

## Search for precursor activity in *Swift*/BAT gamma-ray bursts.

---

**G. Calderone<sup>a</sup>, P. Romano<sup>b</sup>, V. Mangano<sup>b</sup>, and G. Cusumano<sup>\*b</sup>**

<sup>a</sup>*Dipartimento di Scienze Fisiche ed Astronomiche, Università degli Studi di Palermo  
via Archirafi 33, 90123 Palermo, Italy*

<sup>b</sup>*INAF-IASF*

*via Ugo La Malfa 153, 90146 Palermo, Italy*

*E-mail: gcalderone@ifc.inaf.it, romano@ifc.inaf.it,*

*vanessa@ifc.inaf.it, giancarlo.cusumano@ifc.inaf.it*

Some GRBs are characterized by emission episodes occurring before the main event, called precursors. Their emission intensity is usually weaker than the GRB and is often comparable to (or slightly greater than) the background noise, making their detection difficult. We developed a code to automatically detect such precursors and applied it to a sample of 280 *Swift*/BAT GRBs.

The identification procedure of precursor candidates uses a detection algorithm based on the wavelet transforms, and relies on an accurate study of the BAT background noise properties. A rigorous statistical approach is used for detection threshold settings and detection significance evaluation. We show the preliminary results of the analysis.

*7th INTEGRAL Workshop  
September 8-11 2008  
Copenhagen, Denmark*

---

\*Speaker.

## 1. Introduction

A precursor is an emission event occurring before the GRB main prompt emission, whose origin is still under debate. Models explaining precursors are discussed in [6], [7], [8]. The search and study of these events is important to constrain the physics and model parameters.

Precursors are expected to be weaker and softer than the main event, so they will likely be misinterpreted as noise in a gamma-ray instrument such as *Swift*/BAT. We developed a code to automatically detect such weak events in noisy signals using wavelet transforms.

## 2. The wavelet transform

Wavelet transforms, like Fourier transforms, are mathematical operations to be applied on functions to obtain a different representation of them, in which certain features are more easily recognizable. The wavelet transform is performed computing the convolution between the original function and an appropriately scaled and translated localized function (called analyzing function or *mother wavelet*), so the localization information in the original signal is preserved. The new domain comprises not only the “scale factors” (as in the Fourier transform) but also the offsets used to translate the mother wavelet, so the wavelet domain is always larger than the original domain<sup>1</sup>. In our case the signal to be analyzed is a light curve,  $S(t)$ , a one-dimensional data series. The corresponding wavelet transform will be a two-dimensional data series indicated as  $W(a, b)$ , where  $a$  identifies the scaling factor and  $b$  the translation offset. We will often refer to this domain as the “time-scale plane”. The mother wavelet we chose is the so called “Mexican hat” (actually the second derivative of a Gaussian<sup>2</sup>) because it is particularly well suited in revealing Gaussian-shaped peaks in noisy signal and it has already been used in search for GRB precursors in the BATSE archive (see [4]).

Note that no information is lost when performing wavelet transforms as it is always possible to perform the inverse transformation and reconstruct the original signal.

### 2.1 The wavelet transform of a Gaussian noise

It is possible to show (see [1], [5]) that the ratio of the wavelet power to the squared standard deviation (or normalized wavelet power) of a Gaussian noise is, at each point of the time-scale plane, a random variable distributed according to a  $\chi^2$  distribution with one degree of freedom (when the mother wavelet is the Mexican hat):

$$\frac{|W(a, b)|^2}{\sigma^2} \rightarrow \chi_1^2 \quad (2.1)$$

where the symbol  $\rightarrow$  means “is distributed as”. This let us conduct a significance test on the GRB light curve with the null hypothesis that the corresponding normalized wavelet power value is only caused by noise. If the normalized wavelet power is less than  $\chi_1^2(1 - \alpha)$ , with  $\alpha$  being a prescribed significance level, we accept the null hypothesis that the signal is just noise, otherwise we reject the null hypothesis and conclude that some real physical signal is present in the data.

<sup>1</sup>If the original domain has more than one dimension it is also possible to consider rotations of the mother wavelet (see [2]).

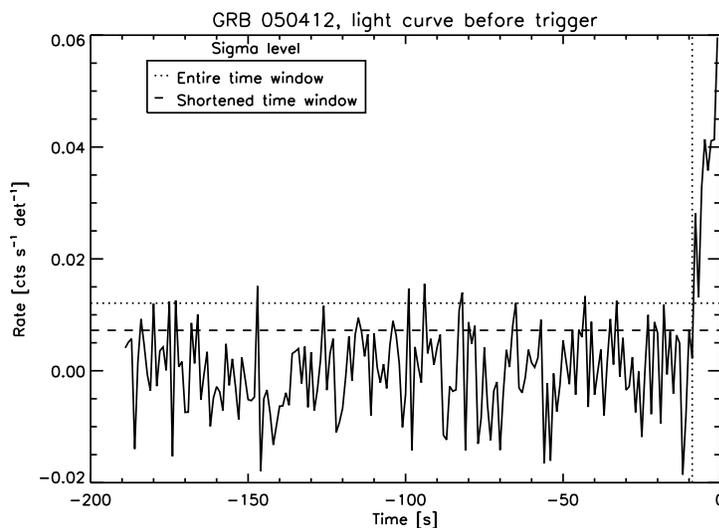
<sup>2</sup>This family of mother wavelets is usually indicated as derivative of Gaussian (DOG) with a parameter representing the order of derivative.

### 3. The *Swift*/BAT data

The search for precursor activity was conducted on a set of 280 light curves (dates ranging from 2004-12-17 to 2008-04-13) from the *Swift*/BAT archive. All data were processed through the standard *Swift*/BAT pipeline<sup>3</sup> for data reduction: production of quality maps, mask weighting and light curve extraction. All light curves were produced with a 1 s time binning.

#### 3.1 The background noise

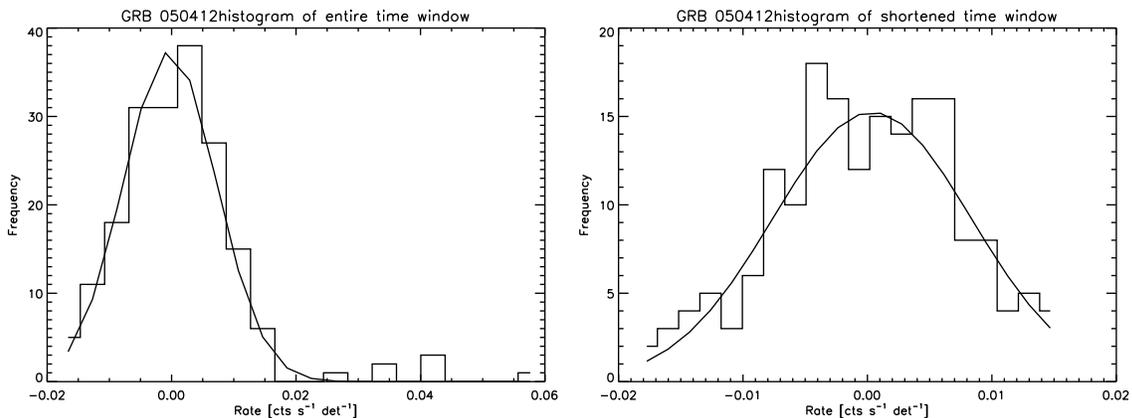
The part of the light curves before the trigger were used to study the BAT background noise. We found that it can be effectively modeled with a Gaussian noise. In particular we started by producing histograms of the part of light curves lying between the beginning of data and the trigger (“time window”), and fit them with a Gaussian function. The error associated with each bin count in the histogram is just the square root of counts (we assume a Poissonian statistic). While the fit was satisfying in many cases, we noticed that the trigger often occurred after the count rate had grown significantly over the noise level, and this could result in an overestimation of the standard deviation in the fit. So we shortened the time window until all points were below a cutoff value calculated so that the probability for a Gaussian variable to have values greater than the cutoff value is  $0.05/N_{\text{point}}$ . Then we ran again the fit to obtain new estimates for the standard deviation. Fig. 1 shows the light curve of GRB 050412 before the trigger (occurring at  $t = 0$ ) and the noise



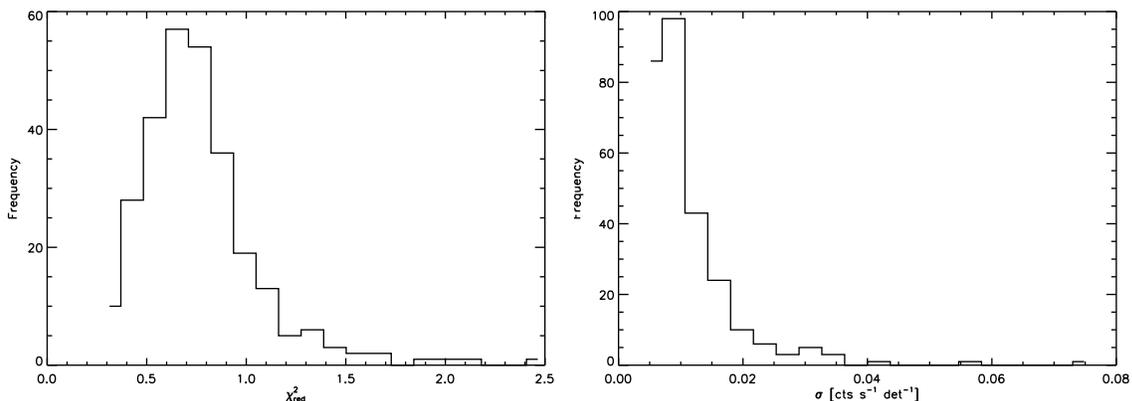
**Figure 1:** Noise levels in the light curve of GRB 050412. The vertical line indicates the end of the shortened time window.

levels calculated using the entire time window and the shortened one. Fig. 2 shows the resulting histograms in the two cases. Finally, in Fig. 3 (left) we report the distribution of  $\chi_{\text{red}}^2$ , the reduced chi-squared, obtained from the fit over all light curves. In more than 80% of cases we have  $\chi_{\text{red}}^2 < 1$ , so we conclude that the *Swift*/BAT background noise is a Gaussian noise. This is a necessary condition to apply the significance test described in Sect. 2.1. Fig. 3 (right) shows the distribution

<sup>3</sup>[http://swift.gsfc.nasa.gov/docs/swift/analysis/threads/bat\\_threads.html](http://swift.gsfc.nasa.gov/docs/swift/analysis/threads/bat_threads.html)



**Figure 2:** Histogram of light curve using the entire time window (left) and the shortened one (right).

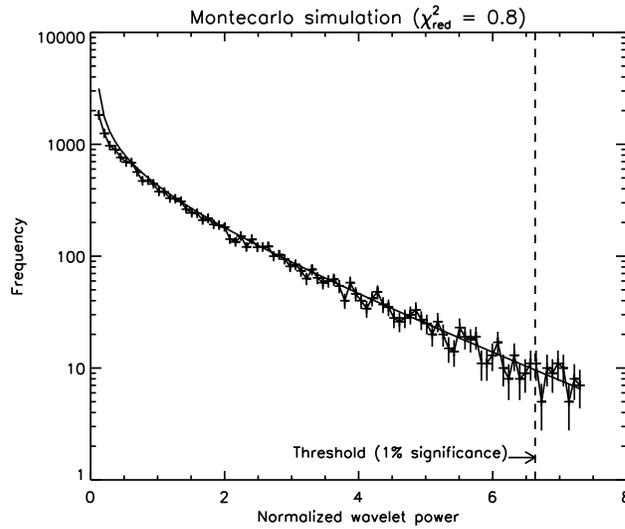


**Figure 3:** Histogram of  $\chi^2_{\text{red}}$  (left) and of standard deviation ( $\sigma$ , right) obtained by fitting the noise in light curves against a Gaussian distribution.

of standard deviation ( $\sigma$ ). The best fit parameters for each light curve were stored on a database for use in subsequent analysis.

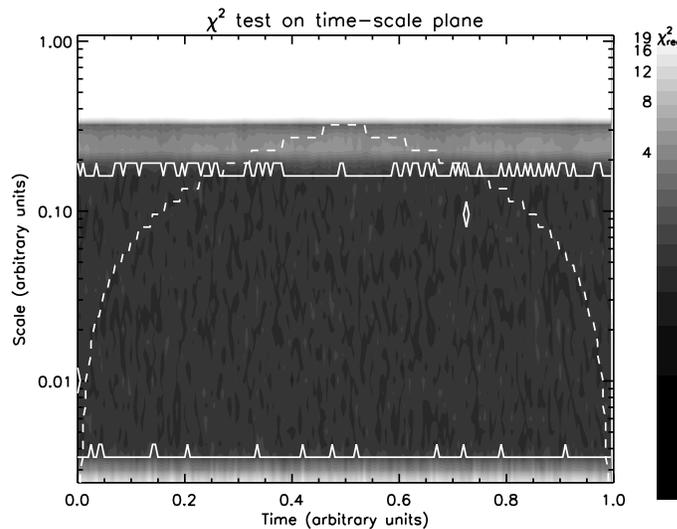
#### 4. Monte Carlo simulations

We performed Monte Carlo simulations to test the validity of Eq. 2.1, in particular if there are some border effects, and if it applies to the entire time-scale plane. We simulated  $2 \times 10^4$  light curves of Gaussian noise with parameters similar to the ones obtained by fitting the original light curves. For each simulated light curve we computed the normalized wavelet power and compared it to the expected  $\chi^2$  distribution with one degree of freedom. Fig. 4 shows the distribution of normalized wavelet power for a fixed point of the time-scale plane, which closely follows the expected distribution. The vertical line indicates the 1% confidence level threshold; some points fall beyond this line because of the very high number of trials. We also performed a  $\chi^2$  test to check if the distribution of normalized wavelet power follows the expected distribution and we found that this is not true in the entire time-scale plane. In Fig. 5 we show a contour plot of the time-scale



**Figure 4:** Distribution of normalized wavelet power for a fixed point of the time-scale plane. The vertical line indicates the 1% confidence level threshold.

plane and the corresponding reduced  $\chi^2$  (as a gray scale gradient) resulting from the comparison of the simulated data with the expected distribution; the solid line identifies the region inside which  $\chi^2_{red} < 1.6$ , while the dashed line identifies the region outside the cone of influence (COI) of the borders (see [3]). Our simulations show that Eq. 2.1 holds only inside the region identified by

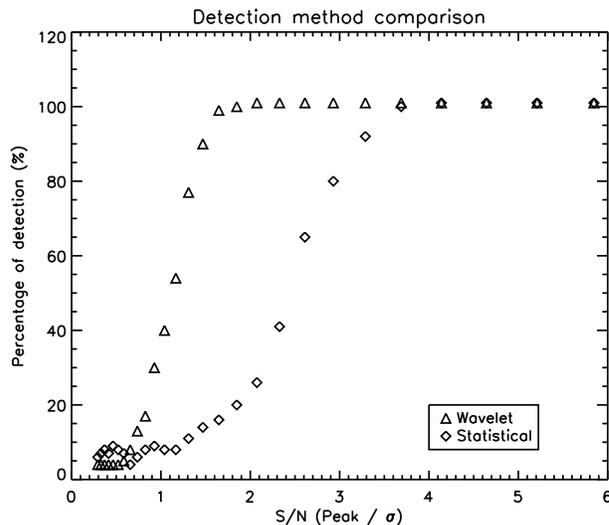


**Figure 5:** Contour plot of the time-scale plane and reduced  $\chi^2$  (as a gray scale gradient) resulting from the comparison of simulated data with expected distribution.

the solid line. This is because at larger scales the wavelet extension becomes comparable with the length of time series itself, so in the evaluation of the wavelet transform the signal  $S(t)$  is seen as an approximately constant term and the result is no longer a random variable. At each scale we can divide the time domain into two parts: inside and outside the COI (respectively outside and inside

the dashed line). We conclude that when dealing with finite length time series (as is always the case in any real experiment) Eq. 2.1 holds only at those scales for which the part outside the COI is longer than the part inside COI.

Finally we performed a simulation about the sensitivity of our feature detection method (Sect. 2.1), and compared it to a more classical statistically-based approach. The simulated data comprise 25 sets of  $10^3$  time series with Gaussian noise and a Gaussian peak (the feature to be detected). Each set has a different signal/noise ratio [ $S/N = \text{MAX}(\text{peak})/\text{noise}(\sigma)$ ]. The statistical approach is based on the presence of significant positive excess in the noise Gaussian distribution. The result of our comparison is shown in Fig. 6. Our feature detection method reaches 100% of detection at a



**Figure 6:** Comparison of our feature detection method with a statistically-based approach. Our method is much more sensitive.

$S/N \approx 1.6$ , while the statistically-based approach reaches the same level at  $S/N \approx 3$ . Therefore our method is much more sensitive than a statistically-based method.

## 5. The precursor detection algorithm

Precursors are emission events separated from the main events. However, a more precise definition of precursor does not exist yet, and different authors use different definitions (see for example [4], [9]). According to our definition, a precursor is an event that:

- occurs before the main prompt emission;
- has a maximum peak weaker than the main event peak;
- the time-integrated flux is less than the integrated flux of the main event;
- is separated from the main peak by at least 2 seconds of noise;

With the first condition we allow the candidate precursor to trigger the instrument. The last condition is instrument based, as it comes from the fact that we chose a 1 s binning time, and we could

not assess whether a peak is “well separated” from the main event if the separation is less than two seconds.

The precursor detection algorithm follows these steps:

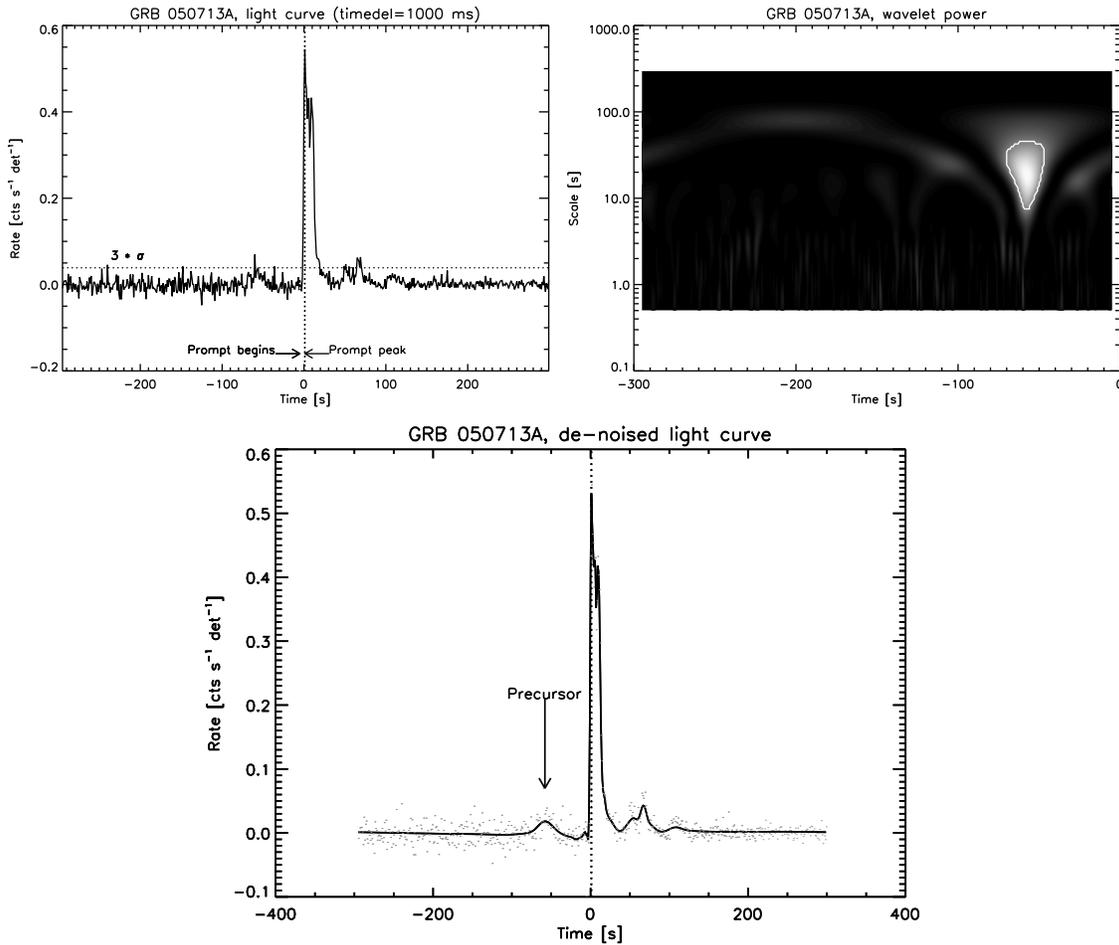
- identification of the part of light curve before the main prompt:  
we search for the main event peak as the highest count rate in the entire time series. Then we move back in time until we identify a 2s-long interval with all points being below a noise threshold, calculated as the level at which a Gaussian noise with a given standard deviation (see Sect. 3.1) has only 1% probability of being higher. The part of light curve from the beginning to this interval is used for wavelet analysis;
- wavelet transform;
- test for significance of normalized wavelet power:  
we apply the significance test described in Sect. 2.1 with a confidence level of 1%. This means that if a significant peak is found, we have a probability of 1% that it is due to a statistical fluctuation;
- candidate precursor confirmation:  
in case a significative signal has been found, the criteria described above will be checked for validation. In case all tests are passed the precursor is confirmed;
- light curve reconstruction:  
if the precursor is confirmed, we perform the wavelet transform on the entire light curve, apply the test again and identify the regions in the time-scale plane for which the significance test is passed, then we extend this region to comprise its COI. Finally, we perform the inverse transform using only the wavelet coefficients inside the identified regions, this way we obtain a de-noised light curve which allows us to easily identify the temporal interval during which precursor emission takes place.

## 5.1 An example

As an example we report the detailed analysis on GRB 050713A. The light curve before the prompt shows only some isolated points exceeding the  $3\sigma$  limit (see Fig. 7, upper-left), however our algorithm found only one considerable region of the time-scale plane above threshold (Fig. 7, upper-right). Wavelet coefficients below threshold can be discarded and the transform can be inverted to reconstruct a “de-noised” light curve (Fig. 7, bottom).

## 6. Results

Our detection algorithm found 113 candidate precursors (over a sample of 280 GRBs), out of which only 31 satisfy the tests described in Sect. 5. Three more GRBs (060210, 061019, 071010B) did not pass the tests; since we think that they should be taken as precursors we added them manually. So a total of 34 precursors have been detected (12%). We detected all precursors cited by [9] except 050401, 060729, 070411 since they do not satisfy our precursor definition. We



**Figure 7:** Upper-left: light curve of GRB 050713A. Upper-right: wavelet transform of the light curve, the white contour indicates the region of points above threshold. Lower: “de-noised” light curve.

excluded 060124 since only the data collected during the precursor phase are in event mode, while the remainder of the light curve is in survey mode.

## References

- [1] Calderone, G. et al., 2009, in preparation
- [2] Farge, M., 1992, *Annual Review of Fluid Mechanics*, **24**, 395
- [3] Torrence, C., Compo, G. P., 1998, *Bulletin of the American Meteorological Society*, **79**, 61
- [4] Lazzati, D., 2005, *MNRAS*, **357**, 722
- [5] Ge, Z., 2007, *Annales Geophysicae*, **25**, 2259
- [6] Li, Li-Xin, 2007, *MNRAS*, **380**, 621
- [7] Ramirez-Ruiz, E., MacFadyen, A.I., Lazzati, D., 2002, *MNRAS*, **331**, 197
- [8] Wang, X., Meszaros, P., 2007, *ApJ*, **670**, 1247
- [9] Burlon, D. et al. 2008, *ApJ*, **685**, L19