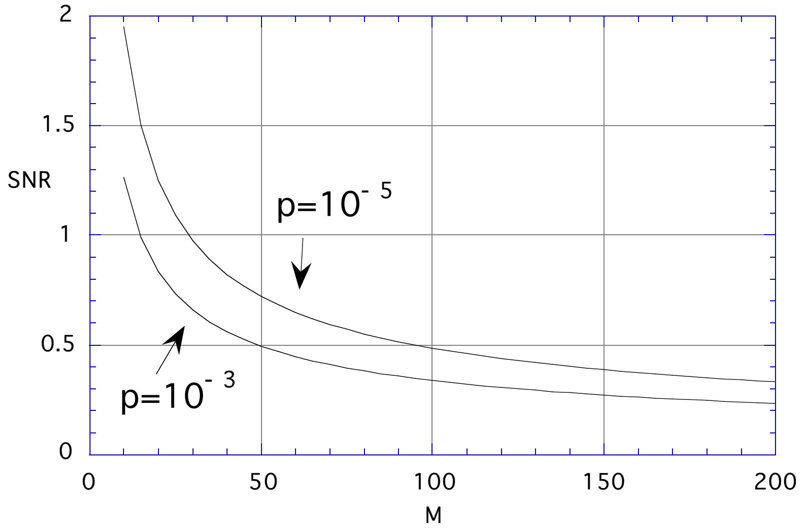

Fig. 5 – The SNR obtained by chance, with probabilities of less than 10^{-5} and 10^{-3}, from a given number M of gamma samples.

From the figure we deduce that, already with M=10 gamma bursts, a signal with SNR near unity can be detected with very good probability that it is not due by chance to the noise.

One way to present the result of the SUM algorithm applied to GW data associated to M gamma bursts could be the following. We calculate the sum over M as defined by (5), then we calculate the χ^2 probability according to (3) that such a value SUM(t_j) could have happened by chance. We plot this probability and the SUM(t_j) as function of the time relative to the gamma burst.

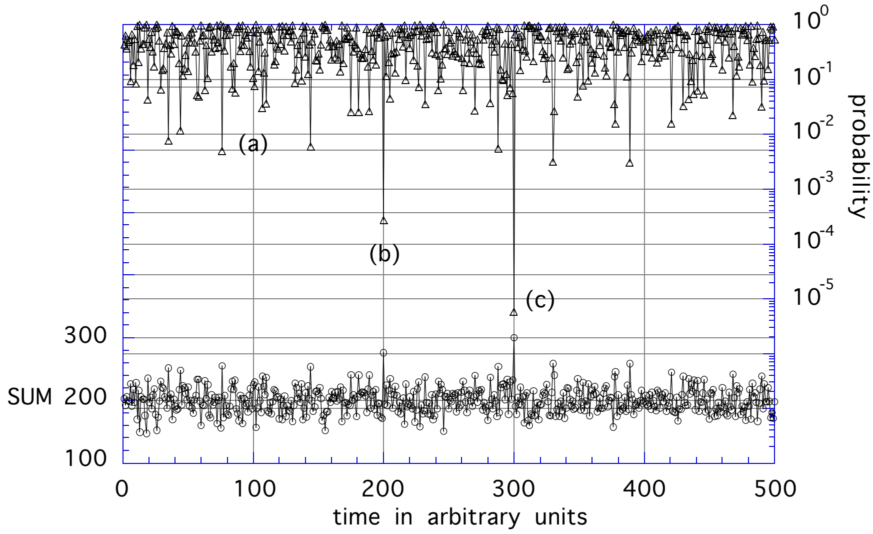

Fig. 6 – The SUM and the relative probability for a peak to have occurred by chance with M=100 gamma bursts as function of time. Signals, with given SNR, were superimposed to the noise at various times: (a) SNR=0.2 at time=100, (b) SNR=0.3 at time=200, (c) SNR=0.4 at time=300.

An illustration of this procedure is given in Fig. 6 for M=100 gamma bursts with 500 samplings, where we have added signals to the noise as explained in the figure caption.

We notice that with M=100 gamma bursts the detection is very good already for signals with SNR=0.3, as seen before in fig. 4.

It must be clear, however, that this probability estimation is based on the hypothesis that the noise be gaussian. A powerful mean to estimate the probability that does not depend on any model is to count how many peaks have amplitude equal or larger to that of the peak we are interested (experimental probability). But, even in this case, the probability considerations apply, for definition of the probability, only to those peaks occurring at times which have been established "a priori". These times can be established by theoretical models, or by previous experimental analysis.

4. THE MINIMUM ALGORITHM

The MINIMUM algorithm has been sometimes used among the GW researchers, because it is the algorithm that reduces the noise most. We consider therefore worthwhile to look into this problem.

The MINIMUM algorithm consists in choosing, for each relative time t, the minimum among the M samples. The idea is the following: suppose that a real GW signal of energy E_s occurs always at a time t=Δt after the beginning of a gamma-burst. Then we expect that the minimum sample energy among the M samples at the time t=Δt have energy of E_s plus (statistically) the noise, while the other samples at the other times, due to noise alone, have energy equal, on the average, to T_{eff} / M. This last statement is valid because the chance to have, due to the noise alone, M simultaneous (in the relative time coordinates) samples, each one with energy equal or grater to E_s, is $\{\exp(-E_s/T_{eff})\}^M = \exp(-E_s /[T_{eff} /M])$. In this way it seems possible to enhance the signal with respect to the noise.

For the MINIMUM algorithm, using simulated data, we get the noise distribution shown in Fig.7.

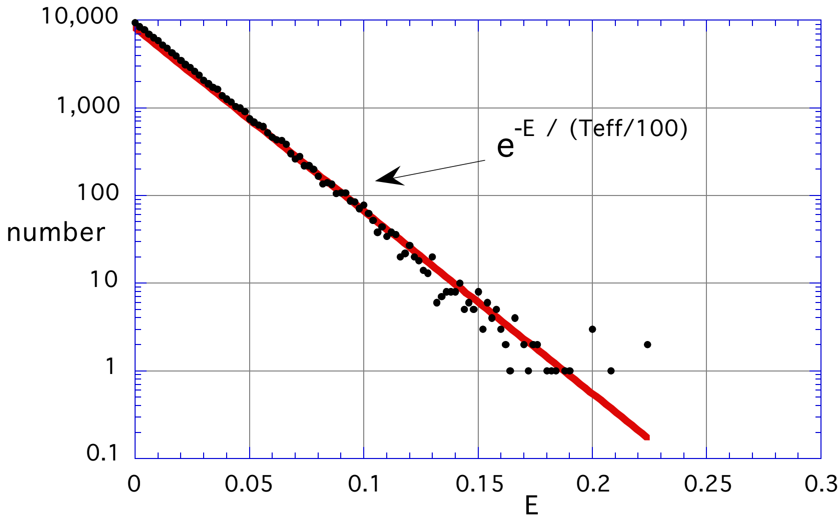

Fig. 7 – The noise distribution for the MINIMUM algorithm with M=100. The average noise is equal to T_{eff} / M, as expected.

The average noise reduces, as expected, to $T_{eff}/100 = 0.02$.

But we have now to estimate how many times the noise alone can simulate a signal with a given SNR when using the MINIMUM algorithm. Applying a procedure similar to that used previously for the SUM, whose result has been shown in fig. 4, we get for the MINIMUM the result shown in fig. 8.

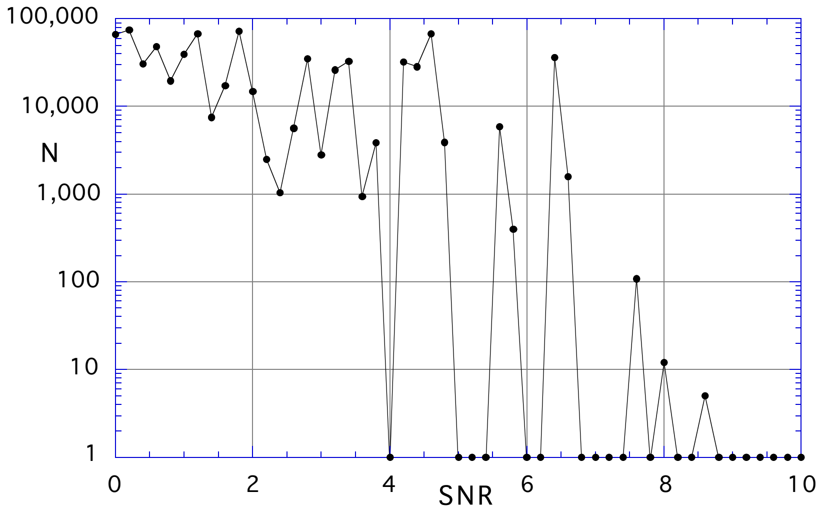

Fig. 8 – The number N for M=100 versus the SNR of the signal for the MINIMUM algorithm (compare with fig.4).

We notice that the MINIMUM algorithm is subjected to large fluctuations. This is due to the strong effect of the noise on the signal, as shown in Fig. 1. As a consequence of this strong effect, which produces large fluctuations, the minimum signal among the M data streams has a great chance to be very small. Therefore, in spite of the large reduction of the noise (T_{eff}/M), the MINIMUM algorithm turns out to be in practice not usable. This conclusion remains valid even if the second or third minimum is considered.

5. FINAL REMARKS

The principal idea at the basis of our search for correlation of GW signals with gamma bursts is, essentially, the consideration that common features in the time behavior of the GW data associated to gamma bursts could exist. Thus the superimposition of several streams of GW data, by summing them up (or taking their average), should enhance this possible features because of the reduction of the noise.

It might be worth to mention that the average algorithm is also often used in many other research fields where, in a situation similar to that of our gamma-GW, a correlation is searched between measurements and excitations produced at given times, like, for instance, in the analysis of bioelectric signals (evoked potentials).

In addition to the previous considerations, the SUM is more convenient also for the very likely cases when, even if a correlation exists, it could show up in a different way for each data stream associated to each gamma burst. Let us suppose that a correlation exists at time t* only for a fraction of the M streams of data. In these cases the MINIMUM algorithm would choose the smallest E value related to a gamma burst when no correlation exists, in this way wiping out a possible effect due to the rest of the gamma bursts. This problem arises also for other algorithms like, for instance, the PRODUCT algorithm which would give a null value by the presence of a null value among the M streams of data.

One reasonable idea is to use of the MEDIAN algorithm in the place of the SUM. This is because the median is less affected by a noisy value than the sum. But the MEDIAN algorithm works out only when the correlation at the same time t* exists for most of the M streams of data. If the correlation exists only for, say, 20% of the streams it is evident that it would not show up when the median is taken, except for a minor effect. Instead the SUM algorithm could reveal a correlation even if it occurs for a fraction of the M streams.

ACKNOWLEDGMENTS

We thank Rinaldo Baldini, Gian Vittorio Pallottino and Francesco Ronga for useful discussions and suggestions.

BIBLIOGRAPHY

1. G.J.Fishman and C.A.Meegan "Gamma ray burst"; Annu. Rev. Astron. Astrophys. 33, 415-458 (1995).
2. Tegmark et al.; The Astrophys. Journal 466, 757-763 (1996).
3. G.Boella et al."Beppo SAX, the wide band mission for X-ray astronomy"; Astron. Astrophys. Suppl. Ser. 122, 299-307 (1997).
4. E.Costa et al. "Discovery of an X-ray afterglow associated with the gamma-burst of 28 February 1997"; Nature, 387,783-785 (1997).
5. Frontera et al.; Astron. Astrophys. Suppl. Ser. 122, 357-369 (1997).
6. B.M.Belli "Evidence for the galactic origin of gamma-ray bursts"; The Astrophys. J. 479, L31-L34 (1997).
7. K.Thorne; In "Recent advances in General Relativity"; eds. A.Janis and J.Porter, Birkhauser, Boston, pag. 196 (1992).
8. G. Pizzella; “Resonant detectors for the search for gravitational waves”. Class. Quantum Grav. 14, 1491-1485 (1997).
9. P.Astone, P.Bonifazi, G.P.Pallottino, G.Pizzella; "Wiener filters for GW antennas" Il Nuovo Cimento, 17C, pag. 713 (1994).
10. A.Astone, G.P.Pallottino, G.Pizzella; "On the efficiency of the search of coincidences in gravitational wave experiments"; In press in the J. of General Relativity and Gravitation (1997).