UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura

www.uniud.it

Corso di dottorato di ricerca in:

Fisica

Ciclo 34°

Two studies concerning the development of the HERMES-TP/SP space mission

Dottorando
Giuseppe Dilillo

Supervisore
Prof. Andrea Vacchi

Co-supervisore
Dott. Fabrizio Fiore

2022

Università degli Studi di Udine

DIPARTIMENTO DI DI SCIENZE MATEMATICHE, INFORMATICHE E FISICHE.
Corso di Dottorato in Fisica

Two studies concerning the development of the HERMES-TP/SP space mission.

Candidato:
Giuseppe Dilillo

Supervisore:
Prof. Andrea Vacchi

Co-supervisore:
Dr. Fabrizio Fiore

Anno Accademico 2021-2022

Dark star crashes, pouring its light into ashes
Reason tatters, the forces tear loose from the axis
Searchlight casting for faults in the clouds of delusion
Shall we go, you and I while we can?

Introduction

Executive Summary

The goal of the HERMES project is to develop a constellation of nanosatellites to study astronomical high-energy transients, such as gamma-ray bursts. The first six HERMES spacecrafts are expected to be launched in orbit during 2023. These units will be deployed in the framework of HERMES-TP/SP, a pathfinder mission intended to demonstrate the feasibility of the HERMES approach to gamma-ray burst observations. In this thesis, two works concerning the hardware and software development of the HERMES-TP/SP payload are presented.

The first of these works is a qualification study of GAGG:Ce scintillation crystals. The HERMES enabling technology is a detector with a “siswich” architecture. In a siswich detector a silicon detector plays the double role of an active X-ray sensor and a passive gamma-ray sensor. GAGG:Ce ($Gd_3Al_2Ga_3O_{12}$:Ce, Cerium-doped Gadolinium Aluminium Gallium Garnet) is a novel inorganic scintillator with features well-matching the requirements of the HERMES payload. GAGG:Ce has been known for its unusually intense and long-lasting afterglow emission, a slow phosphorescence component in scintillation light [122]. Afterglow emission can be induced by the interaction between the scintillator and the energetic trapped particles of the near-Earth radiation environment. It was unclear if afterglow emission could reach levels high enough to compromise the functionality of the HERMES detector electronics. To investigate this point, a proton irradiation campaign of a GAGG:Ce sample has been designed and conducted. These efforts resulted in the introduction of a novel model of GAGG:Ce afterglow. This model was applied to the HERMES-TP/SP orbital scenario, aiming at an upper-bound estimate of the detector performance degradation. We concluded that GAGG:Ce afterglow emission should not endanger the operations of the first HERMES units.

The second work concerns the design, implementation and test of algorithms for detecting gamma-ray bursts. At a fundamental level, algorithmic techniques for detecting gamma-ray bursts have gone largely unchanged through different generations of spacecrafts and monitor experiments. High-energy photons reaching a detector are counted (binned) over a fundamental acquisition time interval and count rates are compared against a background estimate over multiple pre-defined timescales. To minimize the chance of missing a burst due to a mismatch between the event duration and the length of the tested timescales, many different timescales are generally tested, in some case up to as many as allowed by the on-board computer [42].

With the enhancement of on-board computers and CPUs, 'trigger algorithms' have grown to support an increasing number of criteria and parameters to achieve better sensibilities, the most critical of which usually is the timescale of the sample being analyzed [79]. For example, while BATSE allowed for the simultaneous operations of three different logics, more than 120 and 800 different trigger criteria can be specified for Fermi GBM and Swift BAT respectively [81] [79]. This situation is unfavourable to nanosatellites whose available computational resources are penalized by the miniaturized nature of the on-board computer.

We investigated various approaches to the trigger algorithm problem. In particular, we inquired the applicability of a modern and efficient changepoint detection technique to serve as the foundation of a novel trigger algorithm. This technique is named FOCuS (Functional Online CUSUM) and was pioneered by researchers of the University of Lancaster Department of Statistics, with whom we collaborated. Exact implementations of FOCuS enable the design of a trigger algorithm with sensibility matching that of a routine checking over all possible timescale parameters, with per-iteration computational costs growing on average as the logarithm of the number of observations [98][117]. The usage of similar changepoint detection techniques (with linear per-iteration computational costs) has been advocated in the past [103] but never thoroughly investigated. Different trigger algorithm logics have been tested on synthetic and real data, using both conventional and unconventional background estimate techniques. In our tests, we found that an approximated implementation of FOCuS provides better sensibilities and computational performances than a benchmark algorithm modelled after that of Fermi-GBM.

Acknowledgements

I thank my supervisor and co-supervisor, Andrea Vacchi and Fabrizio Fiore, for the support, the inspiration and the tremendous encouragement they provided me during the last four years.

The first study presented in this thesis would have not been possible without the work of Nicola Zampa and Giovanni Pauletta. Nicola's contributions to that work are almost too many to list. Still, I wish to remark on the formal development of the GAGG:Ce afterglow emission model presented through Sec. 3.2, and on the arduous fitting of the experimental data discussed in Sec. 3.4. Beside his scientific contributions and mentoring, I wish to acknowledge Nicola's kindness and patience. Giovanni Pauletta led the development of the experimental apparatus and the operations of the irradiation campaign. These efforts are presently covered in Sec. 3.1.

The mathematics described in the sections on the subjects of CUSUM (Sec. 4.1.7 and Sec. 4.1.4) and FOCuS (Sec. 4.1.5, Sec. 4.1.8 and Sec. 4.1.9) are based on the research work of Kester Ward, under the supervision of Paul Fearnhead and Idris Eckley. Kester did also provide for the first Python implementations of FOCuS and for the idea of using exponential smoothing techniques for background assessment. Paul Fearnhead and Idris Eckley set a common discussion ground for our collaboration. That discussion provided the foundation for Sec. 4.1.4, in which CUSUM is introduced. Riccardo Crupi solved the n-dimensional case of the false detection problem sketched in Sec. 4.1.3 and helped me parallelizing the simulations described in Sec. 4.3.

Fundamental to both studies were the work and the mentoring of Riccardo Campana as well as the precious comments of the reviewers, Elisabetta Bissaldi and Ehud Behar.

The more I write, the more I realize how difficult it is to truly acknowledge the contributions of all the people which accompanied me during these years. So, please pardon my brevity, and thank you! Giovanni Della Casa, Jakub Ripa, Gabor Galgoczy, Daniela Cirrincione, Marco Baruzzo, Marco Citossi, Diego Cauz, Alberto Sabatini, Yuri Evangelista, Fabio Fuschino, Claudio Labanti, Alejandro Guzmann, Moritz Klawin.. Thank you! Finally, one special thank to my family and to my beautiful black and white cat for keeping me almost sane during these weird times.

Thesis Outline

This thesis is divided in two parts. The first part comprises two chapters providing an introduction on the topics of gamma-ray bursts and the HERMES space missions. Compelling to the rest of this thesis are the overview of the literature on the subject of algorithms for GRB detection and the discussion of the scientific background expected for HERMES presented in Sec. 4.1.1 and Sec. 2.5, respectively. In the second part of this thesis, two novel works are presented. The first of these studies is covered in the third chapter and concerns a qualification study of the scintillation crystals selected for the HERMES detector. The second study is presented in the fourth and last chapter of this thesis, and concerns an inquiry of different algorithmic strategies for detecting gamma-ray bursts.

Contents

Part I

State of the art

Chapter 1

Gamma-ray bursts

1.1 A timeline of Gamma-ray bursts observations

The VELA spacecrafts On August 5 1963, after a decade of negotiation, the governments of USA, Soviet Union and UK agreed to sign the “Partial Nuclear Test Ban Treaty”. The treaty provided for the immediate interruption of nuclear weapon tests in the atmosphere, underwater, or outer space. In the same year, the US Air Force deployed to orbit the first Vela satellite. The goal of this satellite, and of the many Vela that followed, was to grant the adherence of other nations, in particular the soviets, to the ban. These spacecrafts orbited Earth well above the Van Allen Radiation Belt and were equipped with X-ray, gamma-ray and neutron detectors to readily detect nuclear blasts.

Having different satellites simultaneously operating made possible for localization of possible nuclear tests: since the light travels at finite speed, time difference between the arrival of a signal to detectors hosted by different spacecrafts would result in a crude estimate of the angle of radiation incidence with respect to the satellite position, therefore the direction of the nuclear explosion.

On 2 July 1967 the satellites Vela 3 and Vela 4 detected a flash of gamma radiation. It lasted for around 2 seconds, six orders of magnitude above the expected duration of gamma flashes resulting from nuclear explosion. The origin was unknown. Supernovae and solar flares both could be responsible for such an event but neither were observed on that day.

The team researching the phenomena was based at Los Alamos Scientific Laboratory and led by Ray Klebesadel. Confirmed the absence of any particular military risk, further studies about the mysterious flashes were delayed to later investigation. In 1972, the same Los Alamos National Laboratory team reanalyzed the data from the previous decade and determined that in fact a number of gamma flashes were observed. Moreover, all these events seemed to be originating well beyond our cosmic neighborhood, since their distribution across the sky soon made clear how those were not related to the Sun, the moon or any other body in our solar system.

In 1973 the discovery of gamma-ray bursts (GRBs) was announced to the public. In "*Observations of gamma-ray bursts of cosmic origin*", Ray Klebesadel, Ian Strong and Ray Olson claimed the discovery of sixteen short bursts of photons in the energy range of 0.2-1.5 MeV coming from sources different from the Earth and the Sun [65].

The announcement of GRBs discovery was met with enthusiasm by the astrophysics community. No other known phenomena reached the GRB observed flux between the hard X and gamma band but no optical counterpart was identified. This immediately steered the attention towards compact, extreme cosmic objects such as black holes or neutron stars. However, the large flux was calling for the sources to be located nearby. If the Solar System had to be ruled out, was it possible that black holes were crowding the disk of our own galaxy? Similar questions surrounded the subject. As a matter of fact, a multitude of theories were proposed in order to explain the origin of GRBs through the 1970's and the 1980's, neither of which providing solid arguments.

It was soon clear that in order to take a leap in the comprehension of GRBs a better understanding of the location of such bursts was needed. To achieve this goal, high-energy detectors were hosted as secondary payload in various interplanetary spacecrafts; the idea being that the localization accuracy could be improved by enlarging the baseline distance between the detectors. The first Inter-Planetary Network (IPN) was completed by the end of 1978. The IPN included 5 new space probes orbiting Earth, Sun and Venus: soviet Prognoz 7, german Helios 2 and the NASA's Pioneer Venus Orbiter, Venera 11 and Venera 12. By 1987, IPN identified more than 200 GRBs. However, the large uncertainties and the difficulties associated to combine and inter-calibrate different instruments hindered the breakthrough scientists were hoping for.

The BATSE experiment. At the very beginnning of the 1990's most theories about GRBs implied that the mysterious bursts were originating from possibly exotic sources inside our galaxy.

In April 1991, as part of NASA's "Great Observatories" program, the 17 tons Compton Gamma Ray Observatory (CGRO) was launched from space shuttle Atlantis. Compton was a large space observatory dedicated to photon detection between 20 keV and 30 GeV. Part of the CGRO payload was the instrument

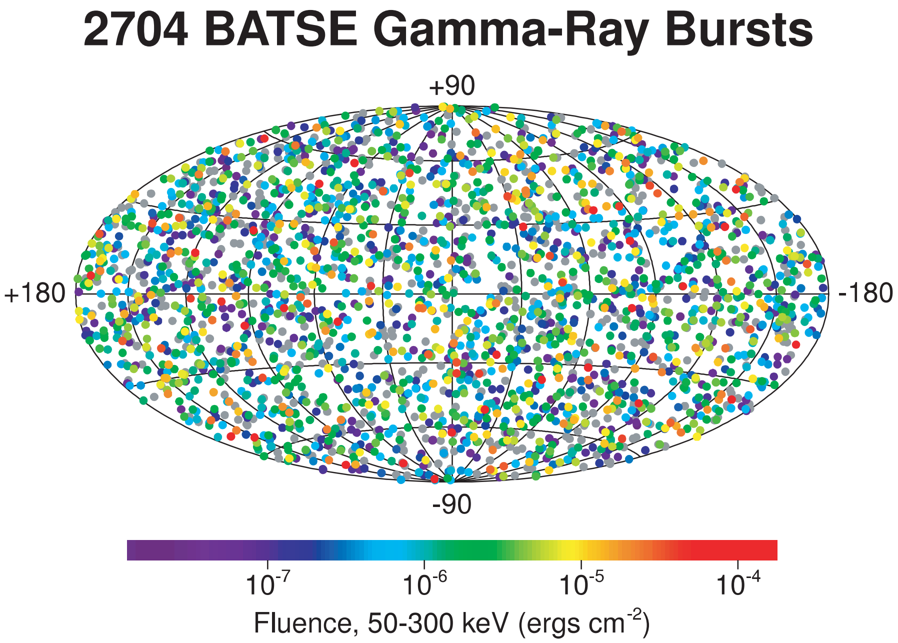


Figure 1.1: Locations of 2704 gamma-ray bursts detected by the BATSE instrument during nine years of observations. Sources are isotropically distributed. Picture from [20].

Burst and Transient Source Experiment (BATSE). Due to unprecedented sensitivity - BATSE was almost 10 times more sensitive than previous IPN detectors - it was able to detect GRBs at a large rate, roughly around one per day. BATSE activity lasted until the end of the mission and CGRO de-orbit in 2000. BATSE allowed to gather a large sample of gamma-ray bursts. Of the almost 3000 GRBs it discovered, no two bursts looked the same. The greatest part of GRBs were dim, while a few were briefly the brightest and most energetic objects of the high-energy sky. Some lasted a few milliseconds, while others shined on for hours. Most importantly, gamma-ray bursts appeared to be randomly distributed across the sky. Figure 1.1 shows the locations in galactic coordinates of 2704 gamma-ray bursts detected by the BATSE instrument during nine years of observations. In 1992 it was already clear that GRBs were isotropically distributed and did not track any known population of galactic objects [80]. Further observations and tests confirmed that the bursts locations were indeed isotropically distributed across the celestial sphere, showing no significant evidences of a dipole or a quadrupole moment [30].

The fact that GRBs were isotropically distributed was already known to scientists. The IPN measures were obviously in agreement with that. However, it was expected that a more sensitive instrument such as BATSE would reveal an excess of faint GRBs towards the galactic centre, where stars are more densely

distributed, in agreement with the idea that GRBs originated inside the Milky Way. BATSE showed instead that even the faintest burst were uniformly distributed across space. This was the breakthrough scientist were waiting for.
At this point, the most likely scenarios were two: 1. GRB sources were located at galactic distances in a massive halo surrounding our galaxy. Even in this scenario a small deviation from isotropy was expected but, since the BATSE catalog was still modestly sized, the galactic hypothesis remained in agreement with the observation. 2. GRB sources were at cosmological distances, in which case the observed distribution was explained by the Cosmological Principle. With BATSE, the long-reigning model of 'local' GRB sources started to crumble while the cosmological theory - a fringe hypothesis for all the 1970's and 1980's - gained momentum.
By looking at the number of bursts within a certain brightness range, the observations of BATSE also confirmed that the capability to look at nearly the 'edge' of the GRB population was already reached [30]. An observer can infer if he is resolving the border of a population of similarly bright, isotropically distributed objects determining if the fainter ones are less frequent than expected. In the galactic and cosmological scenario the existence of a border for the GRB population carries very different meaning. In the galactic picture, the number of sources decreases with increasing distance from the galactic centre out to a distance where there are virtually no more left. In the cosmological scenario, the decline is due to the expansion of the Universe: at earlier times in the history of the Universe the volume of space was smaller than now, showing up in the distribution of the most distant, hence faint, bursts.
The BATSE observations showed also that GRBs were bimodally distributed in their duration and spectral features [69]. Short bursts have durations of few seconds and are dominated by higher-energy photons. On the other hand, long bursts can last up to several tens of seconds and their emission is dominated by lower-energy photons. In order to explain this distinction, astronomers proposed new models of GRB sources. Long GRBs were supposed to be emitted as the result of the collapse of massive stars while short bursts could originate from the merging of compact objects such as neutron stars or neutron stars and white dwarfs in exotic binary systems.

BeppoSAX By the mid 1990's astronomers had searched for counterparts to GRBs for decades. Finding what was emitting the gamma burst also meant finding where the bursts were emitted. In light of BATSE's recent observations, this was more important than ever since it could end the diatribe around the GRB scale distance. However, the lacking localization precision of the available

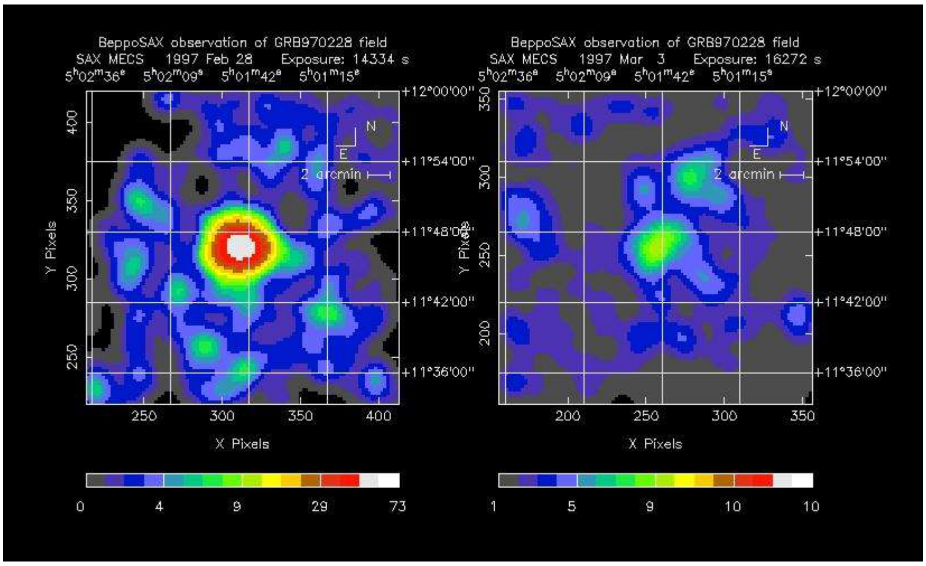


Figure 1.2: Original and follow-up GRB970228 X-ray images by BeppoSAX's Medium Energy Concentrator Spectrometer (MECS) narrow field telescope shows an unknown X-ray source fading over a week. From [38].

gamma detectors prevented this from happening. The best hope seemed to lie in finding a fainter, fading, longer wavelength emission after the burst itself. An *afterglow*, the 'smoking gun' behind the bang.

The italian-dutch satellite BeppoSAX was launched from Cape Canaveral on April 30, 1996. Originally named "Satellite per l'Astronomia X" (SAX), it was later renamed BeppoSAX in honour to the italian astrophysicist Giuseppe 'Beppo' Occhialini.

Through different instruments, BeppoSAX was able to cover a broad-band energy range, spanning 0.1 to 300 keV. The mission was arranged in such a way that after the initial detection in the higher part of the spectrum with coarse accuracy, BeppoSAX would have turned on the more accurate soft-X instruments towards the same area, looking for any fading afterglow with much higher positioning accuracy. This is what indeed did happen on 28 February 1997, leading to the discovery of the first GRB afterglow [38]. On that day, GRB970228 was detected by BeppoSAX's GRBM gamma detectors. BeppoSAX was able to repoint the narrow field X-ray telescope towards the gamma ray source. A few hours after the burst, an unknown X-ray source in the constellation of Orion was discovered and localized with accuracy of one hundreth of a degree. After about two days follow-up observations showed that the source's flux had dropped by about a factor of 20. In figure 1.2 the images from the original and the follow-up measurements are compared. Later images, taken after the point source faded revealed a faint galaxy at almost the same position, the presumed host galaxy of the burst; a chance position coincidence was unlikely yet possible, so the discovery was not conclusive until a few months later. On 8 May 1997, BeppoSAX succeded again at detecting GRB970508 and accurately localizing its X-ray afterglow. On 10 May a brightening variable optical source was reported within the error box of GRB970508. The following day, a team from the California Institute of Astronomy headed by Mark Metzger successfully acquired a spectrum of the source making use of the 10-metre Keck telescope. Metzger and his team reported the first discovery of the redshift of a GRB source to be $z = 0.835$, billions of light years away from our galaxy [82]. Such distances implied that GRBs were the most luminous class of events ever observed. From detected redshift and flux it was estimated that, if isotropic[1], GRBs could radiate between 10^{48} and 10^{55} ergs.

1.2 GRBs prompt emission observational properties

The NASA Fermi space observatory was launched during 2008. Fermi hosts the GBM and LAT experiments. GBM (Gamma-ray Burst Monitor) consists of 12 NaI and 2 BGO scintillation detectors. The NaI are sensitive in the energy ranges spanning a few keV to ~ 1 MeV, while the BGO scintillators are active in the energy range comprised between 150 keV and 30 MeV. The GBM field of view covers almost two-thirds of the celestial sphere, the whole portion of the sky not occluded by Earth. GBM achieves transient localization by algorithmic comparison of the number of counts observed by differently oriented scintillation detectors. GBM localization accuracy is in the order of magnitude of the degree for the brightest GRBs[81]. The Large Area Telescope (LAT) is an imaging gamma-ray detector able to detect photons with energy raging from 20 MeV to 300 GeV over a field of view covering $\sim 20\%$ of the sky. The LAT instrument is able to achieve arcmin localization accuracy, yet the LAT limited field of view will most often require a GBM trigger and a subsequent spacecraft slew for localization to be achieved at all [11].
Swift was launched in 2004. It is equipped with BAT (Burst Alert Telescope), a coded mask instrument operating between 15 and 150 keV with a field of view around 1/6 of the full sky. Alongside BAT, XRT is a narrow field X-Ray telescope able to provide localization of weak sources with arcmin accuracy [50].

---

[1] We will see how and why this seems to not be the case. Emission anisotropy brings back the upper bound to the true amount of energy released during a gamma burst to 10^{52} erg.

Through the 2000's and 2010's the observations of Fermi and Swift greatly improved our knowledge of GRBs. Many of these results are described in the present section whose intent is to provide a brief overview of the literature concerning the observational properties of GRBs.

Time domain properties The duration of a burst is generally quantified by the burst 'T_{90}', the time elapsed between the moment in which the 5% and the 95% of the total measured fluence has been detected in a specific energy range (traditionally set to $50 - 300$ keV). Observed values of T_{90} range from milliseconds to thousands of seconds and follow a log-normal bimodal distribution with a separation of components traditionally set around 2 seconds. Short Bursts have been described having T_{90} smaller than 2 seconds, with distribution mean at 0.2 - 0.3 seconds; while long Bursts have T_{90} greater than 2 seconds and mean around 30 seconds. Long GRBs are more common than short burst and account for almost 70% of the observed bursts [69]. In the latest analysis of the large Fermi-GBM sample wider T_{90} distribution is implied for short bursts, with mean around the 1 s and component separation exceeding 5 s [112]. Long GRBs are supposedly emitted as the result of the collapse of massive stars. Short bursts originate from the merging of neutron stars or neutron stars and white dwarfs in exotic binary systems. In recent years a new ultra-long class of GRBs has been proposed to classify rare events with durations larger than 1000 s [72]. A number of conflicting theories on ultra-long astrophysical progenitors exist such as collapse of blue super-giant stars or tidal disruption events [62]. On average, the photons energies are higher for short bursts than they are in long bursts [97]. The GRB light-curves are very irregular but still present notable features. While some bursts appear to be smooth, temporal variability has been detected down to the millisecond scale [115]. Hydrodynamic simulations and analytic studies both suggest that observations of GRBs lightcurve structure below the millisecond scale could be related to the source activity [86]. Bursts often present distinct emission events separated by periods of quiescence. The distribution of the separation times between emission pulses is log-normal [73]. Pulses in lightcurves are asymmetric and often rise faster than they decay. For this reason "FRED" (fast-rising exponential-decay) time profiles are often used to model individual pulses [70]. There are no evidences for periodicity in GRB lightcurves.

Twelve sample lightcurves from the BATSE catalog are presented in Figure 1.5.

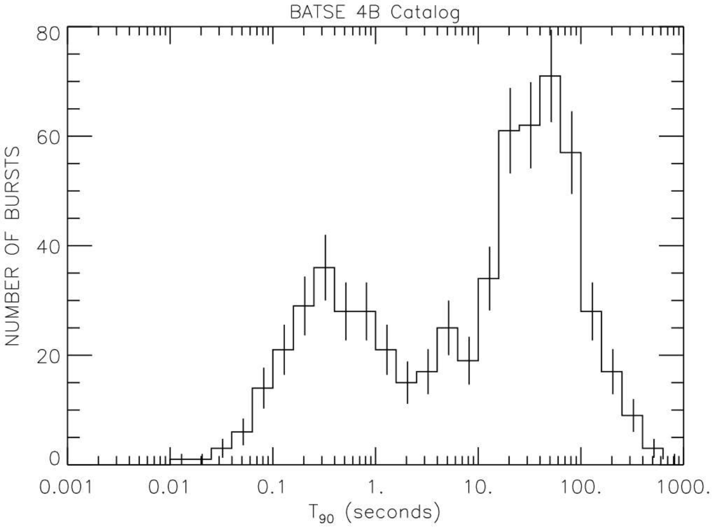


Figure 1.3: T90 duration for energies > 20 keV of gamma-ray bursts in the BATSE 4B Catalog. From [1].

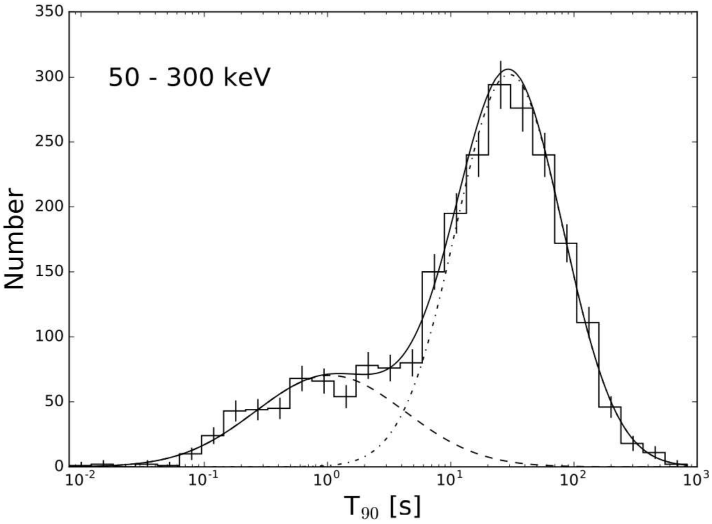


Figure 1.4: T90 duration in band $50-300$ keV of gamma-ray bursts in the fourth Fermi-GBM catalog. From [112].

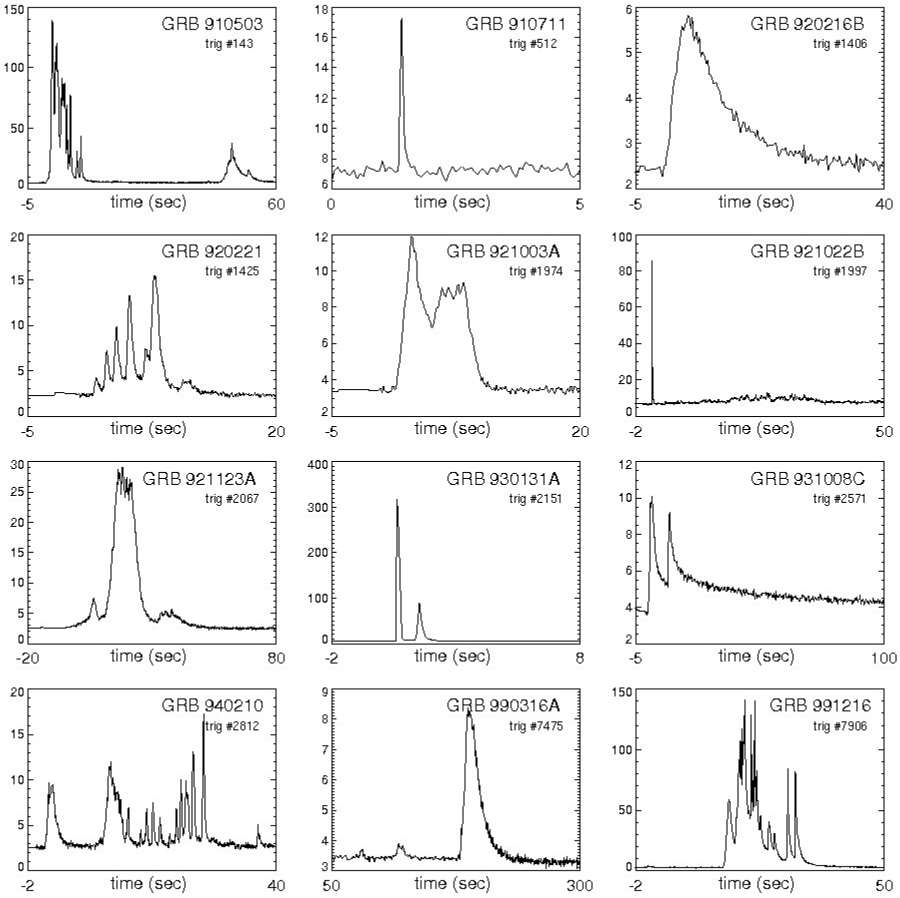


Figure 1.5: GRBs lightcurves samples showing both short and long bursts, events with smooth behaviour and events with high variability. Created with data from the public BATSE archive [1]. Author: Daniel Perley.

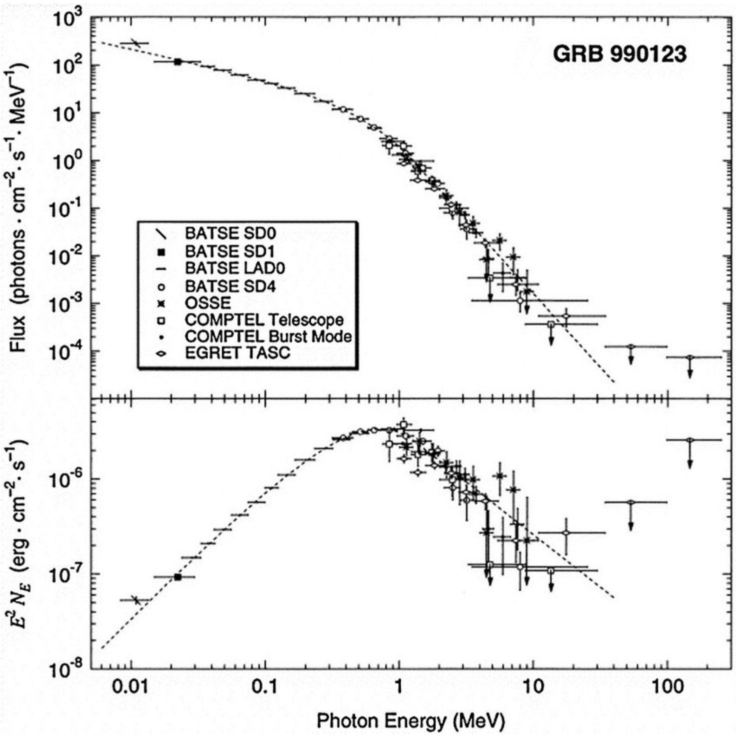


Figure 1.6: GRB990123 is fitted by a Band function. From [29].

Spectral properties GRB spectra are often fit to an empyrical, broken power-law model known as Band function [21]:

$$N(E) = \begin{cases} A\big(\frac{E}{100\,\mathrm{keV}}\big)^{\alpha} \exp\big(-\frac{E}{E_0}\big), & E < (\alpha-\beta)E_0 \\ A\big(\frac{(\alpha-\beta)E_0}{100\,\mathrm{keV}}\big)^{\alpha-\beta} \exp(\beta-\alpha)\,\big(-\frac{E}{100\,\mathrm{keV}}\big)^{\beta}, & E \geq (\alpha-\beta)E_0 \end{cases}$$

The number of photons detected between energies E and $E + dE$ is $N(E)dE$, α and β are the photon spectral indices, and E_0 is the break energy. The corresponding spectral energy distribution (SED) $E^2N(E)$ has a maximum at $E_p = (2+\alpha)E_0$. Typical values of the spectral indices for the bright BATSE sample are $\alpha \sim -1 \pm 1$ and $\beta = -2^{+1}_{-2}$ [96]. The peak energy E_p is widely distributed over a continuum ranging from tenth of keV to a few MeV [28]. GRB spectra do evolve with time and "hard-to-soft" evolution in the peak energy value is a common feature [88]. In Figure 1.6 the time-integrated GRB990123 spectrum is shown to fit the Band function over 4 orders of magnitude. [22].

On the basis of Fermi broad-energy observations, it is speculated that a combination of three spectral components may be shaping the time-resolved GRB spectra. These components are:

- A Band function component, non-thermal and extending over a wide energy range.
- A quasi-thermal component extending to low energy.
- A non-thermal component extending to high energies.

Many bursts spectra may be decomposed into one or more of these elemental components [123]. A small number of low-luminosity GRB have spectra which significantly depart from the Band function. For example, GRB060218 shows a thermal component in its X-ray spectrum, which in time cools and shifts into the optical/ultraviolet band [36]. Another notable example of GRB showing a significant thermal component is GRB101225A, commonly referred as the "Christmas Burst". Both GRB060218 and GRB101225A are characterized by many unusual features, most importantly their long-lasting prompt emission which exceeded 1000 s [51].

1.3 Models of GRB prompt emission

Two main models able to explain the γ-ray prompt emission from GRBs exist. The first and more affirmed model considers a "*fireball*" consisting of photons, electron and positron pairs and an amount of baryons [93]. A fireball could be produced in extreme astrophysical events, such as mergers and collapse of massive stars. The largest part of energy is initially stored in radiation but, as the fireball expands, energy is transferred to the baryons that are in turn kinetically accelerated to a high Lorentz factor. A fraction of the initial thermal energy is expected to be radiated away at the photosphere. At larger distances internal shocks tap into the jet kinetic energy in order to accelerate electrons which in turn produce non-thermal γ-rays via the synchrotron and inverse-Compton radiation. The main competitor model predicts the energy outflow to be dominated by Poynting-flux with particles accelerated from extreme magnetic fields through magnetic reconnection[109].
In this section we will give an outline of the physics behind the former model and discuss its strengths and limitations.

The fireball model We consider an ouflow with luminosity L and initial radius R_0. According to Stefan-Boltzmann law, the initial temperature T_0 is determined by the relation:

$$k_B T_0 \approx k_B \left(\frac{L}{4\pi R_0^2 g_0 \sigma_B} \right)^{1/4} = (1.3\mathrm{MeV}) L_{52}^{1/4} R_{0.7}^{-1/2} \tag{1.1}$$

where k_B and σ_B are respectively the Boltzmann and Stefan-Boltzmann constants, and $g_0 = 2.75$ is half of the effective degrees of freedom for a plasma consisting of photons, electrons and positrons at the thermal equilibrium. We use the notation $X_n \equiv X/10^n$.

We demonstrate now that the Lorentz factor of an adiabatically expanding spherical shell increases proportionally to the distance from the center up until a terminal value Γ_s, for as long as the system is optically thick to Thomson scattering.

Consider a spherical shell of radius r and width δr in the inner engine inertial frame, adiabatically expanding with a Lorentz factor $\Gamma(r)$.

The fireball dynamics is constrained by the conservation of energy flux and entropy.

Luminosity in the inner engine frame does not change as the shell expands:

$$L = 4\pi r^2 g(r) \sigma_B T^{'4} \Gamma^2(r) \tag{1.2}$$

Where $T'(r)$ denotes the shell temperature in the comoving inertial frame.

The entropy contained in the shell is:

$$s = 4\pi r^2 \delta r^{'} g(r) T^{'3} \tag{1.3}$$

The entropy is conserved for an adiabatically expanding shell and for Lorentz transformations. The shell width in the comoving frame is $\delta r^{'} = \Gamma \delta r$. Using equation 1.3 and entropy conservation:

$$\frac{T^{'}}{T_0} = \left[\frac{1}{\Gamma} \cdot \left(\frac{R_0}{r} \right)^2 \left(\frac{g_0}{g} \right) \right]^{1/3} \tag{1.4}$$

Substituting now in equation 1.2:

$$\Gamma(r) = \left(\frac{r}{R_0} \right) \left(\frac{g}{g_0} \right)^{1/2} \tag{1.5}$$

The Lorentz factor Γ continues to increase proportionally to r for as long as the system is optically thick to Thomson scattering so that photons and particles are coupled[2].

The terminal Lorentz factor is given by:

$$\Gamma_s = \frac{L}{\dot{M} c^2} \equiv \eta \tag{1.6}$$

---

[2] i.e. if the photon energy per baryon is much larger than baryon rest energy, $3 k_b T^{'} n_\gamma \gg m_B c^2$

where $\dot{M}$ is the baryonic mass flux associated with the outflow. The Lorentz factor terminal value Γ_s is attained at saturation distance $R_s \sim R_0\Gamma_s$.

In order for the prompt emission to reach us it is needed for the electron/-positron pairs to cease being in thermal equilibrium with the plasma and annihilate, making the fireball transparent to radiation. We will see that this crucial observation comes with a caveat regarding the existence of a baryonic component in the fireball.
The main observation we address now is that pair annihilation/creation freezes when the expanding, cooling shell reaches $T_{freeze} \approx 20.5\,\mathrm{keV}$ or $\Gamma_{freeze} \approx 64$. If the baryonic load is zero, the freeze-out radius equals Thomson photospheric radius while $\Gamma_s \longrightarrow \infty$ and so pair annihilation/creation continues for as long as the shell accelerates.
The cross-section for pair annihilation is

$$\sigma_{e^\pm \longrightarrow 2\gamma} = \frac{\sigma_T}{\langle v\rangle/c}.$$

Thus, the comoving frame time for a positron to annihilate with an electron is:

$$t'_{e^\pm \longrightarrow 2\gamma} = \frac{2}{\sigma_{e^\pm \longrightarrow 2\gamma} n' \langle v\rangle} \approx \frac{2}{\sigma_T n' c}. \tag{1.7}$$

The factor 2 in the numerator is due to the fact we are just considering electrons (and not pairs); n' is the electron-positron pairs number density in comoving frame:

$$n'_\pm = \frac{2(2\pi k_B m_e T')^{3/2}}{h^3} \exp\left(- \frac{m_e c^2}{k_B T'}\right) \tag{1.8}$$

The process of pair annihilation/creation freezes when $t'_{e^\pm \longrightarrow 2\gamma}$ approximates the dynamical time $\sim r/c\Gamma(r)$, which happens at $T'_{freeze} \approx 20.5\,\mathrm{kev}$. From Eq. 1.8 follows that when baryon loading is negligible the freeze-out radius equals the Thomson-photospheric radius. If the freeze-out occurs while the jet is accelerating, then $\Gamma(r)/r \sim 1/R_0$ and, from equation 1.7, we have:

$$\sigma_T n'_\pm R_0 \sim 2.$$

Making use of equation 1.8 we obtain the following equation for the freeze-out temperature:

$$T'^{3/2} e^{-\frac{5.9\times10^9}{T'} \approx 62 R_{0,7}^{-1}}.$$

The solution of which is $T'_{freeze} \approx 20.5\,\mathrm{keV}$. The Lorentz factor at the freeze-out:

$$\Gamma_{freeze} \sim \frac{T(R_0)}{T'_{freeze}} \tag{1.9}$$

Considering that when freeze-out takes place the fireball is dominated by photons and $g = 1$, we can estimate the freeze-out radius:

$$R_{freeze} \sim R_0 \Gamma_{freeze} (g_0/g)^{1/2} \sim 1.7 R_0 \Gamma_{freeze}. \tag{1.10}$$

We consider now the effect of a baryonic component on the fireball dynamics and demonstrate that the presence of minimal baryon loading makes possible for the fireball to continue to accelerate after the pair annihilation/creation freeze-out, up until the Lorentz factor reaches its terminal value or the outflow breaks over the Thomson photospheric radius.
The number density of electrons associated with protons is:

$$n'_p = \frac{\dot{M}}{4\pi r^2 m_p c \Gamma} = \frac{L}{4\pi r^2 m_p c^3 \eta \Gamma} \tag{1.11}$$

The number density of electrons associated with protons at the freeze-out is:

$$n'_p = \frac{L}{4\pi R_0^2 m_p c^3 \eta \Gamma_{freeze}} \tag{1.12}$$

If the fireball dynamics after freeze-out is dominated by electrons associated with protons we have that:

$$n'_p(R_{freeze}) > n'_{\pm}(R_{freeze}) \sim \frac{2}{\sigma_T R_0}.$$

Using the last equation as well as equations 1.12, 1.11 and 1.1 we obtain that whenever the condition

$$\eta = \frac{L}{\dot{M} c^2} < \frac{L \sigma_T}{8\pi R_0 m_p c^3 \Gamma^3_{freeze}} \sim 2 \times 10^6 L_5^{1/4} 2 R_{0,7}^{1/2} \tag{1.13}$$

is satisfied, the jet continues to accelerate for $r > R_{freeze}$ until $\Gamma \sim \eta$ or the outflow reaches the Thomson photospheric radius. It is evident that in order for condition 1.13 to be satisfied a certain baryonic outflow $\dot{M}$ is needed. This minimal amount of baryons is found to be very low, $\sim 10^{-7} - 10^{-5}\,\mathrm{M}_{\odot}$. For this reason, it is reasonable to expect the condition 1.13 to hold true for most GRBs.

Internal shock scenario The last ingredient needed to sort out GRB prompt emission is a mechanism converting the kinetic energy transported by jet baryons into the gamma-ray photons we observe at distance. The most well accepted answer is known as the *internal-shock* scenario.
Consider a relativistic, baryonic outflow regulated by a time-dependent Lorentz

factor. The fastest shells composing the outflow catch up the slower ones, moving ahead of them. Collisions between relativistic shells transform the kinetic energy, in largest part transported by baryons, into thermal energy, thermal energy that is in turn radiated away, in largest part by electrons, via synchrotron radiation in presence of a magnetic field.
From a thermodynamical point of view one may say that collisions make it possible to pass from an ordered relativistic bulk motion to an highly disordered state. Relativistic electrons in stochastic motion are then able to irradiate via synchrotron emission in presence of an external magnetic field.
The fireball-internal shock model for GRB prompt radiation is not flawless. It does in fact suffer from two major issues:

- The GRB prompt emission spectra do not agree with synchrotron emission at low energies (see Sec. 1.2).
- Internal shocks are highly inefficient at converting energy [71]. On the other hand, the external shocks powering afterglow emission are expected to be efficient. This asimmetry has no equivalent in the observed emitted energy of afterglow and prompt radiation, which are similar.

As briefly discussed in the previous paragraphs the most popular emission models interpret GRB prompt emission as synchrotron emission by relativistic electrons embedded in an intense magnetic field. The synchrotron model is also the most natural option, since its basic ingredients - namely accelerated electrons and intense magnetic field - are provided by the internal shock scenario, in which strong shocks determined by collisions between relativistic shells emitted by the inner engine accelerate particles to relativistic velocities and compress and amplify magnetic fields.
Unfortunately, the synchrotron model does not come without its own problems. Shock-accelerated electrons are expected to have the following energy distribution:

$$\frac{dN_e}{d\gamma} \propto \begin{cases} \gamma^{-p}, & \text{if } \gamma \geq \gamma_m \\ 0, & \text{if } \gamma < \gamma_m \end{cases} \tag{1.14}$$

The fast-cooling synchrotron photon flux N_ν is:

$$N_\nu = N_{\nu,max} \begin{cases} (\nu/\nu_c)^{-2/3}, & \text{if } \nu_c > \nu \\ (\nu/\nu_c)^{-3/2}, & \text{if } \nu_c > \nu > \nu_m \\ (\nu_m/\nu_c)^{-3/2}(\nu/\nu_m)^{-(p+2)/2}, & \text{if } \nu > \nu_m \end{cases} \tag{1.15}$$

where $N_{\nu,max}$ is the peak flux, ν_c is the cooling frequency, ν_m is the characteristic synchrotron frequency and $p \sim 2.5$ is the power index characterizing the

electrons energy distribution.
The synchrotron characteristic frequency ν_m is related to the minimum Lorentz Factor by the approximate relation $\nu_m \propto \gamma_m^2 B\Gamma$.
The cooling frequency ν_c is a frequency associated to γ_c, the Lorentz factor of relativistic electrons that, after the cooling time $t_c \propto \gamma^{-1}B^{-2}$, have radiated most of their energy by synchrotron emission. The cooling time is comparable to the dynamical time comparable to the dynamical time $t_{dyn} \sim \frac{R}{\Gamma^2 c}$. Although still lacking in the low energy ($< 10 - 20$ keV) part of the spectrum, prompt emission spectral data suggest the photon index α before the peak to be distributed around a typical value $\langle\alpha\rangle \sim -1$ higher than the value expected in case of fast-cooling synchrotron radiation, $\alpha \sim -3/2$ [89].
The problem has been widely discussed. The largest part of proposed solutions fall into two classes: 1. models advocating corrections to the basic fast-cooling synchrotron scenario; 2. models rejecting the synchrotron emission mechanisms. Fact is that both approaches come with their own theoretical problems. As a result, to date there is still no consensus on the origin of prompt emission.
Recent interesting developements come from the work of Oganesyan, Nava and others [89]. From the analysis of 14 GRBs for which Swift XRT data were available during the prompt emission phase it emerged that for more than half of the sample events the Band function model fails reproducing the 0.5 - 1000 keV spectra. Indeed, low-energy data suggest the presence of a spectral feature around a few keV, compatibly with the theoretical expectations from synchrotron radiation in fast-cooling regime, where a break is expected at the cooling energy $E_c = h\nu_c < E_p$.
Similar analyses have been extended to larger samples, leading to very close results [90].
The XRT aboard Swift is the only instrument online at moment capable of a detailed description of the low energy prompt spectra. However XRT is a narrow-field instrument that needs to be timely slewed to source to gather data. Considering the brief duration of GRB prompt emissions this is not feasible in the vast majority of cases.
We deduce that the activity of a new instrument capable of a detailed and timely characterization of the low energy prompt spectra could result in an improved theoretical comprehension of GRB emission mechanisms.
GRB prompt emission models built around the internal shocks scenario have an efficiency problem. In collisions, efficiency depends on the relative kinetic energy available to be dissipated. Dealing with the internal shock scenario we compare a number of shells accelerated to similar Lorentz factors. As a result only a few percent of the total kinetic energy is available to be transferred to the fireball particles to be irradiated. Moreover, since the synchrotron power is in-

versely proportional to the particle squared mass, protons are poor synchrotron emitters and of this relatively small amount of energy only the fraction available to the electrons is radiated. Quantitatively, the expected efficiency of internal shocks is attested around 5%.

On the other hand, when the relativistic merged shell impacts on the external medium surrounding the source a large fraction of the total kinetic energy can be dissipated. This implies that external shocks are expected to be highly efficient. However observations fail to show the expected difference between prompt emission and afterglow total energy: the observed emitted prompt energy is often comparable, or even greater, than the afterglow emitted energy.

Proposed solutions to the efficiency problem traditionally exploit a distribution or spectrum of Lorentz factors among different shells in order to make possible for internal collisions to happen with very high Γ-contrasts.

1.4 Time variability and the inner engine

The fireball model supposes the existence of an *inner engine* able to accelerate a baryonic wind into highly relativistic motion. The nature of the inner engine is to this day still mysterious. This is expected in the fireball model: a thick photosphere prevents direct observation of the inner engine. When looking at gamma-ray bursts astronomers are left in the frustating situation where they can see the most powerful accelerators of the universe at work while not being able to take a glance at what is causing them.

Numerical simulations of internal shocks suggest that the observed variability could track the activity of the emitting source [67]. If this is the case, GRB light-curves can provide a useful timetable of the processes having place in the inner engine.

In 2002 paper, Nakar and Piran purposed a simple analytical model in order to clarify the relations between the observed light-curve variability and the activity of the inner engine [86].

In the internal shock scenario we assume two shells of matter with widths l_1 and l_2 to be ejected while separated by L at times

$$t_2 \sim t_1 + (l_1 + L) \tag{1.16}$$

with Lorentz's coefficient $\Gamma_1 = \Gamma$ and $\Gamma_2 = a \cdot \Gamma$ respectively. In the previous relation and in the following, c is supposed to be normalized to unity.

We also suppose the second shell to be faster than the first so that $a \gtrsim 2$. The second shell will impact on the first at :

$$R_s \sim \Gamma^2(2a^2/(a^2-1)) \sim 2\Gamma^2 L \tag{1.17}$$

The photons emitted during the collision will reach the observer at:

$$t_{obs} \sim t_1 + l_1 + \frac{R_s}{2\Gamma^2} \sim t_1 + l_1 + L$$

Substituting L from equation 1.16, we observe that $t_{obs} \sim t_2$. This implies photons from the collision will be observed almost simultaneously with an hypothetical photon emitted from the 'inner engine' with the faster shell.
Moreover, since the width of observed pulses can be estimated considering that beaming photon moving towards the observer will arrive before photons moving at an angle $\theta \sim 1/\Gamma$, we have that $\Delta t \sim R_s/2c\Gamma^2 \sim L/c$. Substituting 1.17 in the last equation:

$$R_s \sim 2\Gamma^2 c\Delta t$$

Fast variability gives a scale of the distance R_c between the inner engine and the emitting region, where collisions take place. Considering $\Gamma \sim 100$ and $\Delta t \sim 1\,\mathrm{ms}$, we get a distance in the order of half an astronomical unit.
The GRB variability window below a few milliseconds is, to date, still scarcely explored. The main constraining factor limiting our ability to detect fast variability in GRBs is detector area. For example, the photon flux count in the energy band between 50 and 300 keV for a typical GRB is $\sim 10\,\mathrm{ph\,s^{-1}cm^2}$, in order to provide 1 count every 10 μs a large collecting area of $\sim 1\,\mathrm{m}^2$ is needed; compare with the100 cm^2 effective area of the Fermi-GBM NaI detectors at energy 100 keV.
Still there is evidence for millisecond structure to be a rather common bursts feature. Power density spectrum analysis of a few hundreds joint Fermi and BeppoSAX GRBs sample, using temporal resolution down to 0.5 ms, did find signal up to 10 Hz, suggesting average intrinsic variability timescale $\lesssim 0.1$ s below which the temporal power changes regime [39]. Analysis of GRBs from the BATSE catalog resulted in the observation of 'flickering' on timescales ranging from $33\,\mu$s to $256\,\mu$s [116]. Similar studies have been performed on the Fermi burst catalog resulting in somewhat longer average minimum variability timescales yet supporting a scenario in which millisecond variability is not an uncommon feature [53].

1.5 Gamma-ray bursts progenitors

The fireball model explains reasonably well the physics of gamma bursts. However, it does not answer the most interesting question from an astrophysical point of view: which astronomical object can produce GRBs?
In order to answer this question it will be useful to keep in mind some remarkable features of GRB emission.

- An enormous amount of energy is estimated to be released during a GRB. To release $\sim 10^{51}$ ergs, the inner engine must be able to accelerate $\sim 10^{-5}$ solar masses to relativistic velocities.
- GRB emission is collimated within small opening angles. The inner engine should then be capable to collimate a relativistic flow of energy and matter.
- GRBs show variability down to the millisecond timescale.
- Depending on their duration, GRBs fall in two main categories: long and short GRBs. Since the burst duration is determined by the processes having place in the inner engine the existence of these two categories may imply the existence of two different types of inner engines.
- GRBs take place roughly once per 3×10^5 years per galaxy, about one part in three thousands the rate of supernovae. GRBs progenitors are rare, extreme astronomical objects.

All these clues hint to black holes. GRBs may arise due to rapid accretion from a massive disk (0.1 solar masses per second) onto a compact object. A disk of this dimension can only form simultaneously with the compact accreting object, leading to the conclusion that the emission of GRBs follows the formation of a newborn black hole. This theory is supported by the observations of relativistic jets in active galactic nuclei, which again are powered by accretion onto black holes.
Simultaneous generation of a black hole and massive accretion disk can arise from different scenarios such as compact binaries mergers and ”failed supernovae” in rapidly rotating, massive stars. It has been shown that of all the above scenarios, only neutron star-neutron star and neutron star-black hole mergers could produce short bursts, while only collapsar could produce long bursts [87]. Strong observative evidence came from the simultaneous detection of a short GRB and gravitational signal carrying the neutron start merger “signature” on 17 August 2017.
In these models the duration of the accretion episode depends on the size of the disk. Since mergers could produce only “small” disks, short bursts must originate from such systems. Naively long bursts will of course be expected from

large disks. However, since large disks accretion is inherently inefficient this seems not to be the case. Collapsars are expected to produce a small accretion disk continuously feeding a central newborn black hole for a long (several tens of seconds) period of time. In this case the efficiency can be large and the emission duration long.

Collapsars and long bursts In the collapsar model [120], the GRB progenitor is a rapidly rotating ($\sim 200\ \mathrm{km\,s^{-1}}$ at surface), massive Wolf-Rayet star with mass $\mathrm{M} \gtrsim 30\,\mathrm{M}_\odot$. During their lifetime, Wolf-Rayet stars lose their external hydrogen envelope through intense stellar winds resulting in an exposed helium core. When the nuclear fuel is exhausted, the iron core collapses into a black hole, either directly or during the accretion phase following core collapse.

An accretion disk of mass $\sim 1\ \mathrm{M}_\odot$ rapidly forms around the newborn black hole, funneling matter along the rotation axis where angular momentum is low. The black hole is expected to accrete from the disk over several tens of seconds. Accretion powers up a long GRB extracting energy via neutrino annihilation [77]. The energy deposited in the surrounding matter leaks out along the rotation axis producing jets with small opening angles, eventually penetrating through the stellar material envelope and producing the GRB.

Relativistic jets are collimated by their passage through the stellar mantle. In particular, it is expected that starting with initial half-apertures up to 20 deg, the jet will emerge with half-angles reduced to $\sim$ 5 deg [124]. In order for the jets to reach the star surface, progenitor stars should have lost their hydrogen envelope. Conversion of internal energy into kinetic energy results in a terminal Lorentz factors ~ 150 along the axis. Such a Lorentz factor is compatible with the theoretical predictions.

Observations hint towards an association of long GRBs with Type-Ic supernovae ([35], [59] and more). It should be noted that the contrary does not hold true: a radio survey of supernovae type Ib/c suggests that less than 3% are associated with a GRBs [26]. Long GRBs often occur in the brightest regions of their host galaxy, suggesting very high star formation rates at the burst site henceforth higher occurrence of massive stars [47].

Mergers and short bursts Mergers, a term which we will use to indicate neutron star binary mergers or neutron star - black hole binary mergers, can produce a black hole - accretion disk system.

Mergers take place because of the decay of angular momentum due to gravitational radiation emissions. Two degenerate objects will spiral closer and closer together, until in the last few moments, tidal forces rip the neutron star (or stars) apart and an immense amount of energy is liberated before the matter plunges from an accreting disk into a black hole along the rotation axis, similarly to how it happened for collapsars. The whole process is expected to take place over a fraction of a second, accounting for short GRB's duration.

Mergers do release $\sim 5 \times 10^{53}$ ergs. Most of this energy will be emitted in neutrinos and gravitational waves, but enough will be produced to power up the burst emission.

In agreement with the merger models, searches of supernovae associated with short GRB have been carried out without successful detection [25].

Short GRBs host galaxies have been analyzed and compared with the hosts of long GRBs and Type Ia supernovae [46]. Results show that most short GRB hosts have exponential disk profiles, characteristic of late-type galaxies, but with a median size twice as large as that of long GRB hosts. More importantly, the observed short GRB projected physical offset distribution has a median about a factor of 5 times larger than long GRBs, which instead did show strong concentration into the brightest, central region of their host galaxy. The observed physical offset distribution matches the predicted value of neutron star binaries: compact stars born in asymmetric supernovae most likely received a "kick", so that the binary system drifted away from the star forming regions when mergers occur.

Kilonovae, weaker than supernovae in their transient optical/IR emissions, have been predicted to be associated with mergers events. After observational suggestions of a connection between kilonova events and short GRBs, kilonova emission has been observed in association with the event observed on of 17 August 2017 [110].

1.6 Detecting Gamma-ray Bursts

GRB-monitoring experiments on-bard spacecrafts are able to detect GRBs automatically. A detection signal is generally used to trigger an on-board recording apparatus. This task has been crucial in the past due to the limitations of memory storage devices and communications. Today spacecrafts like Fermi are able to store and communicate to the ground the high-resolution data product gathered during their operations. Still, being able to generate a GRB detection alert, timely and on-board, remains of paramount importance. For example, a timely detection alert can be used to trigger a repointing maneuver enabling the observations of a secondary, fixed-axis instrument (as in the case of Fermi and Swift); or to initiate follow-up observations by ground-based telescopes with a narrow field of view.

At a fundamental level, techniques for detecting GRBs have gone largely unchanged through different generations of spacecrafts and experiments. Discrete, high-energy events are counted over a fundamental acquisition time interval and compared against a background estimate over pre-set timescales. A trigger is issued whenever count rates rise to a level above the background exceeding a pre-set threshold.

In the following paragraphs the literature on the algorithms implemented by Compton-BATSE, BeppoSAX-GRBM and Fermi-GBM are discussed and compared. When possible, particular attention will be given to the operation performances and criticalities of each technique for GRB detection.

Compton-BATSE The computer on-board Compton-BATSE detected GRBs monitoring count rates from each of the eight LAD detectors [68]. Count rates were formed starting from the signal of the four LADs discriminator channels. The most commonly used channels for forming count rates spanned the energy ranges $50 - 100$ keV and $100 - 300$ keV (labels 2 and 3) [92]. Fiducial background count rates were assessed and recomputed every 17.408 s. Count rates were monitored over three different time scales with durations 64 ms, 256 ms and 1024 ms. The on-board computer issued a trigger whenever at least two detectors simultaneously observed a count excess with significance exceeding an adjustable threshold. Different significance thresholds were tested during BATSE operations. The most commonly used threshold value was 5.5σ. Higher threshold values (up to 26σ) have been tested, especially when count-rates comprised the signal of the low-energy discriminator channel 1 (25-50 keV) [92].

BeppoSAX-GRBM The Beppo-SAX Gamma-Ray Burst Monitor was an experiment derived from a secondary function of the anti-coincidence detectors of the BeppoSAX Phoswich Detection System (PDS) experiment. Count rates were formed from events with energies comprised between 20 and 600 keV. The lower and upper band thresholds were adjustable in the ranges 20-90 keV and 200-600 keV. The fiducial background was computed by a rolling moving average with time resolution 7.8125 ms and a long integration time (LIT) adjustable between 8 and 128 s. Count-rates were monitored over a single short integration time (SIT) with value adjustable between 7.8125 ms and 4 s. A trigger was issued whenever an excess in count rate with significance exceeding 4σ, 8σ or 16σ was observed over at least two detectors. At the beginning of the missions, the GRBM trigger operated with lower threshold 32.8 keV, upper energy threshold 604 keV, SIT equal 4 s, LIT equal 128 s and standard deviation threshold 8σ. The low-energy threshold was later increased to 42.5 keV in order to suppress frequent spurious detections. These spurious detections or 'spikes' were due to

CsI scintillators emission induced by cosmic rays with high atomic number Z. The situation was aggravated by the choice of the SIT parameter. Many of the spurious events resulting from cosmic rays could in fact be rejected by the requirement of meeting the trigger conditions simultaneously over two detectors; yet the large SIT value of 4 seconds enhanced the probability of getting a simultaneous trigger by chance. On its own, the 8σ threshold level was deemed unable to provide an efficient discrimination between GRBs and spikes in count-rates due to cosmic rays. The protection against spurious triggers obtained increasing the low-energy threshold allowed to reduce the standard deviation threshold to 4σ. Moreover, the SIT was lowered to 1 s. This setting allowed to achieve better sensibilities to short events. With these parameters, GRBM triggered an average of 12 times per day. Most of these events were false triggers and were still due to high-Z cosmic rays intercepting two contiguous detectors and being therefore in time coincidence[44].

Fermi-GBM The Fermi on-board computer forms and monitors count rates for each of the Gamma Ray-burst Monitor (GBM) NaI and BGO scintillation detectors separately. Count rates are compared against an average background rate accumulated over an adjustable duration, nominally set to 17 s. The signal window over which the background rate is computed does not include the most recent 4 seconds of observations. The GBM algorithms for GRB detection can operate in four different energy ranges: 25 - 50 keV, 50 - 300 keV, > 100 keV and > 300 keV. Ten timescales spanning 16 ms and 16.384 s in step factor of 2 are implemented for the 25-50 keV and 50-300 keV ranges. All these timescales but the 8.192 s and 16.384 s ones are also supported in the > 100 keV range, while the > 300 keV trigger supports only short timescales with duration 16 ms, 32 ms, 64 ms and 128 ms. For each of these algorithm criteria two different phase offsets are selectable with length equal to half the timescale duration [81]. Four extra high-energy (2-40 MeV) trigger criteria with timescale 16 ms have been implemented specifically for the BGO detectors. These criteria are designed for the detection of Terrestrial Gamma-ray Flashes (TGFs) and were introduced through a flight software update which took place during the beginning of 2009 [112]. In total, the GBM flight software supports the simultaneous operations of 119 different trigger criteria, each with adjustable standard deviation threshold levels [113]. For a trigger to be issued a simultaneous exceedance of the trigger threshold over at least two detectors is required.

The Fermi-GBM trigger burst catalog papers present in detail the chronology of the operations of each of the trigger criteria [91] [113] [27] [112]. At launch (June 2008), only 19 trigger criteria were operated. On 2008 July 14th, 66 trigger criteria were operating simultaneously. Higher threshold values were specified for criteria with the shortest timescales or operating in the low-energy range. On 29 October 2009 most of the low-energy trigger criteria were disabled to reduce the computational burden on the on-board computer [91] and to reduce the number of triggers by non-GRB events [113]. In 2011 and 2012 some trigger algorithms were disabled and the low energy thresholds raised for all the GBM sun-facing NaI detectors (labels 0-5). These adjustment were meant to mitigate the effects of high solar activity on the Fermi-GBM false trigger rates and to avoid saturating the memory of the on-board data storage devices [113].

During the first four years of GBM operations approximately 10% of the triggers were due to cosmic rays or trapped particles. A smaller fraction of $\sim 6\%$ triggers were explained as statistical fluctuations or events too weak to be classified [113].

Swift-BAT The Burst Alert Telescope (BAT) supports about 800 different criteria to detect GRBs [79] and three different algorithms for GRB detection. Two of these algorithms are count rate triggers intended for the detections of short and long transients. The last trigger logic exploits the imaging capability of BAT and is presently not taken into account [43].

The BAT short rate trigger forms separate counts for nine different region of the focal plane and four energy ranges: 15 - 25 keV, 15 - 50 keV, 25 - 100 keV and 50 - 350 keV. Count rates are tested over timescales of 4, 8, 16, 32 and 64 ms against the same fiducial background rate. The background rate are determined by the long trigger algorithm. To reduce the computational burden of the short rate trigger, a significance score is computed only for the maximum count observed over the five timescales every 256 ms. The short trigger significance is computed according to:

$$S = \frac{(C_{i,k} - B_i 2^{k-10})^2}{C_{i,k} + \sigma^2_{min}} \tag{1.18}$$

where the i label define one of the 36 region-energy combination, B_i are the background counts expected over 1024 ms and $C_{i,k}$ is the maximum count observed over the 2^k ms timescale. The variable σ^2_{min} is used to ensure for a minimum variance when counts are small. A trigger is declared whenever S exceeds a threshold value σ^2. This logic allows for the definition of 180 different trigger criteria [43].

The BAT long rate triggers manage 36 time series in circular buffers, one for each energy-region combination. These time series are divided in a background and a foreground section. To avoid trends in background estimates, a polynomial is fitted to the background sections of each circular buffer. The background region can either predate the foreground section (extrapolation) or “bracket” it (interpolation). To obtain the long rate trigger score, the count and variance in the foreground section are first evaluated, then the background section is fitted and the expected background rate computed accordingly, finally the trigger score is computed according to an equation similar to Eq. 1.18. Each BAT long rate trigger is controlled by about 30 adjustable parameters defining, among others, the length of the foreground and background sections, threshold values and several parameters to control the CPU usage [43].

1.7 Multi-messenger astronomy

On August 17 2017 12:41:06 UTC, the Gamma-ray Burst Monitor (GBM) instrument aboard the Fermi Space Telescope was triggered by a GRB, later known as GRB170817 [8].

Just 6 minutes later, based on a single-detector analysis of the Laser Interferometer Gravitational-wave Observatory (LIGO) Hanford data, a gravitational-wave event was registered. The same signal was detected at LIGO-Livingston and Virgo interferometers. Similarly to the GRB, the gravitational event was named GW170817. The signal was soon determined to be consistent with a binary neutron stars coalescence event with merger time t_c =12:41:04 UTC, ending 2 seconds before the GRB170817 trigger time. A Gamma-ray Coordinates Network (GCN) notice was issued at 13:21:42 UTC, reporting that a highly significant gravitational wave candidate event consistent with a binary neutron star coalescence was associated with the observation time of GRB170817. A rapid re-analysis of data from LIGO-Hanford, LIGO-Livingston and Virgo confirmed highly significant, coincident signals. The data from the three interferometers were combined and the first skymap was produced at 17:54:51 UTC, five hours after the first detection. It placed the source in the galactic neighborhood, at a luminosity distance about 40 ± 8 Mpc from Earth, somewhere in an elongated region of the celestial sphere covering about 31 deg^2. The source position was also constrained by Fermi-GBM data analysis to a region of $\sim$ 1000 deg^2 overlapping that identified by the ground interferometers. The observed time delay between the signal detected by Fermi and INTEGRAL SPI spectrometer made it possible to triangulation of GRB position. The electromagnetic signal was located in a region of $\sim$ 100 deg^2 extended along the area identified by the ground interferometers, as in Figure 1.7.

All evidences pointed towards GRB170817A and GW170817 being generated by the same source, thus implying that for the first time a gravitational wave event was observed with an electromagnetic counterpart.

Both Fermi and INTEGRAL analysis agreed in classifying the burst as a *short* GRB. During the 2000's, strong theoretical hints suggested an association between short GRBs and mergers of neutron stars with other neutron stars or black holes. Theory predicts neutron stars binary mergers to be associated with

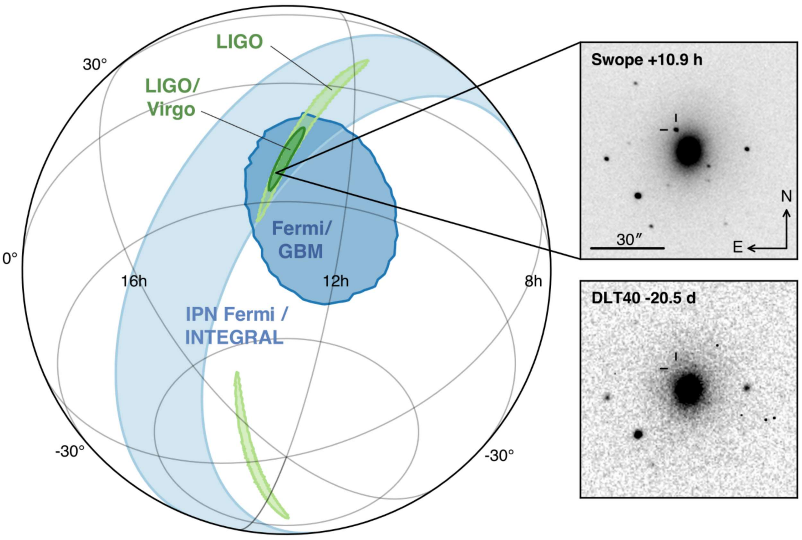


Figure 1.7: Localization of the gravitational-wave, gamma-ray, and optical signals for the 17 August 2017 event. Left panel shaded regions identify 90% confidence level. From [8].

isotropic electromagnetic signal which included optical and infrared emission, the kilonova, and delayed radio emission from interaction between ejecta and the circumburst matter. The announcements of Fermi and LIGO-Virgo detections was hailed with great enthusiasm from different research groups all around the world. Astrophysicists were finally having a chance to directly prove their models of bursts and binary mergers, and the existing relations between the two. However, by this time, the fact that the two signals were in fact coming from the same source was still to be established with statistical certainty. The localization of an optical counterpart to the event would have provided such evidences. In order to rule out the possibility that the two signals were related by chance, scientists had to identify an host galaxy that was: 1. located at a distance compatible with the LIGO data and 2. exhibiting unexpected activity due to the burst emission.

An unprecedented large broadband observation campaign soon started [8]. An optical transient counterpart, SSS17A, was observed from Chile for the first time at $t_c + 10.87$ hr by the One-Meter, Two-Hemisphere team. The transient was associated with the galaxy NGC4993 at distance ~ 40 Mpc, compatibly with the gravitational-wave source distance estimate. In the following two weeks, different telescopes from all over the world investigated the fading electromagnetic counterpart of GW170817 spanning the IR to UV spectrum, looking for activity related to the burst emission. With exception of an excess in the blue part of the spectrum [95], the evolution of the spectral energy distribution of the SSS17a emission was coherent with the theoretical models of kilonovae. For the first time direct observations supported the theoretical links between gravitational events, short GRBs and kilonovae.

Chandra and VLA were the first observatories to claim the detection of the afterglows of GRB170817 in X and Radio bands respectively. NGC4993's X and Radio emissions were closely monitored in the years following August 2017, allowing to constrain the geometry and dynamic of the event [56].

The standard model for short GRB afterglows predicts synchrotron emission from a decelerating and decollimating relativistic jet. If the jet propagates through a baryon contaminated region surrounding the merger site, a hot *cocoon* forms from which the jet either emerges collimated (successful jet) or not (choked jet). The cocoon subsequently expands almost isotropically producing its own prompt emission and external shock powered afterglow.

X-ray and very long-baseline interferometry radio observations suggested a scenario in which early-time radio emission was powered by the cocoon while the late-time emission was dominated by an energetic and narrowly collimated jet, observed off-axis from a viewing angle of about 20 degrees [107][84].

1.8 Challenges to the next generation of multi-messenger observations

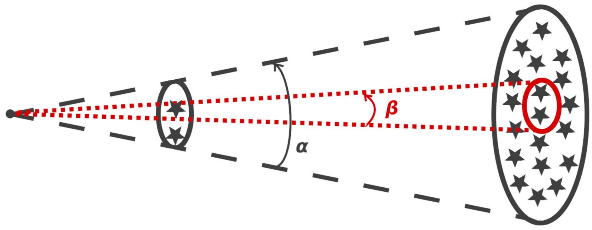


Figure 1.8: Number of sources in a patch of sky grows with the third power of the distance. In order to guarantee prompt optical counterpart detection, in the next decade localization accuracies in the order of the arcmin will be necessary.

The key point that made possible the identification and characterization of GW170817 is the rapid detection of an optic counterpart to the event in galaxy NGC4993. The unambiguous identification of a source has in turn been possible only thanks to the relative closeness of the event at 40 Mpc and the combined

Advanced Ligo/Virgo - GBM ~ 30 deg^2 accuracy. These circumstances made it possible to reduce the source pool to just ~ 50 candidate galaxies.
In the next decade LIGO is expected to improve its sensitivity while two new detectors will become operational: the japanese KAGRA [19] and the indian third detector of the LIGO family, IndIGO [108]. If, on one hand, the enlarged detector network will make localization accuracy below ~ 10 deg^2 possible, on the other hand, a large fraction of the binary neutron stars merger events will be discovered at increasing distance, up to almost 200 Mpc, because of the improvement to LIGO detector sensitivity expected starting from the O4 observation segment [9]. With increasing source distance comes an increasing source candidate pool, hosting possibly thousands of galaxies. A larger candidate pool makes it harder to promptly and unambiguously identify the sources. Will the new interferometers positioning accuracy be enough to keep the pace with increasing detection distance? Thanks to a simple geometry argument we can assert that this will not be the case and that arcmin localization represents the goal accuracy for the next generation of multi-messenger astronomy observations. With reference to Figure 1.8, we assume the spherical region at left, representing the GRB-LIGO/Virgo area of detection, to be distant $d \sim 40$ Mpc and containing ~ 50 candidate sources, as during August 2017. The angular extension of this region is $\alpha \sim \sqrt{30}$ deg. At right, in black, we suppose to have a spherical region at distance $d' \sim 200$ Mpc with angular extension still $\sim \alpha$. Suppose the galaxies to be distribuited uniformly through the sky. The number of candidate galaxies in this region is $50 \cdot d'^3/d^3 \sim 3200$. Depicted in red we have a second, smaller, region at distance $d' = 200$ Mpc with angular extension $\sim \beta$. In order to have this region containing 50 sources, just as supposed for the first one, the angular extension should be $\beta \sim \alpha \cdot d/d' \sim 10$ arcmin.

In addition to the new generation of gravitational interferometers, during the 2020s The Cherenkov Telescope Array (CTA)—a new ground facility for studying very high-energy electromagnetic sources—will come online. CTA will consist of two arrays of telescope covering the northern and southern hemisphere hence providing access to the entire sky. The telescopes will operate in the energy band comprising 20 GeV and 300 TeV and are expected to improve the sensitivity level of current instruments (MAGIC, HESS, VERITAS) by an order of magnitude at 1 TeV [10]. The CTA fast re-positioning capabilities (20 seconds) and the largely improved sensitivity as compared to Fermi-LAT makes the study of GRB high-energy afterglow radiation possible and appealing. Since CTA will have limited field of view while operating at GeV energies ($\sim 3-4.$ deg), in order to guarantee the telescopes afterglow observation, it is necessary to employ an instrument able to first provide source localization with accuracy

better than a few degrees.
What about the next generation of space observatories? The ESA L2 mission Athena is expected to launch in 2034 [12]. Athena will host the X-ray Integral Field Unit (X-IFU), a X-ray narrow-field X-ray instrument with unprecedented energy resolution, and the Wide Field Imager (WFI) [24][23]. Both these instruments will operate between ~ 0.1 and ~ 10 keV. THESEUS was proposed in response to ESA call for medium-size mission M5 within the Cosmic Vision Programme [17]. In March 2018, ESA selected THESEUS for a three-year assessment study. The main object of interest to THESEUS are early universe, high-redshift GRBs. In June 2021, ESA selected EnVision as the fifh medium-class mission of the Cosmic Vision. Many other proposal have been submitted to NASA and ESA. Much like THESEUS and Athena, these missions are based on single spacecraft hosting powerful, large-area detectors.

No present X-ray astronomy facility provides an all-sky monitor with arcmin or better localization capability to support the next generation of multi-messenger observations. To date, four active satellites host instruments able to provide GRB and X-ray transients localization: The NASA Fermi space observatory was launched during 2008. Fermi hosts the GBM and LAT experiments. GBM (Gamma-ray Burst Monitor) consists of 12 NaI and 2 BGO scintillation detectors. The NaI are sensitive in the energy ranges spanning a few keV to ~ 1 MeV, while the BGO scintillators are active in the energy range comprised between 150 keV and 30 MeV. The GBM field of view covers almost two-thirds of the celestial sphere, the whole portion of the sky not occluded by Earth. GBM achieves transient localization by algorithmic comparison of the number of counts observed by differently oriented scintillation detectors. GBM localization accuracy is in the order of magnitude of the degree. The Large Area Telescope (LAT) is an imaging gamma-ray detector able to detect photons with energy raging from 20 MeV to 300 GeV over a field of view covering $\sim 20\%$ of the sky. The LAT instrument is able to achieve arcmin localization accuracy, yet the LAT limited field of view will most often require a GBM trigger and a subsequent spacecraft slew for localization to be achieved at all.
Swift has been launched in 2004. It is equipped with BAT (Burst Alert Telescope), a coded mask instrument operating between 15 and 150 keV with a field of view around 1/6 of the full sky. Alongside BAT, XRT is a narrow field X-Ray telescope able to provide localization of weak sources with arcmin accuracy.
AGILE was launched in 2007. It hosts Super-AGILE, a coded mask instrument operating between 15 and 45 keV with ~ 3000 deg^2 or about 1 sterad field of view with localization accuracy of a few arcmin for bright sources [104].
INTEGRAL has been launched in 2002. IBIS camera operates between 10 and

150 keV and is able to achieve arcmin positioning accuracy through a coded mask. IBIS field of view is below 1000 squared degrees and below 100 squared degrees for the fully coded sections [119].

Moreover, events detected on instruments hosted by multiple satellites can be used jointly by the InterPlanetary Network (IPN) to achieve typical localization accuracy of a few degrees reaching arcmin in rare cases. All these instruments have worked for more then 10 years. The equipment and instruments aboard are ageing and it is unknown for how long they will still be serviceable.

Chapter 2

The HERMES Mission

The HERMES (High Energy Rapid Modular Ensemble of Satellites) project goal is to develop a constellation of nanosatellites to study GRBs and other high-energy transients [45]. In its final configuration HERMES aims to provide an all-sky monitor with arcmin localization capability, μs time resolution and keV to MeV energy coverage, allowing to:

1. Support the next generation of multi-messenger astronomy issuing reliable and timely GRB localization data.
2. Investigate the sub-ms time variability of GRB light-curves.
3. Investigate the nature of the GRB radiative mechanism providing systematic measurements of the GRB emission below 10 keV.

For a more in-depth discussion of these points, the reader can refer to Sec. 1.8, Sec. 1.4 and Sec. 1.3.

2.1 Gamma-ray burst localization

Since light travels at finite speed, the delay between the arrival time of a signal to detectors hosted by different spacecrafts can be used to measure the angle between the detectors normal and the incident radiation. Starting from this information, the location on the celestial sphere of the burst source can be estimated. This is the localization strategy adopted by HERMES and pioneered by the VELA satellite network [45].

Consider a network of $N > 3$ spacecrafts[1] in Earth orbit separated by a baseline

---

[1] With $N = 3$ or less spacecrafts, localizations are ill-defined or degenerate. For $N = 2$, candidate sources are located along an arc of the celestial sphere. For $N = 3$, localizations are constrained to two distant areas of the sky; these locations being the intersection of two arcs of the celestial sphere.

distance d, each spacecraft hosting a detector. The localization accuracy is approximated by the equation:

$$\sigma \sim c \cdot \frac{\sqrt{\sigma_{delay}^2 + c^{-2} \cdot \sigma_{\vec{r}}^2 + \sigma_t^2 + \sigma_{sys}^2}}{d \cdot \sqrt{N-1-2}} \tag{2.1}$$

where σ_{delay} is the error estimate associated to the signal delay between different units, σ_t is the uncertainty in absolute time, $\sigma_{\vec{r}}$ is the error on the units location, σ_{sys} is a systematic uncertainty and c is the speed of light. The HERMES spacecrafts will fly in Low Earth Orbit (LEO) where the use of miniaturized spaceborne GPS receivers is possible. GPS localization allows for temporal and positioning accuracy within a few hundreds nanoseconds and tens of meters respectively, $\sigma_t \sim 10\,\text{ns}$ and $c^{-1} \cdot \sigma_{\vec{r}} \sim c^{-1} \cdot 10\text{m} \sim 30\text{ns}$. Plugging these numbers in Eq. 2.1 and taking into account a baseline distance between different spacecrafts similar to the Earth radius $d \sim R \sim 7000$ km, one obtains:

$$\sigma \sim 2.4\,\text{deg} \cdot \frac{\sqrt{\sigma_{delay}^2 + \sigma_{sys}^2}}{\sqrt{N-3}} \tag{2.2}$$

where σ_{delay} and σ_{sys} are expressed in units of ms. This relation implies that 1. with six spacecrafts a localization accuracy < 1 degree is attainable provided the values of σ_{delay}^2 and σ_{sys}^2 to be smaller than about 1 ms; and 2. one hundred spacecrafts would allow to achieve arcmin localization accuracy assuming contributions to systematic and cross-correlation uncertainties of about 0.1 ms. The entity of σ_{delay} depends on multiple factors, both peculiar to the detector and the astrophysical signal, such as the detector temporal resolution, the background noise and the temporal structure of the transient. Using cross-correlation techniques and Monte Carlo simulations of GRB lightcurves as expected for HERMES and based on Fermi GBM observations, realistic estimates of $\sigma_{delay,s} \sim 1.54$ ms for short bursts and $\sigma_{delay,l} \sim 0.10$ ms for long bursts have been obtained [100].

Estimating the systematic uncertainties σ_{sys} is not straightforward. Contributions are expected from diverse phenomena such as detector angles and calibrations, signal reflection from Earth atmosphere and background modulation. In this regard, the experience gained from the in-orbit operation of a small fleet of prototype units will be crucial.

2.2 Networks of small satellites

The HERMES platform is designed as a decentralized, scalable network of miniaturized satellites. Redundancy between components makes decentralized systems generally more resistant to system failure than their monolithic counterparts. A decentralized system may even benefit from a moderate degree of failure: a failing node brings knowledge on its vulnerabilities from which the whole system can take advantage. Indeed, this interesting property does not require a failure to unfold. Consider for example the problem of gauging the uncertainties σ_{delay} and σ_{sys} discussed in the previous section. A few operative units can be used to study in detail the limiting factors on σ_{delay} and σ_{sys} from which subsequent units design can be improved. Building a large decentralized system from scratch is often a complex and expensive operation. In the space segment ESA Galileo makes for a suggestive example. Galileo is the long-awaited european civil global navigation satellite system [60]. In its final configuration, it will be composed by 26 units of relatively small spacecrafts ($\sim$ 700 kg) in medium Earth orbit plus a ground segment. The problem of building a global navigation satellite systems may be considered as solved in 2020. Still the development of Galileo took more then 20 years and a budget of 10 billion of euros.

A decentralized system which is scalable i.e, able to grow efficiently, may amortize its cost over time or even pay for its own growth. Scaling can be achieved either vertically, upgrading the nodes already in place, or horizontally, deploying new nodes to the existing network. It is easy to realize how in the past the problem of a scalable satellite system may have seemed insurmountable to most entities. On one hand the inherent difficulty of upgrading an already existent satellite made—and still makes—vertical scaling impossible. On the other hand, the overhead cost involved in node deployment (e.g., sitting a new satellite onboard a rocket and launch it to space) made horizontal scaling an option only for the richest national space agencies and defence departments.

In the past decade two events radically changed this scenario. First, strong competition in the space launch market lowered the launch costs per payload unit mass steeply. Second, miniaturized spacecraft standards allowed to reduce the payload minimum mass requirements. Today it is possible to develop a small satellite and launch it on a budget, sharing a rocket 'ticket' with many other projects. As a result many nanosatellite constellations have been planned or deployed already, as shown in Fig. 2.1.

The CubeSat Standard Launch prices have been brought down to the point that it is feasible for many universities to launch their own miniaturized spacecraft to space. Indeed it's no coincidence that the CubeSat nanosatellite standard was initially developed for educational purposes. A Cubesat spacecraft unit has a mass smaller than 1.33 kg and dimensions $10 \times 10 \times 10$ cm^3. The

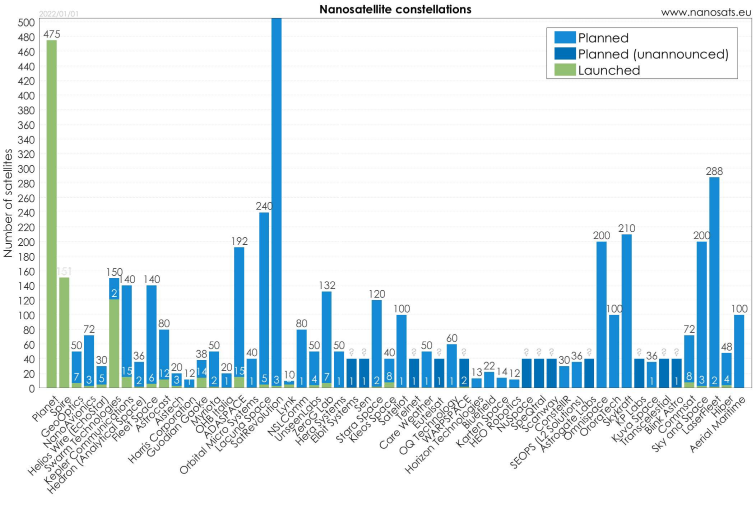


Figure 2.1: Number of spacecraft for different planned and deployed nanosatellite constellations. Credits: Erik Kulu, Nanosats Database, `https://www.nanosats.eu/`.

standardized Cubesat form factor allows stacking multiple units into a compact enclosure and spring-powered launcher tube called P-POD (Poly-Picosatellite Orbital Deployer). Through the P-POD, the interactions between the payload and the launching vehicle are minimized. Other than universities, the Cubesat standard has caught the attention of both research institutes and enterprises. In March 2018 NASA launched the first CubeSat to deep space: Mars Cube One (MarCO) [18]. MarCO is a flyby mission consisting of two nanospacecraft, intended to follow and assist NASA InSight Mars lander mission. The largest existent nanosatellite constellation is composed of 452 nanosatellites and is called Planet [105]. Planet is a commercial mission intended to gather and analyze image data to predict crop yields. Planet's "Dove" satellites are 3U CubeSats. Of course, not all scientific targets can be pursued with a miniaturized spacecraft. For example the observation of faint astrophysical sources requires large mirrors, infrared sources are hard to observe without complex active temperature control system and so on. Gamma-ray bursts represent a good scientific target since their high brightness allows detection also by small detectors.

2.3 The HERMES roadmap

In 2016 the Italian Ministry of University and Research (MIUR) and the Italian Space Agency (ASI) funded a first project to study, design, and prototype the detector and the Service Module of HERMES. The project – HERMES Technology Pathfinder - foresaw the deployment of three 3U CubeSat flying units by mid 2020. During 2018 the European Union granted 3.3 million euros for the development of three additional units in the framework of the Horizon 2020 programme for research and innovation. Finally, ASI approved and funded the participation to the project SpIRIT (Space Industry Responsive Intelligent Thermal), founded by the Australian Space Agency (ASA), and led by the University of Melbourne. SpIRIT will host an HERMES-like detector and S-band transmission systems. The HERMES-TP/SP (HERMES technological-scientific pathfinder) mini-constellation of six satellites plus SpIRIT is expected to be launched with a vector provided by ASI during 2022. HERMES-SP will be exploited to carry out a Phase A feasibility study for a future large constellation system (HERMES-FC) composed by tens of units.

The pathfinder program The main goals of the pathfinder program is to demonstrate the feasibility of the HERMES approach and in particular to:

- Develop miniaturized scientific instrumentation and technologies for detecting and studying GRBs and other high-energy transients.
- Demonstrate that accurate GRB localization can be achieved by measuring the delay time of arrival of the signal of different detectors hosted by miniaturized satellites.
- Demonstrate the applicability of disruptive technologies and COTS products to a space mission with challenging scientific goals.

These targets translate in two main high-level scientific goals:

- Detect GRBs with peak flux $> 0.5 - 1$ ph/s/cm^2 in the $50 - 300$ keV band.
- Detect ≥ 40 long GRBs and ≥ 8 short GRBs simultaneously in at least 3 units during a nominal phase of 2 years and to be able to assess their position through the analysis of the delay time in the signal arrival time on different detectors.

The first scientific constraint requires a detector with collecting area ≥ 50 cm^2 and background level ≤ 1.5 counts/s/cm^2 [48] [41]. The second scientific goal results in constraints for both the payload and the service module. Indeed, a large detector field of view (FOV) > 1 steradians is needed for the simultaneous observation of the same burst over multiple detectors. Additionally each spacecrafts must be able to slew with an attitude determination accuracy better than a fraction of the FOV. To verify the feasibility of these targets, a detailed mission analysis has been performed [37].

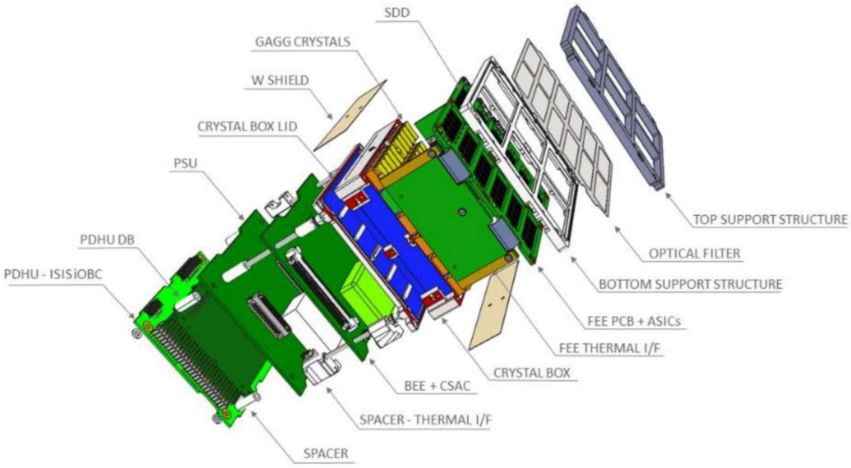


Figure 2.2: An exploded view of the HERMES detector.

HERMES Full Constellation The HERMES-FC scientific goals are:

- Cover each location of the celestial sphere with 20-30 detectors with maximum off-axis angle equal 60 degrees and at any given time.
- Achieve sensitivity to weak GRBs with flux ~ 0.2 ph/s/cm^2 in the energy band $50-300$ keV allowing the observation of events similar to GRB170817 at distance $120 - 150$ Mpc.
- Achieve localization accuracy better than 15 arcmin for long bursts and 1 degree for short bursts.
- Disseminate transient position data within minutes from the trigger time.
- Achieve better than μs photon timing accuracy.
- Deploy enough units to achieve a total collective area better of the order of a few m^2.

2.4 The HERMES payload

Siswich detectors The HERMES detector is based on the "siswich" architecture [48][41][78]. In a siswich detector, thin silicon detectors are coupled to a scintillator crystal. Soft X-ray photons are absorbed in the silicon bulk. Hard X-ray and gamma photons have enough energy to cross from face to face the HERMES 450 μm thick silicon detectors, hence to reach the scintillator where they are absorbed. In the scintillator the energetic photons are converted to visible light which is then collected by the silicon detectors. The discrimination between the two signals is achieved using a segmented design: a single scintillator crystal is coupled to two different silicon detectors, so that events detected by a single silicon detector will be most likely associated with soft X-Rays, while events detected simultaneously in multiple adjacent detectors will be associated to the optical light produced in the scintillator by hard-X and gamma photons. The scintillation crystal selected for HERMES is GAGG:Ce ($Gd_3Al_2Ga_3O_{12}$:Ce, Cerium-doped Gadolinium Aluminium Gallium Garnet). It is a recently developed material presenting a wide array of desirable features such as mechanical robustness, extremely low intrinsic background and no hygroscopy [121]. Most importantly, GAGG:Ce displays high light yield of $\sim 50 \cdot 10^3$ ph/keV and short decay times $\sim$ 100 ns [121] [122]. In siswich detectors selecting a scintillator with high light yield is particularly important since it enables lower energy thresholds for high-energy detections. Fast decay times on the other hand are important to achieve high photon timing precision. GAGG:Ce has a troublesome property in its particularly intense and long-lasting phosphorescence. In Sec. 3 a study of this effect and its expected impact on the HERMES detector is described in detail.

The silicon detector technology employed by HERMES is Silicon Drift Detectors (SDD) [49]. The main feature of SDDs is that the photon collection area is decoupled from the region where charges are collected, the anode, which can then be built extremely small. Since the anode size determines noise levels, extremely low noise level can be achieved with SDDs. When compared to alternatives (e.g., silicon photomultipliers), SDDs are relatively slow in response since the charge produced by light or by X-rays must 'drift' through the device towards the collecting anode. Faster response times can be achieved reducing the SDD size or increasing the intensity of the electric field in the depletion volume. The HERMES SDDs are state-of-the-art detectors developed by the italian ReDSoX collaboration. Within ReDSoX cooperate the designing and manufacturing structures and the expertises of INFN-Trieste and Fondazione Bruno Kessler (FBK, Trento).

Front-end and back-end electronics The SDDs anode signal is read out and processed by the detector front-end electronic (FEE). The HERMES FEE design is called LYRA and stems from VEGA, an application specific integrated circuit (ASIC) developed in the context of the LOFT mission Phase-A assessment study [14]. The HERMES detectors hosts four independent LYRA ASICs. Each LYRA ASIC is divided in two stages called LYRA-FE and LYRA-BE. The LYRA-FE is the first stage of the LYRA ASIC and includes 32 integrated circuits (ICs) with functions of preamplifier, signal shaper and transmitter. In order to keep the input capacitances (hence series noise) as low as possible, the LYRA-FE ICs are installed in close proximity to the SDDs, one IC per SDD cell. The LYRA second stage is composed of a single chip and is placed in a farther position with respect to the SDDs. Indeed, closeness to SDDs is not critical for the LYRA-BE, the signal having been amplified already at the first stage. The LYRA-BE completes the signal processing chain providing second shaping stage, discriminators and peak and hold. Moreover, it is responsible for the ASIC configuration and for the multiplexing of the processed signal from the 32 LYRA-FE input channels.

The signals from the four LYRA-BE chips is routed towards the HERMES back-end electronics (BEE). The BEE hosts a FPGA, four ADCs, a miniaturized atomic clock as well as a number of ancillary components (e.g., pulse generators, temperature sensors..). At back-end level the analog signal is digitized and "tagged" with data such as the photon arrival time and the triggered SDD cell addresses. The photon event information is stored and curated in two buffers, the housekeeping and event buffers, physically hosted by the BEE. The BEE is also responsible for executing the Payload Data Handling Unit (PDHU) on-board computer commands onto the power supply unit (PSU) interface.

The Payload Data Handling Unit The HERMES-TP/SP Payload Data Handling Unit (PDHU) is built around the Innovative Solutions in Space (ISIS) On-Board Computer (iOBC) [55] [41]. The iOBC is a flight-proven processing units hosting a 400 MHz ARM9 processor alongside a mass memory and numerous standardized interfaces. A custom-made 'daughter board' has been developed and implemented to provide all the payload-bus electrical interfaces as well as the PDHU internal interface with the other payload subsystem. The PDHU manages the PSU and multiple analog payload temperature sensors. Moreover it runs the software to manage the spacecraft operation modes; generate, compress and filter photon list; prepare the scientific and housekeeping data and more. Most importantly for this work, the PDHU runs the trigger algorithms for GRB detection.

The HERMES demonstration model In numbers, the HERMES-TP/SP detector is composed of:

1. 60 scintillation crystals with dimensions $14.5 \times 6.94 \times 15.0$ mm^3.

2. 120 SDD with surface area 7.44×6.05 mm and thickness $450\,\mu$m.

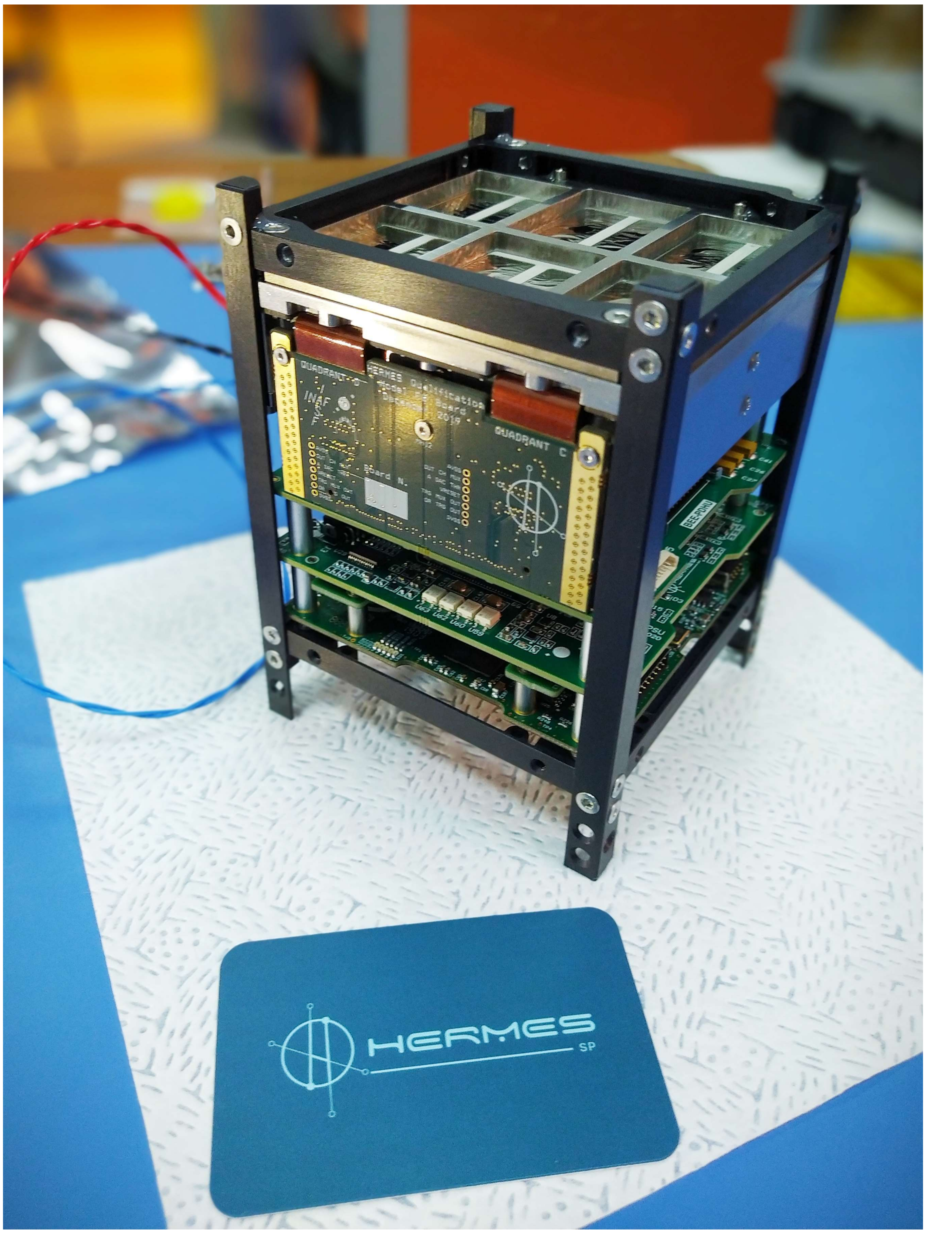


Figure 2.3: The HERMES payload demonstration model. From top to bottom the detector assembly, the BEE board and the PDHU's main and daughter boards are distinguishable.

Parameter	Value
Peak effective area	52 cm^2
Field of view	3.2 sr FWHM
Lower energy threshold	$\leq$ 3 keV
Energy resolution X-mode (6 keV)	$\leq$ 800 eV FWHM
Energy resolution S-mode (60 keV)	$\leq$ 5 keV FWHM
Time resolution	320 ns
Time accuracy	181 ns
Payload mass	1588 g
Payload power consumption	4791 mW

Table 2.1: Technical and performance budget of the HERMES-TP/SP [41].

3. A tungsten-shielded crystal box and a detector support structure mounted on top of the sensitive plane.

4. A front-end electronics board. The FEE board hosts 12 coupled SDD-scintillator arrays, each composed by 5 crystals and 10 SDD in a 5 × 2 configuration; 120 LYRA-FE ASIC dies, one per SDD channel; discrete electronic components; and electrical connectors providing the electrical interfaces with the BEE board;

A demonstration model (DM) of the HERMES payload complete with the detector assembly, the FEE, the BEE and the PDHU, has been assembled and tested to validate the payload mechanical and thermal design as well as the FEE performances [41], see Fig. 2.3. The results of these tests are summarized in Tab. 2.1.

2.5 The HERMES background simulations

The first six HERMES spaceccrafts will be launched as secondary payloads in a near-equatorial, Low-Earth Orbit (LEO). The exact orbital parameter are still unknown but inclinations smaller than 10 degrees and altitudes around 550 km are expected. Once in orbit, the nanosatellites will be subject to high-energy radiation fluxes mostly composed of primary, secondary and trapped charged particles as well as diffuse photon background. All these sources may contribute to the HERMES detector scientific background count-rate. Geant-4 Monte Carlo simulations have been performed to estimate the mean rate of background events as well as to estimate the contributions from each component [33]. These simulations take into account a mass model of the spacecraft geometry and assembly providing a faithful description of the detector and a simplified representation of the service module.

Background sources The HERMES detector will operate in LEO, near-equatorial orbit and no operations will be held during transits through the inner Van Allen Radiation Belt i.e., the South Atlantic Anomaly (SAA). Hence the major sources to the HERMES scientific background are expected to be primary and secondary cosmic rays particles and photons from the cosmic X-ray background (CXB) and the Earth albedo.
Primary cosmic ray particles. The primary cosmic rays full spectrum at a given location of the magnetosphere and at a given phase of the solar cycle is determined by the particle nature, solar modulation effects and the shielding effect of the local geomagnetic field [32][83]:

$$F(E) = F_U(E + Ze\phi)\, F_M(E, M, Z, \phi)\, C(R, h, \theta_M)$$

where M and Ze are the mass and charge of the cosmic ray particle, E is its kinetic energy, ϕ is a solar modulation factor, h is the orbit altitude and θ_M is the geomagnetic latitude. The function F_M comprises effects of the solar modulation on the cosmic ray particles. The term $C(R, h, \theta_M)$ is the geomagnetic cut-off function, a quantity which describes the local rate of charged particles transmission through the magnetosphere. This factor depends on the geomagnetic position and the cosmic particle magnetic rigidity $R = pc/Ze$, where with p we note the cosmic particle momentum. Most importantly $F(E)$ is determined by the primary cosmic ray spectrum $F_U(R)$. Empiric or semi-empiric parametrizations of $F_U(R)$ for protons, alpha-particles, neutrons and leptons exist in literature [101][85].
Secondary particles. The measurements of the experiment AMS on-board the International Space Station, show that the secondary cosmic rays component is composed of both short-lived and long-lived particle populations originating from the regions near the geomagnetic pole [13]. A secondary, quasi-trapped proton radiation component exists in low-altitude, near-equatorial orbits. These particles are produced in the impacts of cosmic rays on the atmosphere. In literature they are sometimes referred to as the “splash” or “re-entrant” component. In literature, the low-altitude, secondary proton and lepton spectrum has been modelled through broken power-laws [32] [83]. There is presently no standard model of the secondary neutrons or atmospheric albedo neutrons. In the simulation of Campana et al. [33], Monte Carlo radiation transport software packages have been used to account for this component.
Photon background. A parametrization of the cosmic X-ray and gamma-ray dif-

Energy band [keV]:	(3, 25)	(25, 50)	(50, 100)	(100, 300)	(300, 1000)
Count-rate [s^{-1}]:	544.4	67.1	41.0	35.1	12.1

Table 2.2: The mean HERMES scientific background count-rate in different energy bands [33].

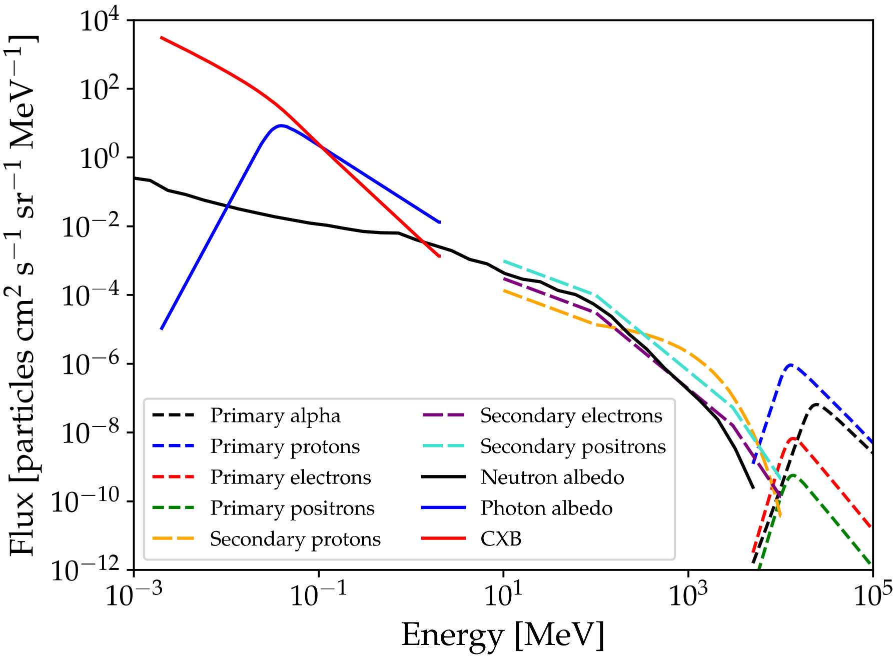


Figure 2.4: Contribution from different sources to the HERMES scientific background rate [33].

fuse background exists for energies between 3 keV and 100 GeV, based on the measurements of HEAO-1 A4 and COMPTEL [54]. Energetic photons may also be produced in the impacts of cosmic rays with the atmosphere. This albedo radiation component has been measured by SWIFT/BAT and parametrized [15]. *Nuclear activation.* Scattering high-energy proton radiation may exceed the Coulomb barrier of a target, inducing a level or radioactivity into the target material. Contributions from nuclear activation have not been taken into account by the simulations of Campana et al. [33].

The expected contribution of each background source to the HERMES scientific background mean rate is plotted in Fig. 2.4 [33]. Up to 500 keV the most intense source is the photon component while at higher energies the particle background becomes dominant. The scientific background rate over the low and high energy channels of the HERMES detector are represented in 2.5.

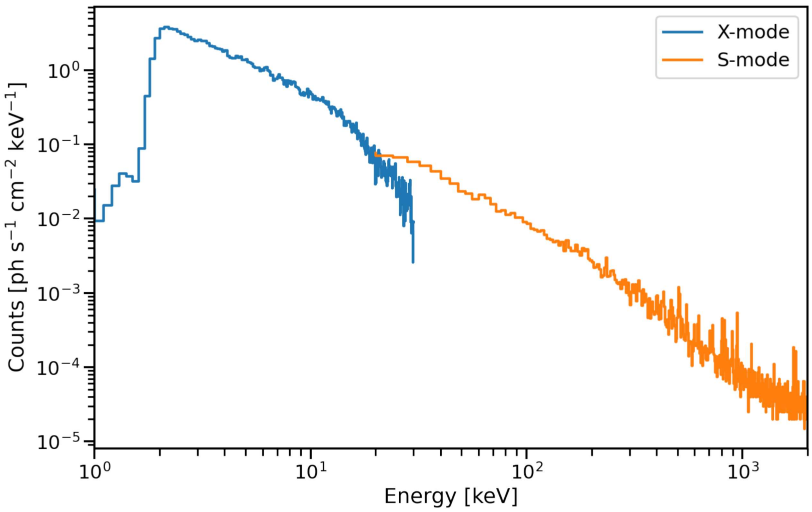


Figure 2.5: The HERMES scientific background [33]. The step line labelled 'X-mode' represents the expected background differential rate of those events which do not reach scintillators. On the other hand, the step line labelled 'S-mode' represents the differential rates for those events which have enough energy to reach the scintillation crystals. See Sec. 2.4 for more details on the HERMES detector segmented design.

Background count rate in various energy bands are reported in Tab. 2.2.

Background reduction strategies The HERMES segmented detector design may allow to discriminate particle events from photons with good accuracy. Impacting on the HERMES detector, energetic particles are expected to intersect multiple scintillation crystals simultaneously. These unwanted events may be suppressed to some degree on the basis of them resulting in a simultaneous trigger from many different SDDs. Moreover, for background particle events large energy deposits in the crystal bulk are expected, eventually exceeding (saturating) the 2 MeV detector sensitivity band limit. The impact of a baseline filtering strategy has been evaluated in the Monte Carlo simulations of Campana et al. [33]. The filter conditions were 1. an energy deposit in a single crystal below 2 MeV; and 2. simultaneous triggers over a maximum of 5 SDD channels. This logic was able to correctly reject up to 94% of the background particle events while keeping rejection rates under 0.1% for photon events.

Part II

Two studies concerning the development of the HERMES-TP/SP space mission.

Chapter 3

Space applications of GAGG:Ce scintillators: a study of afterglow emission by proton irradiation

The content of this chapter has been published in "Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms" *during 2022 [40]. An integral version of that article, adjusted to fit the context of this thesis, is presented hereafter.*

HERMES detectors are designed around the so-called "siswich" concept [31] in which silicon detectors play the double role of sensors for the scintillation light emitted by suitable scintillator crystals and of independent detectors for low energy X-rays. The scintillator selected for use on HERMES units is GAGG:Ce ($Gd_3Al_2Ga_3O_{12}$:Ce, Cerium-doped Gadolinium Aluminium Gallium Garnet). It is a promising scintillation crystal displaying a wide array of appealing features for space applications: very high light-yield, fast-decay times, very low intrinsic background and mechanical robustness.
However, GAGG:Ce is still a recently developed scintillator [63] and, as a consequence, literature is lacking on points crucial to its applicability in space. For example, GAGG:Ce is characterized by unusually intense and long-lasting afterglow emission [76], a slow phosphorescence component in scintillation light. Afterglow emission is a source of background noise and is induced by the exposure of GAGG:Ce crystals to electromagnetic and particle radiation. Hence, in space applications, an effective degradation of the detector energy resolution

should be expected as a result of the phosphorescence induced by the interaction of the energetic particles in the near-Earth radiation environment with the GAGG:Ce scintillators; the extent and dynamics of such phenomena depending on both the host spacecraft orbit and the crystal intrinsic properties. Besides degradation of the energy resolution, the current induced in the SDD sensors by the background light due to the afterglow emission may become too large, impairing the functionality of the Front-End Electronics (FEE). To tackle this last concern we conducted an irradiation campaign at the Trento Proton Therapy Center (TPTC) in which a GAGG:Ce sample was irradiated with 70 MeV protons. The choice of particle specie, energy and fluences was driven by the need to simulate the nature of the radiation environment of near-equatorial LEO orbits and the constraints of the cyclotron particle accelerator available at TPTC and of our equipment.

This chapter is arranged in three parts. In the first part we describe the experiment set-up and timeline. In the second part we discuss GAGG:Ce phosphorescence, introduce our model of the afterglow emission, examine the impact of activation on our observations and outline the fit procedure and results. The afterglow model development is discussed in detail in Appendix A. Finally we make use of the afterglow model, supported by the AE9/AP9 trapped radiation belt models [52], to estimate the impact of the afterglow emission resulting from the orbital radiation environment on the performance of the instrument.

3.1 Experiment Outline

A GAGG:Ce scintillation crystal—dimensions $3 \times 1 \times 1$ cm^3—was housed in a lightproof metal case hosting two PMT detectors. The large faces of the crystal were wrapped in thin white teflon to minimize scintillation light dispersion, while both the unobstructed small faces were coupled to 5 cm long quartz light guides by means of optical grade silicone grease. The same coupling technique was used to interface the light guides to the photocatodes of the Hamamatsu R4125 PMTs. The anode signal of one of the two PMTs, labelled PMT1, was measured by a Keithley 6487 picoammeter. The signals from the last dynode of both PMTs were amplified by a factor of 20, brought to a discriminator with thresholds set to -50 mV and redirected to a counter and a multi-channel digitizer for waveform acquisition, along with the anode signal of PMT2. Waveform acquisition could be triggered by either a single PMT or both, the latter through a programmable logic unit. Irradiation took place at the Trento Proton Therapy Centre (TPTC), Trento, Italy. At this facility a cyclotron (IBA, *Proteus 235*) serves two medical treatment rooms and an experimental area, accelerating protons to a kinetic energy selectable in the range 70–228 MeV. The proton

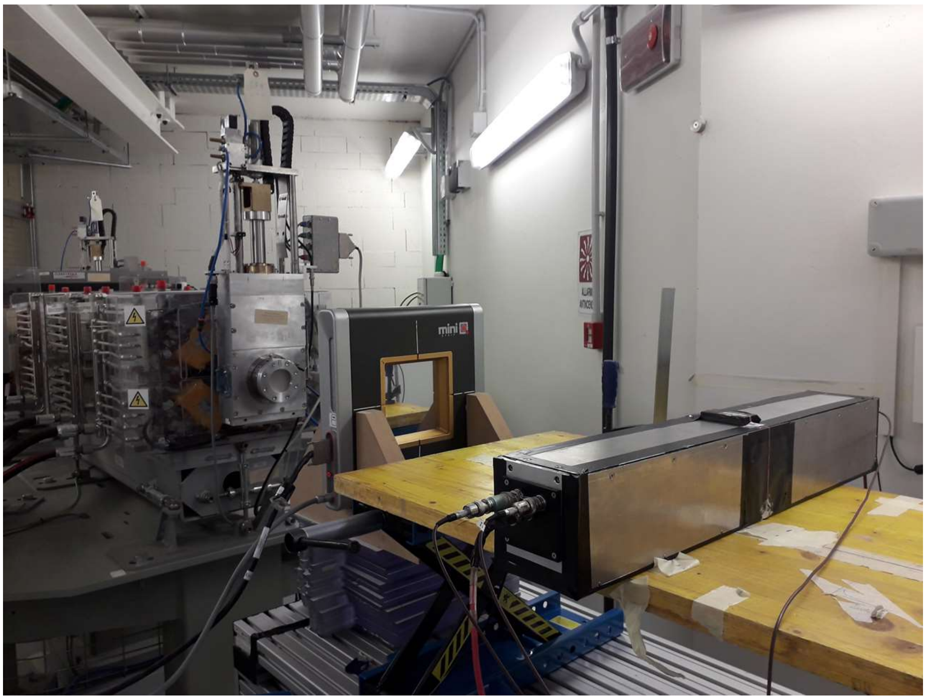


Figure 3.1: A shot from the Trento irradiation campaign. The chassis containing the GAGG:Ce sample and the photomultipliers is aligned to the beam nozzle.

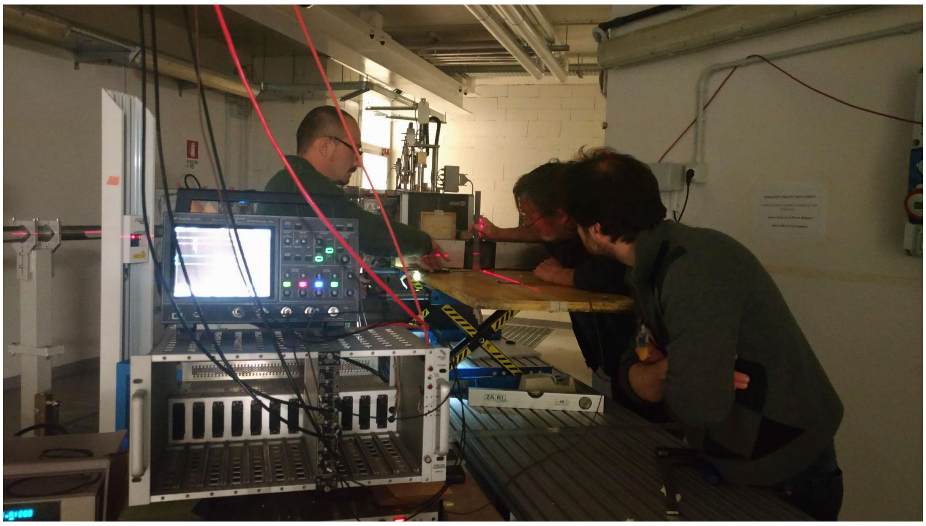


Figure 3.2: A shot from the Trento irradiation campaign. Giovanni, Riccardo and Fabio work at the setup alignment.

Figure 3.3: Giovanni, Fabio, Benedetto and Francesco working at the experiment set-up. HERMES should probably hire a better photographer.

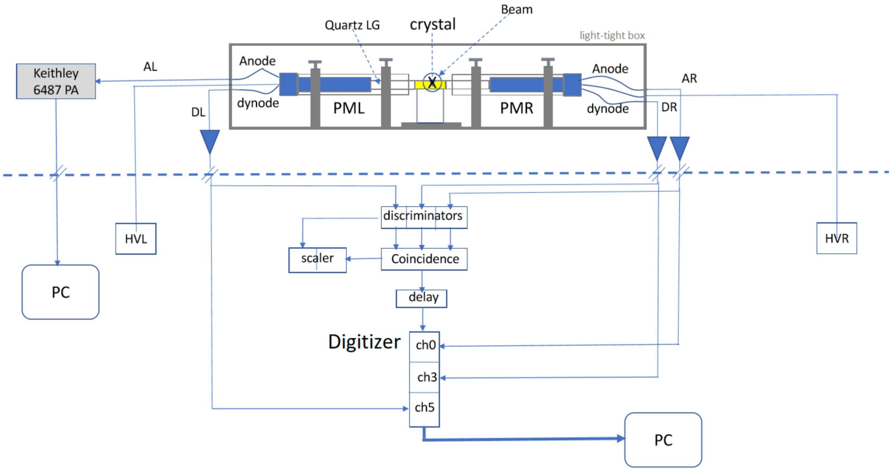


Figure 3.4: A schematic view of the experimental setup.

beam has a Gaussian profile with σ = 6.9 mm at 70 MeV energy [106]. Two operation regimes are available: the so-called ‘dark’ mode with very low beam intensity ($<$ 10 protons s^{-1}), and the normal mode with high beam intensity. In normal mode, the extraction current is adjustable in the range 1–320 nA and is modulated by a 50% duty-cycle square wave with a period of 100 ms. We used both modes for different purposes.

	Ext. Current [nA]	Irr. Duration [s]	Intercepted protons [p]	Dose [rad]	E.O.A.[1]
Irradiation 1 ●	1	90	$(1.19 \pm 0.23) \times 10^{8}$	6	1y EQ
Irradiation 2 ●	1	90	$(1.20 \pm 0.23) \times 10^{8}$	6	2y EQ
Irradiation 3 ●	1	270	$(3.56 \pm 0.69) \times 10^{8}$	19	5y EQ
Irradiation 4 ●	10	100	$(1.37 \pm 0.26) \times 10^{9}$	73	10y EQ
Irradiation 5 ●	10	144	$(1.95 \pm 0.37) \times 10^{9}$	104	2y SSO
Irradiation 6 ●	100	115	$(1.52 \pm 0.29) \times 10^{10}$	807	10y SSO

Table 3.1: Detailed table of the irradiation runs. Each row represents an irradiation step. Runs are identified by current log start time and irradiation steps are color coded as in the article body. The number of protons intercepted by the target and the total absorbed dose were estimated from a GEANT4 simulation of the experimental setup. [1]: Equivalent Orbital Age. The reported values represent the number of consecutive years of orbits needed to the tested sample to achieve the same proton dose accumulated over different steps of the irradiation campaign. Computed according to AP8MIN models of the proton radiation environment (kinetic energies $>$ 0.1 MeV) of an orbit with altitude 550 km and inclination 10 (EQ) or 98 (SSO) degrees.

For our tests we selected a proton energy value of 70 MeV. This choice was driven by the following reasons. First of all, the need of a pure proton beam with an energy representative of the trapped proton spectrum characteristic of low-Earth, nearly equatorial orbit. Most of the trapped protons in the regions of the inner Van Allen radiation belt near the equator have energies which spans tenths to hundreds MeV (see Fig. 3.10). Moreover, 70 MeV is the smallest energy attainable without further degradation at TPTC. Finally, simulations showed that 70 MeV protons would most often release their whole kinetic energy within the GAGG:Ce crystal bulk.

Regarding the irradiation duration and the cyclotron extraction current selected for each irradiation step, hence the dose to be irradiated, we chose parameters such that the radiation dose accumulated by the GAGG:Ce target would match the end-of-life levels expected from different orbital operation scenarios. We remark on the 1 nA beam minimum extraction current resulting in irradiation flux condition exceeding the levels expected in orbit by at least two orders of magnitudes (compare Tab. 3.1, Fig. 3.9 and Fig. 3.10). Working with fluxes outside the expected conditions, possible saturation effects in afterglow

emission may occur. Such phenomenon would result in an underestimate of the phosphorescence to be expected in orbit, where particle radiation fluxes are generally lower. This eventuality is addressed, at least in an approximate way, by our model of the crystal luminescence.
For each irradiation step the schedule of operations involved:

1. In dark mode: collection of $\sim$ 5000 PMT2 anode current waveforms, triggered by the coincidence of the signals from both PMTs operated at 1100 and 1350 V (proton events).

2. In normal mode: irradiation of the crystal with the required beam intensity and exposure duration. During this step the PMTs were turned off.

3. Beam off: 60 seconds after the end of each irradiation, measured by a chronometer, the anode current of PMT1, now operated at 1500 V, was sampled each second for about 15 minutes (800 seconds), see Fig. 3.5. At the same time, count data were acquired both for individual PMTs last-dynode signal and their coincidence.

Proton waveforms were acquired after the last irradiation step. Afterwards, monitoring of the anode current was started once more and continued for $\sim$ 11 hours during the night. The scintillator temperature was monitored by means of a thermocouple and ranged between $21.0 \pm 0.5\,^{\circ}$C. Specifications of each irradiation step are reported in detail in Tab. 3.1. The number of intercepted beam protons was estimated through GEANT4 simulation of a GAGG:Ce crystal irradiation experiment. In this simulation, a 70 MeV proton beam with a 2D Gaussian profile of width $\sigma = 6.9$ mm was modelled, according to the beam characteristics at the isocenter point [106]. The crystal was placed at the beam center and possible positioning errors were taken into account. The total number of simulated protons was selected according to the flux measured by the beam monitor thus the computed intercepted fluence, total energy deposit and resulting dose.

3.2 Afterglow emission models

Long-lived afterglow emission in scintillators is attributed to the existence of intrinsic or impurity defects within the crystal lattice. Some of the charge carriers (electrons or holes) liberated by the ionizing radiation can be trapped at defect sites into metastable states. At later times, charge carriers escape these sites by different processes (e.g. to the conduction band by thermal energy absorption

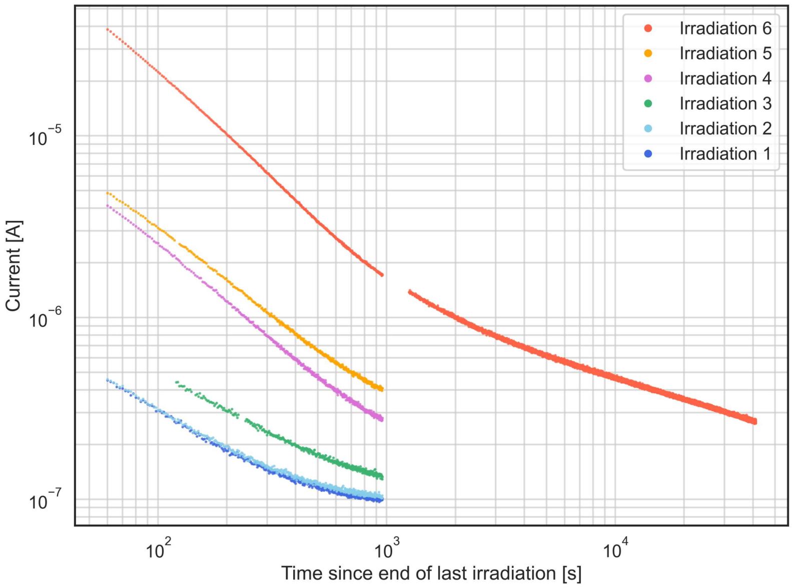


Figure 3.5: PMT1 anode current induced by the afterglow resulting from irradiation with 70 MeV protons of a GAGG:Ce sample, versus time elapsed from the irradiation end. Color coding as in Tab. 3.1. The first 60 s of data in the third measurement (green) have been excluded due to an improper initialization of the PMTs power supply.

[66] or to nearby recombination centers by direct or thermally assisted tunneling [61]). Ultimately all of the charge carriers recombine, mainly through radiative paths in good scintillators, giving rise to luminescence. Different scintillating materials display different afterglow characteristics. Although mitigation techniques (e.g. Mg-codoping) proved successful [76], GAGG:Ce is known for its intense afterglow emission which may last up to several days [122].
In the existing literature GAGG:Ce afterglow emission has been reported decaying as a power-law of time [76][122]. Although a power law can adequately describe the decaying emission within short times after the end of an exposure, we found such decay profile unable to fit the whole duration of our measurements. Deviations from power law behaviour are evident already by inspection of Fig. 3.5. In Fig. 3.6 we report on the power-law fit of one measure in our dataset.

In this work we seek a different approach to model the proton induced afterglow emission which allow us to accurately fit the data and to make predictions on the amount of afterglow to be expected in space applications. At the core of this representation lies the assumption that afterglow arises from the delayed re-

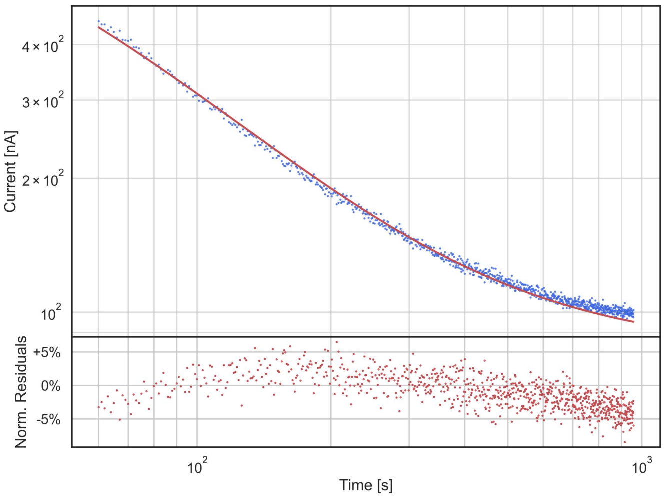


Figure 3.6: Fit to inverse power-law plus a background constant of the PMT anode current from Irradiation 1, cf. Tab. 3.1. Best fit power-law index is 1.54. Time-dependent patterns in residuals are apparent.

combination of charge carriers kept in metastable states by charge 'traps', which can be classified in species characterized by a mean *capture rate* and a mean *lifetime*. We expect that at the smallest time scales (seconds) accessible through our data the characteristic times of de-excitation of the metastable states do not form a continuum, so we assume there exists a discrete set of trap species. The dimension of this set is not known beforehand and has to be determined by data analysis. At any given time the number of occupied traps can be determined balancing the rate at which new traps get occupied during irradiation versus the rate at which charges free themselves from traps to recombine. This mechanism is described through a system of differential equations of the form:

$$\frac{d}{dt}N_i(t) = n_i\,\phi - \frac{N_i(t)}{\tau_i} \tag{3.1}$$

with n_i and τ_i being the characteristic capture rate and mean lifetime of an i-labelled trap species. These equations can be analytically solved for arbitrary irradiation flux profiles ϕ, hence for the irradiation timeline of our campaign. This simple model is able to describe observations following the first four, low-dose, irradiation steps but fails when applied to the whole dataset. In order to support the observation which followed the fifth and sixth irradiation steps we modified the model to support linear variations in the traps capture rates.
A mathematical derivation of the model is given in A.

3.3 Crystal activation

The energies of the protons, both in the near-Earth radiation environment and in the irradiation campaign, are high enough to induce a level of activation in GAGG:Ce crystals. In these settings, activation poses different challenges. In orbit, crystal activation is expected to result in a decay spectrum from the unstable nuclides that will interfere with the observation of astronomical gamma-ray sources. From the point of view of the irradiation tests, the scintillation accompanying the decays is superposed to the afterglow emission and the two contributions can not be disentangled from our observations.

The most direct (yet limited) information about the activation occurred during our irradiation campaign come from coincidence count data. A coincidence event was registered whenever current from both PMTs exceeded a pre-set signal threshold. Such events are more likely to result from the sudden, bright scintillation light produced in nuclear decays than they are from the incoherent yet persistent afterglow emission.

The activation effects of proton irradiation on GAGG:Ce has been investigated already for energy of 70 MeV [122] or higher [99]. The work for Yoneyama et al., in particular, identified a number of lines in the GAGG:Ce activation spectrum with energies up to 1038 keV resulting from a 10 krad, 70 MeV proton irradiation.

Using this information we estimated the contribution of scintillation to the measured currents at the beginning of each measurement (60 s after the end of the corresponding irradiation step) assuming that on average each detected decay deposited 1 MeV of ionization energy in the crystal. The decay rates needed to calculate the scintillation component were estimated from the coincidence count curves by fitting them to a simplified model that considers one of the decay modes to be dominant. Taking into account our estimate of light transport efficiency, PMT gain and the picoammeter integration time, analysis of count data from our campaign lead to an estimated contribution to the PMT anode current from activation at the beginning of the measurements following the first irradiation equal to 25 nA, or 6.7% of the current value observed at that time. The same reasoning, applied to the data observed after the last, high-flux step of irradiation, lead to a slightly larger ratio estimate of 8.2%. Far from being conclusive, this analysis shows that the contribution from activation to the measured PMT anode current was not negligible.

3.4 Fitting data to the afterglow emission model

Due to a significant presence of scintillation light in the measurements the application of the afterglow models we developed is not justified because it entails a precise physical meaning to the fit parameters. Nonetheless, being interested in a conservative estimation of the afterglow emission during the mission, the additional component due to crystal activation can be considered as yet another source of overestimation and we can still apply our modelling in a purely empirical way to reproduce the light curves, provided we do not attach any precise meaning to the model parameters. For this reason we choose to assign new names to the parameters, according to the following scheme: instead of trap species we talk about emission modes, the capture rates become emission amplitudes, and the variation of the capture rates turn into the variation of the emission amplitudes.

The PMT used in our measurements intercepted only a fraction of the photon flux emitted by the crystal. The anode current $I(t)$, as measured by the picoammeter, can be expressed as:

$$I(t) = eG_V f_q \phi_q(t) \tag{3.2}$$

where e is the elementary charge, G_V is the gain of the PMT at operational voltage V, f_q is the fraction of photons that extract an electron from the PMT photocathode, and $\phi_q(t)$ is the photon flux from Eq. A.8.
Through PMT calibration we found a gain value $G_{1500} = (1.04 \pm 0.04) \cdot 10^6$. The quantity f_q depends on both the wavelength dependence of the light transport efficiency from the crystal to the PMT and the PMT spectral response [64]. It can be expressed as the average of the product of these efficiencies, weighted by the GAGG:Ce emission spectrum [122]:

$$f_q = \langle \epsilon_{LG}(\lambda) \epsilon_{PMT}(\lambda) \rangle_{GAGG:Ce}$$

In our analysis we assumed that ϵ_{LG} had a negligible variation over the relevant wavelength interval. Therefore it is possible to decouple contributions from the transport efficiency and the PMT response:

$$f_q = \epsilon_{LG} \epsilon_{PMT} = \epsilon_{LG} \langle \epsilon_{PMT}(\lambda) \rangle_{GAGG:Ce} \ .$$

The average PMT quantum efficiency was calculated using data in the literature to be $\epsilon_{PMT} = 0.06 \pm 0.01$ [64].
The fraction of light emitted by the crystal reaching the PMT photocathode

was obtained comparing the photons collected by the PMT with the scintillation light emission expected from irradiation of GAGG:Ce with Cs-137:

$$\epsilon_{LG} = \frac{Q_\gamma}{e\epsilon_{PMT} G_V E_\gamma} \Big/ LY$$

where Q_γ is the integral of the current pulses corresponding to the gamma ray photons of energy E_γ, and $LY = 53 \pm 3$ photons/keV is the light-yield of the crystal measured by its manufacturer [2]. For the Cs-137 662 keV line we found that PMT1 collected 15 ± 4 photons/keV, leading to an estimate of the light transport efficiency $\epsilon_{LG} = 0.29 \pm 0.09$.

To fit the model to the observations we first estimated the uncertainties associated to each measurement, which we assumed to be well represented by the fluctuations in the data. This procedure has been accomplished in two steps. In the first we smoothed the light curves by fitting them with cubic splines and then recovered the fluctuation by subtracting the smoothed curves from the observations. In the second step we determined how the fluctuations changed as a function of the intensity of the light reaching the PMT. We performed a calculation of the moving RMS of the fluctuations, with different window sizes to determine its robustness. The RMS values with a window size of 8 samples are plotted in Fig. 3.7 as a function of the smoothed PMT anode current, to which we subtracted our best estimate of the dark current (83 nA). We fitted the RMS values to a power law and a residual constant term:

$$RMS = A + B \cdot I^C$$

For the measurements immediately following each irradiation step we found that the RMS changed in a way approximately proportional to the square root of the light flux, while in the night measurement the decaying emission was followed linearly.

The fit of the observation dataset to the afterglow models was performed in an iterative way, progressively adding the data of the irradiation steps to see which changes in the model were needed to adequately describe the observations. We found that a model with constant emission amplitudes is able to reproduce the dataset only up to the fouth irradiation step. We looked for evidences of radiation damage analyzing the scintillators light-output before and after the irradiations. In fact, radiation damage is expected to affect light-output in scintillators by reducing trasmittance through creation of color centers absorbing scintillation light [66]. For GAGG:Ce the color centres absorption band appears

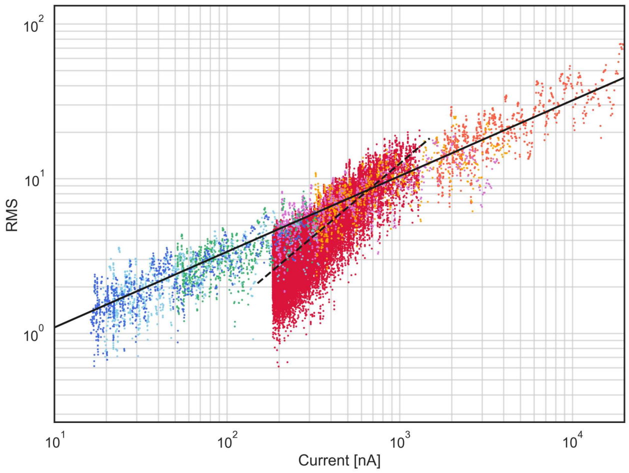


Figure 3.7: RMS values computed from the residuals of cubic spline fits of the light curves. The set in red behaving differently corresponds to the night measurement of the afterglow following the last irradiation. The RMS values were modelled as $RMS = A + B \cdot I^C$. Best fit to this law is shown in black solid line for the measurements immediately following the irradiation steps, and refers to the values: $A = 0$, $B = 0.35$, and $C = 0.49$. The black dashed line represents the best fit of the night RMS dataset with parameters $A = 0$, $B = 0.019$, and $C = 0.94$. See Tab. 3.1 for color legend.

to be in the UV range [16]. Given the small overlap between the absorption spectrum and the afterglow emission spectrum, it is difficult however to appreciate evidences for this phenomenon from our observations, hence to unambiguously attribute a variation in the emission amplitudes to radiation damage. From the analysis of the proton event waveforms we were unable to infer any significant ($> 5\%$) reduction in light-output. Before addressing the complete observation set we performed a fit of the measurements taken following the last irradiation, both immediately after and during the night, to the model with constant emission amplitudes, taking into account all the irradiation steps. We repeated the fit several times to determine the number of emission modes, and we found an optimal value of seven: with fewer modes evident patterns remained in the fit residuals, while a larger number of emission modes resulted in worse estimation of the best fit parameters without significant improvements in the residuals.

Fit results are graphically reported in Fig. 3.8, while the best fit parameters values are reported in Tab. 3.2.

The information about emission modes with time constants larger than about one hour comes almost exclusively from the night measurement. For this reason we assigned constant amplitudes to these emission modes. The uncertainties in the model parameters obtained through the fit procedure are only due to the

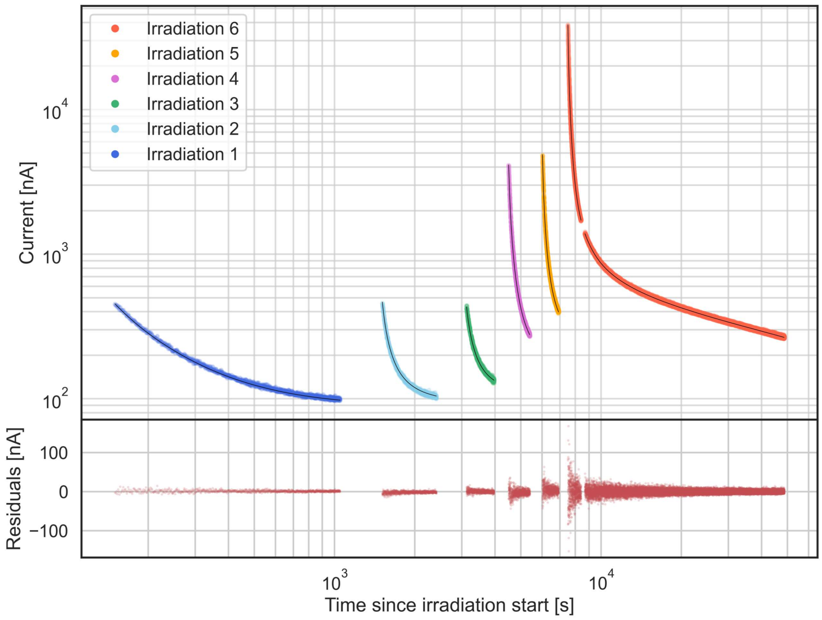


Figure 3.8: Best fit of the January 30th dataset to the model of Sec. 3.2. Parameters estimates are reported in Tab. 3.2. Temperatures in range $21.0 \pm 0.5\,^{\circ}\mathrm{C}$

fluctuations in the data, since we replaced all the other quantities in Eq. 3.2 and the irradiation fluxes with their most probable values. Both the fraction of emitted photons that extract an electron at PMT1 cathode and the fraction of the proton beam intercepted by the crystal are constant multiplicative quantities for which the error propagation is straightforward. However, the model is non-linear in the fluences and gain parameters. Hence a different approach was needed to estimate the contributions to uncertainties from these quantities. We used Monte Carlo techniques, starting from the estimated uncertainties, to evaluate the variation of the model parameters subject to compatibility with the measured data. Results are summarized in Tab. 3.2.

3.5 In-orbit impact of GAGG:Ce afterglow on silicon drift detectors

On average the GAGG:Ce afterglow emission manifests itself as a continuous stream of optical photons with monotonically decreasing flux after stimulation. The randomly arriving photons are able to induce detectable anode current pulses on PMTs but not on the highly-efficient yet unamplifying SDDs. Hence

τ_i $[s]$	σ_τ $[s]$	n_i	σ_n	Δn	$\sigma_{\Delta n}$
24.9	0.8	197.1	80.9	72.0	31.5
69.1	2.8	121.1	49.4	−27.6	12.5
194.1	4.3	93.4	38.2	−30.6	13.4
697.1	29.7	48.7	20.2	−27.3	12.9
2239.5	106.2	35.0	14.8	−18.7	10.9
8315.7	242.8	44.8	18.2	N/A	N/A
70138.5	937.5	375.5	152.8	N/A	N/A

Table 3.2: Emission mode time constants (τ), amplitudes (n) and their relative variations (Δn) as estimated from fit to the models of Sec. 3.2. The amplitude value is fixed for the last two emission modes, while it changes proportionally to the fluence for the other modes. The uncertainties include all the known sources of error, see the text. Afterglow data from GAGG:Ce sample at temperature $21.0 \pm 0.5\,°\mathrm{C}$.

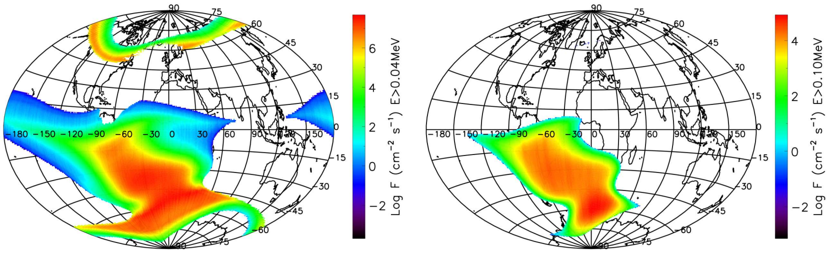
scintillator afterglow will not result in triggering HERMES SDD front-end electronics. Instead it will behave as an equivalent leakage current component adding to the true device current.

In space most of the afterglow emission will be induced by the interactions of the particles trapped in the Van Allen radiation belts and the scintillator material. At LEO altitudes the highest concentration of particles are observed in the South Atlantic Anomaly (SAA) and the polar regions. The HERMES foreseen orbit is at low inclination, thus spanning a restricted range of geomagnetic latitudes and grazing the SAA in its outermost regions [33].

Afterglow emission generated during SAA fly-overs is expected to lead to a periodic modulation in device current, thus a periodic degradation of the HERMES detector energy resolution. Most importantly the total device current may exceed the maximum value the SDDs FEE is capable to cope with. The latter contingency is investigated in the remainder of this section through the application of the model described in Sec. 3.2 and Sec. 3.4 and the solution of Eq. 3.1 for irradiation flux profiles representive of those to be expected during space operations.

In order to calculate the trapped particle fluxes expected along the orbit we used the IRENE (International Radiation Environment Near Earth) AE9/AP9 models [52]. These are empirical models for computing proton and electron orbital fluxes. AE8/AP8 versions were developed by NASA and are regarded as the standard tool for radiation belt modeling [5]. In the near future, the

IRENE AE9/AP9 models—which are built upon much more recent trapped radiation observations—are expected to fully replace their predecessors. Both models have been shown to be in disagreement with recent trapped particle radiation observations in low-altitude, low-inclination orbits. In particular, in-situ measurements of trapped particles fluxes have been found to sit in between the predictions of the two models for these orbits, with AE8/AP8 underestimating and AE9/AP9 overestimating particles fluxes [34] [114].
The AE9/AP9 models were run in Monte Carlo mode. In this mode the flux data contains all of the perturbed mean uncertainty plus an estimate of the variations due to space weather processes. In the Perturbed and Monte Carlo modes, the user sets a number of iterations and has the choice to compute the aggregated mean or percentile across the different runs at each time-step [52]. The results of this work were obtained considering 100 iterations at 90% confidence level with time-step 10 s and duration 30 days for an orbit with altitude 550 km and inclination 10 degrees, coherently with the HERMES-TP/SP mission profile. The resulting average integral flux of trapped protons and electrons are represented in Fig. 3.10. In Fig. 3.9 the integral flux maps of trapped particles at altitude 550 km are shown, as expected according to the IRENE AE9/AP9 models.



(a) Trapped electrons with kinetic energy exceeding 0.04 MeV. (b) Trapped protons with kinetic energy exceeding 0.1 MeV.

Figure 3.9: Integral flux maps. Results obtained from IRENE AE9/AP9 model in pertubed Monte Carlo mode, at 90% confidence level for 100 different iterations.

The HERMES detector will accomodate 60 GAGG:Ce scintillators, each of dimensions $12.10 \times 6.94 \times 14.50$ mm^3. We considered a single crystal of the same dimensions. All the crystal surfaces were supposed to be reflective, except for one small face coupled to a pair of SDD cells covering the entire face of the crystal, each of effective area 6.94×6.05 mm^2. Considering the crystal completely shielded on the side opposite to the SDDs, the particle fluxes were integrated upon the remaining five faces of the scintillator.
Since the afterglow model was developed to explain data acquired for a GAGG:Ce sample at temperature $21.0 \pm 0.5\,^{\circ}$C, the results of this section refer to a crys-

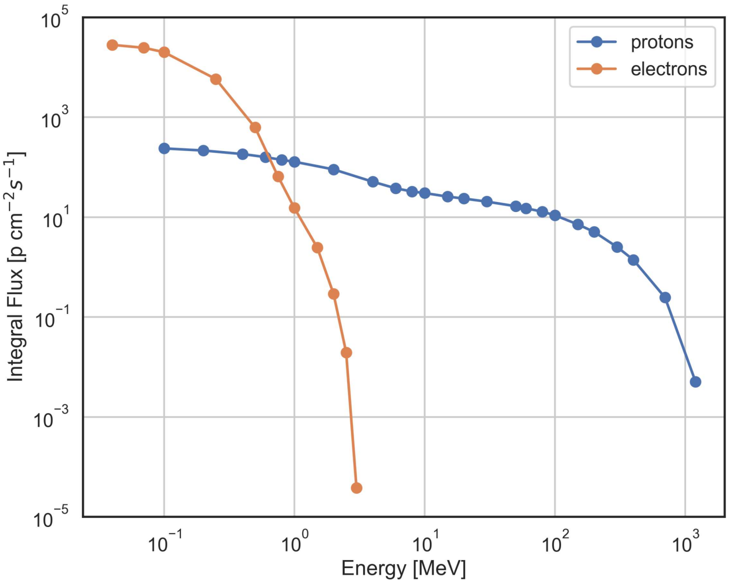


Figure 3.10: Proton and electron average integral flux for an orbit with altitude 550 km and inclination 10 degrees. Results obtained from IRENE AE9/AP9 model in pertubed Monte Carlo mode, at 90% confidence level for 100 different iterations.

tal in the same temperature range. The on-orbit scintillators temperature is expected to be lower, ranging between $-20°$ and $0°$. We expect the intensity of the afterglow emission to decrease with the temperature as a consequence of the increased mobility of the charge carriers.
We used the model of Sec. 3.4 to estimate the leakage current induced by afterglow emission on a single SDD. In particular, referring to Eq. A.8 and 3.2, the expected value of the leakage current is:

$$I_L(t) = e\, f\, \epsilon\, \phi_q(t) \tag{3.3}$$

Where ϵ indicates the quantum efficiency of the SDD, f is the photon transport efficiency from crystal to SDD and e is the elementary charge. The SDD quantum efficiency ϵ was assumed to be constant at value 1 over the afterglow emission spectrum. The photon transport efficiency from the crystal to an SDD cell was set to 0.5 since on average the two cells should receive an equal number of photons. Considering the scintillator as continuosly irradiated at constant average rate for periods of duration equal to the orbital simulation time-step and taking into account the flux values calculated with AE9/AP9, we estimated $\phi_q(t)$ according to a worst-case parametrization of the afterglow emission in which the emission modes have constant amplitudes:

$$n_i^{w.c.} = \max\left(n_i, n_i + \Delta n_i\right)$$

where n_i and Δn_i are the best fit values reported in Tab. 3.2.
Instead of using the ionization energy loss of the particles in the crystal, we considered their kinetic energy and normalized the differential fluxes of trapped particles to the kinetic energy of the protons in the irradiation campaign, i.e. we scaled the flux of particles at energy E (expressed in MeV) by 70/E [1]. Since more energetic protons are not stopped inside the crystal, scaling was performed only for kinetic energies lower than 70 MeV. Particles with energies over 70 MeV were assigned an energy deposit equivalent to the one of a 70 MeV proton.

In Fig. 3.11 we report the resulting estimate of the equivalent leakage current induced by GAGG:Ce afterglow on the SDD, as expected over 100 orbits

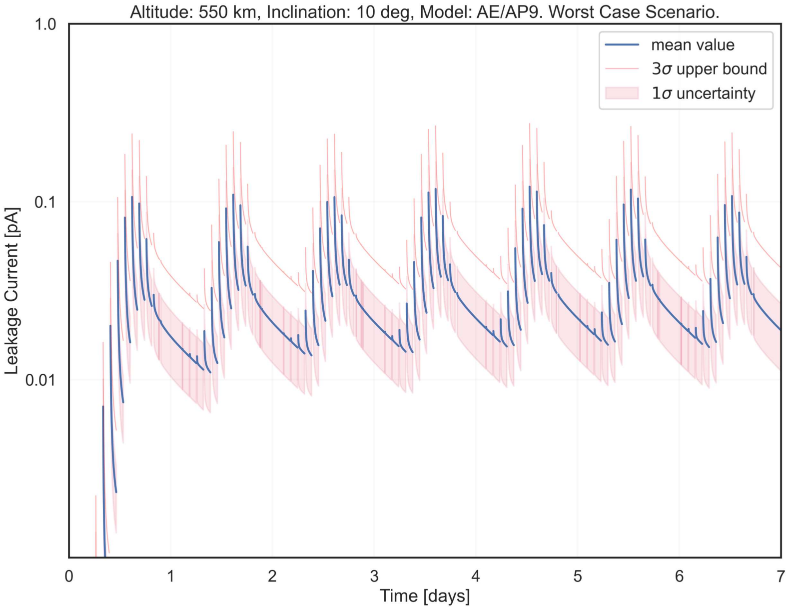


Figure 3.11: Estimated worst-case leakage current of an SDD cell of dimension 6.94 × 6.05 mm^2 induced by GAGG:Ce afterglow emission, as expected from orbital irradiation of a 12.1 × 6.94 × 14.50 mm^3 scintillator at a temperature of $21.0 \pm 0.5\,^\circ$C in 550 km, 10 degrees orbit over $\sim$ 7 days period (100 orbits). Values during transits over SAA were not computed. The 1σ uncertainty region is represented as a shaded band. The 3σ upper-bound is reported as a red solid line. Orbital populations of protons and electrons were modelled through AE9/AP9 packages.

---

[1] For example, a 70 p cm^{-2} s^{-1} flux of 1 MeV particles is converted to a 1 p cm^{-2} s^{-1} flux of 70 MeV protons

at altitude 550 km. We do not display values inside and up to one minute after SAA transits. This choice is due to our model not being valid at times earlier than one minute after the end of an irradiation. Moreover HERMES instruments will be turned off during SAA fly-over. The reported minimum, mean and maximum values of leakage current were computed starting after 24 hours of orbital lifetime. Given the exponential nature of the afterglow model, these numbers are equally well representative when a larger number of orbits are taken into consideration. For instance, considering the full 30 days simulation (~ 400 orbits) we found an increase in the computed maximum leakage current of about 2%, which is comparable to the geographical fluctuations of trapped particle fluxes on the same interval. In our calculations the increase in leakage current due to displacement damage in the SDDs was not taken into account.
The HERMES low-noise front-end electronics (FEE) is able to grant nominal performance up to ~ 100 pA of leakage current, a value well above the estimated maximum (by about two orders of magnitude). We conclude that GAGG:Ce afterglow should not endanger the well-functioning of the FEE. However, when the HERMES fleet will be enlarged—eventually hosting spacecrafts in orbits at higher inclinations and delving in regions of much higher trapped particles concentration—the impact of afterglow on detector performance will demand further investigation.

Chapter 4

Algorithms for Gamma-Ray Bursts detection

Gamma-ray bursts manifest as sudden, transient increases in high-energy detectors count rates. These events will appear unexpected, their activity not explainable in terms of background activity or any known sources. Any automated procedure for detecting GRBs is generally concerned with monitoring the significance of count excess relatively to background in observations. For space-born observatories it is often critical for this task to be held online i.e., as the data are collected. A timely alert may in fact serve to trigger the acquisition of a recording apparatus or to initiate follow-up observations from the ground.
The core logic of trigger algorithms (TA) has gone largely unchanged across different generations of spacecrafts hosting experiments for GRB detection. Photon counts are averaged over a pre-set timescale and a trigger is issued whenever the photon count exceeds the expected background count by a threshold margin. The threshold value is conventionally expressed in units of standard deviations. Generally multiple logics with different timescale parameters are operated simultaneously. A physical reason motivates this fact: GRBs have been observed with much different durations. Since the duration and timing of photons count buckets affect the significance estimates, multiple algorithms are useful to minimize the chance of a burst being missed due to a mismatch.
Over time trigger algorithms have grown to support an increasing number of criteria and parameters to achieve a better coverage of the GRBs parameter space, hence greater sensibilities to different bursts. For example, while BATSE allowed for the operation of only three different logics, more than 120 and 800 different trigger criteria can be specified for Fermi GBM [81] and Swift BAT [79] respectively, see the discussion of Sec. 1.6. The simultaneous operation of many

trigger algorithms can require most of the available CPU time [42] [91]. This picture is unfavorable to nanosatellites mission. Indeed, many of the constraints peculiar to miniaturized spacecrafts lead to limited computational resources.

Several alternative strategies for detecting GRBs are possible. CUSUM (cumulative sum control charts in some texts) for example is a well established and efficient changepoint detection technique around which a TA for GRB detection could be designed[94]. However, much similarly to conventional TAs, a parameter characteristic of the anomaly intensity must be specified for any CUSUM search and detection efficiency is negatively affected whenever the burst brightness does not match such parameter. In a 2012 work [103], J.D. Scargle and B. Jackson proposed using a dynamical programming, changepoint detection algorithm for TA. The algorithm discussed in that paper generalizes a technique previously known to the astrophysics community as Bayesian Blocks [102]. Bayesian Blocks can be thought as a procedure for constructing a histogram with unfixed bin-lengths; the histogram bin edges dynamically determined through a changepoint detection method. It is a flexible and powerful technique which overcomes biases introduced by pre-defined time scales. However, the computational cost of Bayesian Blocks grows quadratically with the number of data and hinders applicability when data are collected at high frequencies or computational resources are on a budget.

In the first section of this chapter different fundamental trigger logics will be discussed. These include the application of a novel changepoint detection technique called FOCuS to the problem of detecting Gamma-Ray Bursts [117]. In a nutshell, this algorithm evaluates count data observations over time-scales which are dynamically determined through a computationally efficient change point detection technique rather than being pre-set (and limited) by the user. All TAs require the assessment of a reference background count-rate. The traditional approach to this problem can be reduced to that of computing a simple moving average of past observations. We discuss different moving average approaches to background estimate and advocate the use of computationally efficient, trend-resistant technique based on exponential smoothing. In the last section of this chapter, we discuss the implementation and testing of different trigger algorithms over both synthetic and real data addressing metrics such as detection efficiency, false positive rates and computational performances.

4.1 Trigger Algorithms

4.1.1 The problem of GRB detection

Consider the problem of detecting a GRB in the following, simplified form. Over time, high-energies photons from a number of unknown sources reach a spacecraft detector. The detected photons are counted over fundamental acquisition times or sampling periods of duration τ. The count values are recorded so that, after a time $t\tau$ from the start of the measurement, a number t of values is collected in a time series. We note the i-th value of the count time series x_i so that the count time series consisting of the first t values $X_t = \{x_1, .., x_t\}$. For most of the time, photons reaching the detector are stochastically emitted by a number of unknown background sources in a predictable way. We suppose to be able to estimate or forecast the number of photons expected from background over each sampling interval. The background estimates are represented in a parallel time series with values b_i, $B_t = \{b_1, .., b_t\}$.

Each interval or segment of the time series X_T and B_T is unequivocally identified by two indices i and h representing the interval last value index and the interval length respectively. For example, the observation interval comprising all the values between x_{i-h+1} and x_i can be represented using the compact notation $X_{i,h} = \{x_{i-h+1}, x_{i-h+2}, .., x_i\}$. At time $t\tau$ a number $t(t+1)/2$ of unique intervals exists; t of these intervals are 'recent' meaning that they contain the most recent observed count value x_t. We note the total number of counts in a given interval with a lowercase letter, $x_{i,h} = \sum_{j=i-h+1}^{i} x_j$. The same notation is adopted for identifying intervals of the background series, so the number of photons expected from background over the same interval is $b_{i,h}$.

Each interval is associated with a statistical significance score $S_{i,h}$. The significance measures the 'extraordinariness' of the number of photons collected relatively to the number of photons expected from background over the same interval of time so that $S_{i,h} = S(x_{i,h}, b_{i,h})$. In physics it is customary to express significances in units of standard deviation. For example, consider a Poisson process where b counts are expected over a certain interval of time. The probability of observing more than n counts over the same interval is given by the Poisson cumulative distribution function:

$$\alpha = \sum_{i=n}^{\infty} \frac{b^i \exp(-b)}{i!}$$

Incidentally we note that α can be expressed:

$$\alpha = 1 - \frac{\Gamma(n+1, b)}{\Gamma(n+1)}$$

where $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ and $\Gamma(a,x) = \int_x^\infty t^{a-1}e^{-t}dt$ are the Γ and incomplete Γ functions. For this process, the significance in standard deviation units is obtained solving in S the equation:

$$\int_S^\infty n(x)dx = 1 - \phi(S) = 1 - \alpha \tag{4.1}$$

Where $n(x)$ is the normal distribution density, ϕ is the standard normal cumulative distribution function and α is the p-value associated with the observation under the background (null) hypothesis.

This work inquiries the problem of designing an automatic procedure to detect the eventual onset of a bright, transient event during observations. Furthermore, we will restrict ourselves to strategies which can be run online i.e., sequentially and on data of indefinite length.

The approach we adopt in this section is purely probabilistic. This implies that crucial aspects of an astrophysical signal such as features of the transient spectra or temporal evolution will be disregarded. Moreover, all the details of the process of background estimate will be neglected; a reasonable estimate of the background count rate b_i is assumed to be given each time a newly observed count value x_i is available. The limitation of this approach means that we will be free of any oblige related to assigning qualitative meaning to our observation: we will not be looking for GRBs, but for bright episodes which are unlikely to stem from background by pure chance. In these terms, the problem of detecting a GRB may be recast as an interval search problem. Our goal is to identify the most efficient strategy to reliably detect intervals of the observation time series $X_{i,h}$ whose associated standard deviations significance $S_{i,h}$ exceeds an arbitrary, positive threshold value T.

4.1.2 The conventional approach to GRB detection

The core logic behind TA for GRB detection has gone largely unchanged through different generations of spacecraft and experiments: count-rates are formed, a background rate estimate is computed and the number of photons observed over recent intervals of different length (timescales, binning) is compared against the photon number that would be expected from background. This process goes on until an excess in observed counts high enough is observed. Which timescales are scheduled for a check is set by the user. The same holds for eventual waiting times (time offsets) between different checks at a particular binning. Most often timescales are set logarithmically equispaced in the range between milliseconds and seconds. For longer timescales, longer waiting times are commonly em-

```
while not triggered do
    $i \leftarrow i + 1$;
    for $h \in H$ do
        if $i\%h = 0$ then
            Compute $S_{i,h}$;
            if $S_{i,h} > T$ then
                triggered $\leftarrow$ True;
            end
        end
    end
end
```

Algorithm 1: Sketch of a conventional algorithm for GRB detection. The maximum significance in excess counts is computed over a grid of observation intervals with bin-lengths $h \in H$.

ployed. For example, during the first year of Fermi operations, the GBM trigger algorithm tested up to eleven logarithmically equispaced timescales spanning 16 to 16384 ms at offsets equal to half the acquisition length [81] [91]. The logic of an elementary TA algorithm is sketched in Alg. 1. A visual representation of the algorithm operation is proposed in C.3 and discussed in Sec. C.

The conventional algorithm's approach to the GRB detection problem is a brute-force one. A grid search of the observation interval significance space is performed over time until an interval significant enough is eventually tested. If an interval $X_{i,h}$ with significance $S_{i,h} > T$ exists, there is no guarantee that a conventional TA operating at threshold T will result in a trigger. For this to happen, either the burst timing should match the grid search or multiple, significant enough intervals must exist. For a bright source underlying an excess in counts, this implies to either be *on time* or to be *bright enough.* In other words, detections from conventional TAs are biased towards events with timings and durations matching those of the grid search.

The per-iteration computational costs of conventional TAs is bound by a constant. Many of these computations are effectively wasted. Indeed, such logic makes no use of the information it acquires during its operations. This information can be used to avoid computing significance scores for intervals which can not possibly result in a trigger. This statement is best illustrated through a simple numerical example. Consider a trigger algorithm like the one in Alg. 1 operating with a positive standard deviation threshold $T > 0$ and checking bin-lengths $H = \{1, 2\}$ over the count observations $X_2 = \{90, x_2\}$ and Poissonian background count-rates $B_2 = \{100.0, 100.0\}$. The computation of the significance $S_{2,2}$ is wasted since either $x_2 > 100$ and $S_{2,2} < S_{2,1}$ or $x_2 \leq 100$ while both $S_{2,2} \leq 0$ and $S_{1,2} \leq 0$. Under these conditions either $S_{2,1} > T$ or no trigger are possible at all, making the actual computation of $S_{2,2}$ unnecessary. We used high background count-rates values to allow for the normal approximation of the Poisson distribution, under which the previous relations are trivially verified due to $S_{i,h} \sim (x_{i,h} - b_{i,h})/\sqrt{b_{i,h}}$. Similar considerations are possible when background count-rates are small. In order make use of the information acquired during its operations, a TA must be equipped with a memory state.

4.1.3 Exhaustive search

It is easy to design an exhaustive search algorithm for solving the GRB detection problem: at the i-th iteration step all the recent intervals (i.e. all $X_{i,h}$ such that $0 < h \leq i$) have their significance score $S_{i,h}$ computed and tested against a threshold T. A sketch of an exhaustive search algorithm is presented in Alg. 2. A visual representation of the TA operations is proposed in Fig. C.2 and discussed in Sec. C. Since the total number of intervals in an observations series X_t is $t(t+1)/2$, the computational cost of an exhaustive search algorithms grows as the square of the number of observations. For this reason, exhaustive search algorithms are of no practical interest for online applications. Still, an exhaustive search TA is a useful benchmark against which the detection performances of other techniques can be evaluated. The detection power of an exhaustive search algorithm for GRB detection is in fact ideal, meaning that it would always be able to meet the trigger condition over the earliest intervals whose significance exceeds the algorithm threshold. An exhaustive search TA running over data for which the background count-rate is known is biased only by the fundamental binning of the data. Within the framework we choose, detection from an exhaustive TA are as unbiased as they get.

```
while not triggered do
    $i \leftarrow i + 1$;
    for $h \leq i$ do
        Compute $S_{i,h}$;
        if $S_{i,h} > T$ then
            triggered ← True;
        end
    end
end
```

Algorithm 2: An exhaustive search algorithm for GRB detection

Type-I errors. The false positive rate is not uniquely determined by the threshold at which a TA is operated and is not equal to the frequency expected by the normal cumulative distribution. Different TA logics operating with the same threshold do in general have different rates of type-I error. This is a consequence of different TAs performing a different number of tests. The highest possible false positive rate is realized for algorithms performing the largest number of tests, such as exhaustive search TA.
We do consider the problem of quantifying the maximum probability of encountering a false positive at an arbitrary iteration of a trigger algorithm. For simplicity sake we will consider data which are normally distributed. Consider the problem in the following form. Over time a number of independent sampling of a standard distributed variable are performed. Observations are collected in a time series. Using the same notation introduced in Sec. 4.1.1, we note the observation series after the first n samplings $Y_n = \{y_1, y_2, ..., y_n\}$ with $y_i \sim N(0,1)$, an arbitrary interval is represented with $Y_{i,h} = \{x_{i-h+1}, x_{i-h+2}, .., x_i\}$ and the sum of the interval observation with $y_{i,h} = \sum_{j=i-h+1}^{i} y_j$. The random variables $y_{i,h}/\sqrt{h}$ are normally distributed, $y_{i,h}/\sqrt{h} \sim N(0,1)$. For an arbitrary threshold T we want to determine the probability $P\{\sqrt{n}y_{n,n} > T \,\text{or}\, \sqrt{n-1}y_{n,n-1} > T \,\text{or}... \text{or}\, y_{n,1} > T\}$, i.e. the the probability that a false positive will trigger the n-th iteration an ideally sensible algorithm operating at threshold value T.
It is easy to obtain an analytical solution to the problem for the case $n = 2$. In this case, the probability P equals:

$$P = 1 - p(y_1 + y_2 < \sqrt{2}T | y_1 < T)p(y_1 < T)$$

The term $p(y_1 + y_2 < \sqrt{2}T | y_1 < T)$ can be computed geometrically. Noting with $N(x)$ the standard normal density function and with $F(x)$ the standard partition function:

$$\begin{aligned} &p(Y_1 + Y_2 < \sqrt{2}T \,|\, Y_1 < T) = \\ &= \int_{-\infty}^{\sqrt{2}T-T} \mathrm{d}x N(x) + \int_{\sqrt{2}T-T}^{\infty} \mathrm{d}x N(x)\Big[1 - \int_{\sqrt{2}T-x}^{T} \mathrm{d}y N(y)\Big] = \\ &= 1 - \int_{\sqrt{2}T-T}^{\infty} \mathrm{d}x N(x)\Big(\Phi(T) - \Phi(\sqrt{2}T - x)\Big) = \\ &= 1 - \Phi(T)\big(1 - \Phi(\sqrt{2}T - T)\big) + \int_{\sqrt{2}T-T}^{\infty} \mathrm{d}x N(x)\Phi(\sqrt{2}T - x) \end{aligned}$$

Naming $H(T) = \int_{\sqrt{2}T-T}^{\infty} \mathrm{d}x N(x)\Phi(\sqrt{2}T - x)$ the last integral term, the probability P:

$$P = 1 - \Phi(T)\Big(1 - \Phi(T)\big(1 - \Phi(\sqrt{2}T - T)\big) + H(T)\Big)$$

For example, for threshold value $T = 3$ we recover $P \sim 0.0025$, a value larger than $1 - \Phi(T)$ by a factor ~ 1.823.
What for arbitrarily large n? We may treat the array $\mathbf{Y_n} = (y_{n,n}, y_{n,n-1}, .., y_{n,2}, y_{n,1})$ as a multivariate random variable with mean equal to the n-dimensional null vector and covariance matrix $M_{i,j} = \mathrm{cov}(y_{n,n-i+1}, y_{n,n-j+1}) = \min(i, j)$.

$$M_{i,j} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{bmatrix}$$

Given the n-dimensional threshold array $\mathbf{T} = (T, \sqrt{2}T, \sqrt{3}T, ..., \sqrt{n}T)$, we can compute the probability $P(\mathbf{Y_n} < \mathbf{T})$ using the multivariate partition function $\Phi(\mathbf{T})$ i.e. integrating on the hypercube defined by the threshold vector:

$$P = 1 - \Phi(\mathbf{T}).$$

Unfortunately evaluating this quantity is computationally hard, even for small n values.

4.1.4 CUSUM for normal data

CUSUM is a technique for the identification of changes in the parameters of a continuous stochastic process. It was pioneered by statistician E.S. Page in 1954 for applications targeted to industry and quality control [94].
To introduce CUSUM and relate it to the conventional approach to GRB detection, we consider a simplified version of the GRB detection problem built over the assumption that the expectation number of background photons is high enough for the Poisson-normal approximation to hold true. This assumption allows us to replace the observation and the background time series with a new time series object comprising the standardized count observations. As usual, we note the i-th value of the count time series y_i and the count time series consisting of the first t values $Y_t = \{x_1, .., x_t\}$. The i-th value of the standardized time series, y_i, is defined:

$$y_i = \frac{x_i - b_i}{\sqrt{b_i}}$$

Before the eventual onset of a transient event, values in the standardized series are distributed according to a normal distribution with mean 0 and standard deviation 1, $y_i \sim N(0, 1)$. Noting the sum of the values in an interval of the standardized series with $y_{i,h} = \sum_{k=i-h+1}^{i} y_k$, the significance score $S_{i,h}$ associated to the interval $Y_{i,h}$ is:

$$S_{i,h} = \frac{y_{i,h}}{\sqrt{h}}$$

The conventional approach to TA involves sequentially checking for significant enough count excess over intervals with length h. Ideally one would check all the possible lengths. Computational constraints imply that in real settings interval checking is performed only over an user-defined set of h parameters. Under the normal approximation, a conventional TA will issue a trigger whenever an interval $Y_{i,h}$ is tested and its significance exceeds a pre-set threshold T:

$$S_{i,h} = \frac{y_{i,h}}{\sqrt{h}} > T \tag{4.2}$$

If the test of equation 4.2 relies on the durations or time-scale h of a possible transient event, the CUSUM sequential test takes a different approach and focus on the intensity or *post-change mean* μ of a transient. CUSUM is in fact a sequentially performed hypothesis test looking for positive, statistically significant change in the signal mean intensity. We note the value of the post-change mean with the letter μ. In CUSUM, the test statistic $Z_{i,\mu}$ is computed:

$$Z_{i,\mu} = \max\left(0, Z_{i-1,\mu} + 2y_i\mu - \mu^2\right) \text{ and } Z_{0,\mu} := 0 \tag{4.3}$$

The CUSUM test statistic acts as a reservoir for statistical evidences supporting the hypothesis of a change having happened some time in the past. At a given time-indeces i, the CUSUM test statistic exists in one of three possible states:

1. $Z_{i,\mu} = 0$; no statistical evidence towards a change with post-change mean μ exists.
2. $0 < Z_{i,\mu} < T^2$; partial evidences towards a change with post-change mean μ exists.
3. $Z_{i,\mu} > T^2$; enough evidences for a change with post-change mean μ have been collected, within T standard deviations significance level.

The time-indices at which states change are meaningful. Whenever the μ-CUSUM test statistic passes from the first to the second state we talk of a *candidate changepoint* in the standardized time series. Compatibly with the T threshold value, whenever the CUSUM test statistic passes from the second to the third state we have a trigger event. Each trigger event is associated with a *changepoint* which is the latest candidate changepoint prior the trigger. We remark on the fact that the post-change mean μ must be a positive value. This is indeed the case of a GRB or other high-energy astrophysical transient which manifest as transient increases in detectors count-rates. For different applica-

tions, it is easy to modify the recursion of equation 4.3 to search for negative changes in the signal mean.
We can derive the CUSUM test statistic using the hypothesis test formalism. We define the two hypothesis:

- $\mathbf{H}_0$, null hypothesis - no transient up to time-index i: $y_k \sim N(0,1)$ for $k \in \{1,...,i\}$.

- $\mathbf{H}_1$, alternative hypothesis - transient onset at time-index j with post-change mean μ: $y_k \sim N(0,1)$ for $k \in \{1,...,j-1\}$ and $y_k \sim N(\mu,1)$ for $k \in \{j,...,i\}$.

The negative log-likelihood for the null hypothesis:

$$\log\{L(Y|H_0)\} = \sum_{k=1}^{i} \frac{y_k^2}{2}$$

And for the alternative hypothesis:

$$\log\{L(Y|H_1)\} = \sum_{k=1}^{j-1} \frac{y_k^2}{2} + \sum_{k=j}^{i} \frac{(y_k-\mu)^2}{2}$$

Hence the log-likelihood ratio:

$$\log\Big[\frac{L(Y|H_1)}{L(Y|H_0)}\Big] = \Big(\sum_{k=1}^{j-1} \frac{y_k^2}{2} + \sum_{k=j}^{i} \frac{(y_k-\mu)^2}{2}\Big) - \sum_{k=1}^{i} \frac{y_k^2}{2} = \sum_{k=j}^{i} \big(y_k\mu - \frac{\mu^2}{2}\big)$$

Since we are interested in all possible transients start-index j, the CUSUM test statistic is defined:

$$Z_{i,\mu} = \max_{0<j\leq i} \Big[\sum_{k=j}^{i} \Big(y_k\mu - \frac{\mu^2}{2}\Big)\Big] \tag{4.4}$$

E.S. Page showed that this test statistic can be computed recursively according to Eq.4.3 [94].

The CUSUM test statistics $Z_{i,\mu}$ and the binned significance score $S_{i,h}$ relate to each other through two statements.

Statement 1. *If* $S_{i,h} > T$ *then* $Z_{i,\mu} > T^2$ *with* $\mu = T/\sqrt{h}$.

Proof. We can 'unfold' the CUSUM test statistic obtaining:

$$\begin{aligned} Z_{i,\mu} &= \max\left(0, Z_{i-1,\mu} + 2y_i\mu - \mu^2\right) \\ &= \max\left(0, \max\left(0, Z_{i-2,\mu} + 2y_{i-1}\mu - \mu^2\right) + 2y_i\mu - \mu^2\right) \\ &\geq \max\left(0, Z_{i-2,\mu} + 2y_{i-1}\mu - \mu^2 + 2y_i\mu - \mu^2\right) \\ &\geq Z_{i-2,\mu} + 2y_{i-1}\mu - \mu^2 + 2y_i\mu - \mu^2 \end{aligned}$$

Rearranging we obtain:

$$Z_{i,\mu} \geq Z_{i-2,\mu} + 2\mu(y_{i-1} + y_i) - 2\mu^2$$

Similarly, unfolding up to the $i - h$ index and using the definition of $S_{i,h}$, we obtain:

$$Z_{i,\mu} \geq Z_{i-h,\mu} + 2\mu y_{i,h} - h\mu^2 = Z_{i-h,\mu} + 2\mu\sqrt{h}S_{i,h} - h\mu^2$$

Using $S_{i,h} > T$, $\mu = T/\sqrt{h}$ and the fact that the CUSUM test statistic is always non-negative, we have the statement thesis:

$$Z_{i,\mu} \geq Z_{i-h,\mu} + 2T^2 - T^2 > T^2$$

$\square$

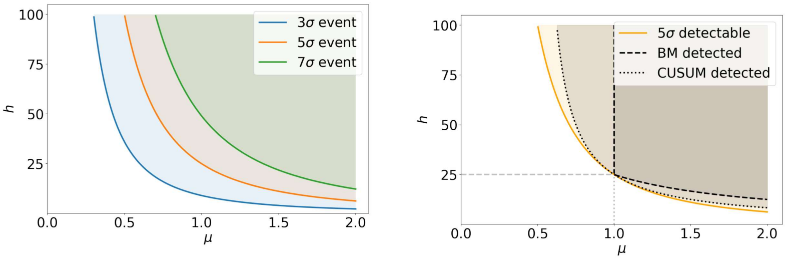
Statement 2. *If $Z_{i,\mu} > T^2$ with $\mu = T/\sqrt{h}$ then $S_{i,k} > T$ for some interval length k.*

Proof. Since $Z_{i,\mu} > T^2$ and by definition $Z_{0,\mu} = 0$, there exists an interval length $i - k > 0$ for which $Z_{i-k,\mu} = 0$ and $Z_{i-j,\mu} > 0 \, \forall j \mid 0 \leq j < t - k$.
Given this observation, we can unfold the test statistic definition up to index $t - k$ disregarding the nested maximum operators:

$$\begin{aligned} Z_{i,\mu} &= \max\left(0, Z_{i-1,\mu} + 2y_i\mu - \mu^2\right) \\ &= Z_{i-1,\mu} + 2y_i\mu - \mu^2 \\ &= Z_{i-2,\mu} + 2(y_i + y_{i-1})\mu - 2\mu^2 = .. \\ &= Z_{t-k,\mu} + 2y_{i,k}\mu - k\mu^2 \\ &= 2y_{i,k}\mu - k\mu^2 \end{aligned}$$

Substituting $y_{i,k} = \sqrt{k}S_{i,k}$ and rearranging:

$$S_{i,k} = \frac{Z_{i,\mu} + k\mu^2}{2\sqrt{k}\mu}$$



(a) Observation domain and detectability regions corresponding to threshold values 3σ, 4σ and 5σ

(b) The 5σ detectability region along with CUSUM and binned detection subdomain.

Figure 4.1: Observation domain

We can cast $Z_{i,\mu} > T^2$ and $\mu = T/\sqrt{h}$, obtaining:

$$S_{i,k} > \frac{T^2}{2\sqrt{k}\mu} + \frac{\sqrt{k}\mu}{2} = \frac{T}{2}\left(\sqrt{\frac{h}{k}} + \sqrt{\frac{k}{h}}\right)$$

Hence $S_{i,k} > T$ independently from the values of k and h . □

The observation domain. An effective visual interpretation of the previous statement is possible. Consider a generic standardized count interval $Y_{i,h}$. This interval can be represented as a point with coordinates $(y_{i,h}/h, h)$ of the space spanning all possible interval intensities and durations. We refer to this space as the *observation domain* [117]. A significance threshold T divides the observation domain in two complementary regions: one containing the intervals whose significance exceed the threshold value, the detectability region, and one containing all other intervals. These regions are separated by the curve $\mu = T/\sqrt{h}$ since observations comprised by an interval with duration h needs to sum up to at least $T\sqrt{h}$ for the interval significance to exceed the threshold value. In these terms, the detection efficiency of a TA can be quantified in terms of the coverage it provides of the detectability region. For a binned search, whatever the bin-length parameter h, a CUSUM search exists whose detection efficiency is strictly better. This CUSUM search has mean change parameter $\mu = 5/\sqrt{h}$. This is the meaning of the statements 1 and 2. In Fig. 4.1b we graphically represent the 5-sigma detectability region along the detection sub-domain spanned by a conventional search with bin-length $h = 25$ and a CUSUM search with post-change mean $\mu = 1.0$.

4.1.5 FOCuS for normal data

FOCuS is a changepoint and anomaly detection algorithm developed by researchers of the department of statistics of University of Lancaster for applications unrelated to astrophysics [98]. Consider again the CUSUM test statistics of Eq. 4.3, this time as a function of the post change mean parameter μ:

$$Z_i(\mu) = \max\left(0, Z_{i-1}(\mu) + 2y_i\mu - \mu^2\right) \text{ and } Z_0^{\mu} := 0 \tag{4.5}$$

In its normal variant, FOCuS provides a recipe to efficiently compute the test statistics :

$$\max_{\mu>0} Z_i(\mu) \tag{4.6}$$

Since for all μ it is true that $Z_{i,\mu} = Z_i(\mu)$, the FOCuS test statistic is equivalent to the maximum value of the CUSUM test statistics over all possible post-change means. Moreover this is equivalent to testing over all bin-lengths and offsets compatible with a series of observations, given the content of the statements 1 and 2.

In contrast to the other logics discussed in this chapter, FOCuS has an internal memory state. This memory state consists of an ordered set of tuples. Each has form (a, b) with a a negative integer and b a positive real. Solutions $Z_i(\mu)$ to Eq. 4.5 are piece-wise quadratic and tuples (a, b) in FOCuS memory state are used to represent different quadratic branches with form $q(\mu) = a\mu^2 + b\mu$, see Fig. 4.2. This implies that the FOCuS test statistics can be expressed as:

$$\max_{\mu>0} Z_i(\mu) = \max_{q_i \in Q_i} -\frac{b_i^2}{4a_i} \tag{4.7}$$

where b_i and a_i represent the coefficients of the quadratic q_i and Q_i is the memory state at the i-th algorithm step. Given this equivalency, it makes sense to refer to tuples and quadratics interchangeably. We will do so for the rest of the present section.

At each iteration, FOCuS updates its memory state according to a few rules.

In subsequent iterations, quadratics in the memory state (a, b) are updated subtracting the a components by 1, while the b components is added two times the most recent standardized observation available.

input : a 2-tuple `q` $= (a, b)$ and a standard observation $y_i \sim N(0, 1)$.
output: the updated tuple `q` $\leftarrow (a - 1, b + 2y_i)$.

Algorithm 3: The function, `update(q,` y_i`)` performs an in-place update over `q` and returns it.

A quadratic may become obsolete and be removed from the memory state. For a quadratic to be obsolete, it must either be negative or *dominated* by an older quadratic. For a quadratic to be negative over all possible post-change mean positive values it is necessary and sufficient that $b_i < 0$. We say that a

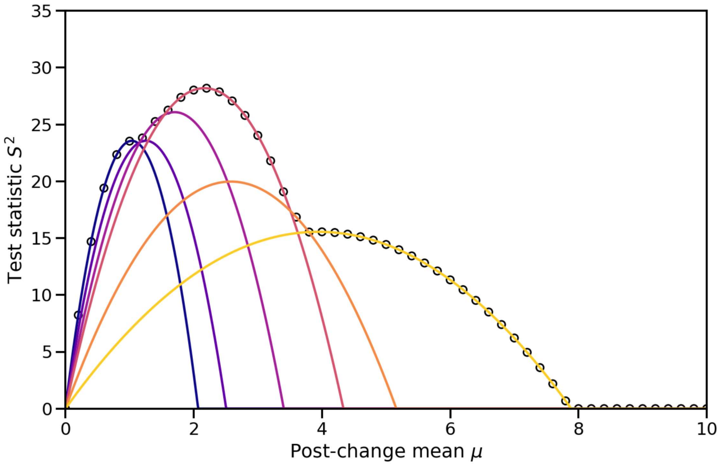


Figure 4.2: FOCuS normal quadratics (solid lines) and CUSUM test statistics (dots) as function of the post-change mean μ after a 5σ trigger event. The CUSUM test statistics values were directly computed over an equispaced grid of post-change mean parameter values. For each value, the CUSUM and FOCuS test statistics are equal, as expected from theory. Time series data were generated ad hoc and comprise a signficant anomalous event.

quadratic q dominates a second quadratic p when $q(\mu) > p(\mu)\ \forall\mu > 0$. This happens whenever q has the derivative at zero and the non-null root exceeding those of p. Whenever a quadratic dominates another, not only that quadratic but all those following it may be removed from the memory state. This operation is called *pruning* and is the key reason behind FOCuS high computational efficiency.

```
input  : Two 2-tuples $q = (a, b)$ and $p = (c, d)$
output: False if $b < d$ or $\frac{b}{a} > \frac{d}{c}$ else True.
```

Algorithm 4: The function `dominate(q,p)` checks if `q` dominates `p`.

Finally, at each algorithm step, a new quadratic may be added to the memory state. New quadratics are added at the right end of the memory state (appended) so that older quadratics always precede newer ones. The order of the memory states must be preserved for the pruning step to work as intended.
FOCuS maximizes each quadratic in the memory state to compute the test statistic of Eq. 4.6. At the end of each step, the value of the test statistics is compared with the squared value of the threshold, expressed in units of standard deviations.
A pseudocode implementation of FOCuS is given in Algo. 6. This particular implementation exploits the use of a dummy head quadratic $(1, 0)$. The head quadratic bears no statistical meaning, it is never checked for a maximum and

```
input  : A 2-tuples q = $(a, b)$.
output: Value $-b^2/4a$.
```

Algorithm 5: The function `ymax(q)` returns the maximum of q.

can not be pruned out of the memory state. It can serve as the head element in a standard linked list implementation of the memory state array. The computational resources needed to complete an iteration of the algorithm is proportional to the number of quadratics in the memory state. Since the average number of components increases as the logarithm of the number of observations, the expected per-iteration time and memory complexity after t observations equals to $O(\log t)$ [98]. A minimal functional implementation of FOCuS Normal is proposed in B.1.

```
init:
an ordered list of 2-tuples Q = $[(1, 0)]$;
global_max ← 0;
time_offset ← 0;
while not triggered do
    $i \leftarrow i + 1$;
    set q to the first element in Q;
    Q.append($(0, 0)$);
    while next(q) is not null do
        p ← update(next(q),$y_i$);
        if dominate(q,p) then
            delete all elements after q from Q;
            break;
        else
            if ymax(p) > global_max then
                global_max ← ymax(p);
                time_offset ← p.a;
            end
        if global_max > $T^2$ then
            triggered ← True
        end
    end
end
```

Algorithm 6: FOCuS for normal data

4.1.6 Limits of the normal approximation

The implementation of FOCuS proposed in Algo. 6 is designed to operate over normally distributed data points. Hence, a sound application over Poisson-distributed data requires high mean count-rates. In this regimen in fact the Poisson distribution is well approximated by a normal distribution. For a given detector such requirements would impose a lower bound to the count acquisition frequency. This is undesirable, especially for detectors with small effective area, since the detection power of count-rate TAs decreases steeply when the characteristic duration of the anomalies is shorter than the fundamental acquisition time. For HERMES for example on average about 70 background photons are expected per second in the energy band 50-300 keV. Requiring an average count rate of 100 photons would limit the fundamental acquisition time to $\sim$ 1.5 s. For reference, about the $\sim$ 15% of the bursts in the GBM catalog have a T_{90} smaller than 1.5 s.
What are the dangers when running a normal algorithm over count data with small mean rate? In this condition, a TA has a rate of type-I errors (false positives) higher than one would expect. To illustrate this point consider a very simple normal TA operating at threshold T and sequentially computing the significance of the individual observations x_i according to:

$$S(x_i, b) = \frac{x_i - b}{\sqrt{b}} \tag{4.8}$$

If the random variable x_i are normally distributed, $x_i \sim N(b, \sqrt{b})$, the probability that $S(x_i, b) > T$ is:

$$\phi(T) = \int_T^\infty f(x)dx$$

where $f(x)$ is the density function of a standard normal distribution. The probability of the algorithm running up to the i-th step follows the geometric distribution:

$$P_i = \phi(T)(1 - \phi(T))^{i-1}$$

Hence the algorithm *average run length*, the mean number of steps before the first trigger when no change occurs:

$$ARL = E(X = i) = \sum_{i=1}^{\infty} i(1-\phi)^{i-1}\phi = \frac{1}{\phi} \tag{4.9}$$

Where we used the fact that $\sum_{n=1}^{\infty} nx^n = \frac{1}{1-x}$. In order to produce the results of Fig. 4.3, a TA as discussed was run over two streams of simulated data.

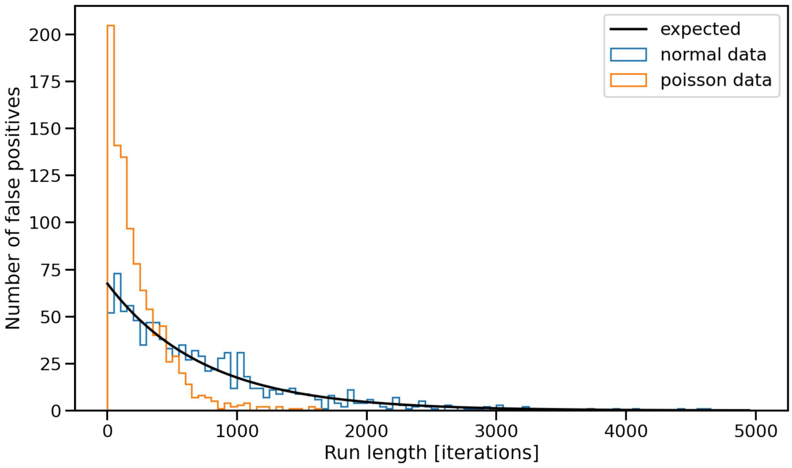


Figure 4.3: Empiric distribution of the run length of a normal TA operating at threshold $T = 3\sigma$ and sequentially computing the significance of individual observations according to Eq. 4.8. The algorithm was run over Poisson-distributed data with mean rate 2.5 (orange steps) and simulated normal data with $\mu = 2.5$ and $\sigma = \sqrt{2.5}$, (blue steps). For each case, the test was repeated 1000 times. The theoretical run length distribution is plotted with a black solid line.

The first stream was made of random variables sampled from the distribution $x_i \sim N(2.5, \sqrt{2.5})$, while the second stream was obtained sampling a Poisson distribution with mean rate 2.5, $x_i \sim P(2.5)$. The algorithm ran for as long as $S(x_i, 2.5) < 3$ and the stop index was recorded. The same test was repeated 1000 times for each data stream. The average run length over the Poisson data streams equals 232.5 steps while for the normal data a value of 740.8 was observed, compatible with the theoretical 733.9 ± 23.4 expected according to Eq. 4.9.

Techniques able to mitigate this problem exist. For example, the Li-Ma formula for assessing the significance of count excess yields a method to standardize count data performing well for count-rates $\lambda \gtrsim 10$ [74], provided a background measure with long enough exposure. Similar techniques for standardizing count data when a background estimate is obtained from a model do exist [111].

4.1.7 CUSUM for Poisson data

The distinctions between the Poisson and normal distributions make the choice of a multiplicative anomaly model more natural than an addictive one. For a known background rate λ, the mean of an observation interval is classified $\mu\lambda$, with intensity multiplier $\mu \geq 0$. Hence an excess in count will be charatcterized by intensity multiplier values $\mu > 1$ while for lack in counts $0 \leq \mu < 1$.

Techniques to adapt the CUSUM method to Poisson-distributed count data exist in literature [75][**?**]. The CUSUM-Poisson test statistic can be recovered from a hypothesis test, as we did in the normal case. Assuming the background count rate to be known for each observation, the hypothesis we test are:

- $\mathbf{H}_0$, null hypothesis - no transient up to time-index i: $y_k \sim P(\lambda_k)$ for $k \in \{1, ..., i\}$.
- $\mathbf{H}_1$, alternative hypothesis - transient onset at time-index j with intensity multiplier μ: $y_k \sim P(\lambda_k)$ for $k \in \{1, ..., j-1\}$ and $y_k \sim P(\mu\lambda_k)$ for $k \in \{j, ..., i\}$.

Through arithmetic essentially equal to that required to obtain Eq. 4.4, we can write the log-likelihood ratio:

$$\log\left[\frac{L(Y|H_1)}{L(Y|H_0)}\right] = \sum_{k=j}^{i}\big(x_k \log\mu - \lambda_k(\mu - 1)\big)$$

Maximizing over the anomaly start index j, one obtains:

$$\zeta_{i,\mu} = \max_{0<j\leq i}\sum_{k=j}^{i}\big(x_k \log\mu - \lambda_k(\mu - 1)\big)$$

The CUSUM Poisson test statistic can be computed recursively through:

$$\zeta_{i,\mu} = \max\left(0, \zeta_{i-1,\mu} + x_i \log\mu - \lambda_i(\mu - 1)\right) \tag{4.10}$$

How does one set a threshold for the Poisson CUSUM test statistics? In order to identify a count excess within T standard deviations for an interval $X_{i,h}$ it is necessary that:

$$\sum_{x=x_{i,h}}^{\infty}\frac{(\lambda_{i,h})^x \exp^{-\lambda_{i,h}}}{x!} \leq p_T = 1 - \phi(T)$$

The Poisson survival function term at the LHS is generally inconvenient to compute. Significant streamlining can be achieved using a result known as Wilks theorem [117] [118]. The Wilks theorem states that the log-likelihood test statistic:

$$TS = 2\log\left[\frac{L(Y|H_1)}{L(Y|H_0)}\right]$$

is asymptotically distributed as the chi-squared distribution χ^2_d, where the d equals the difference in degrees of freedom between the null and the alternate hypothesis. In our case, the only degree of freedom comes from the intensity multiplier parameter μ so that $d = 1$. Given the relation between the chi-squared and normal distribution, the Wilks theorem ensure that a reasonable

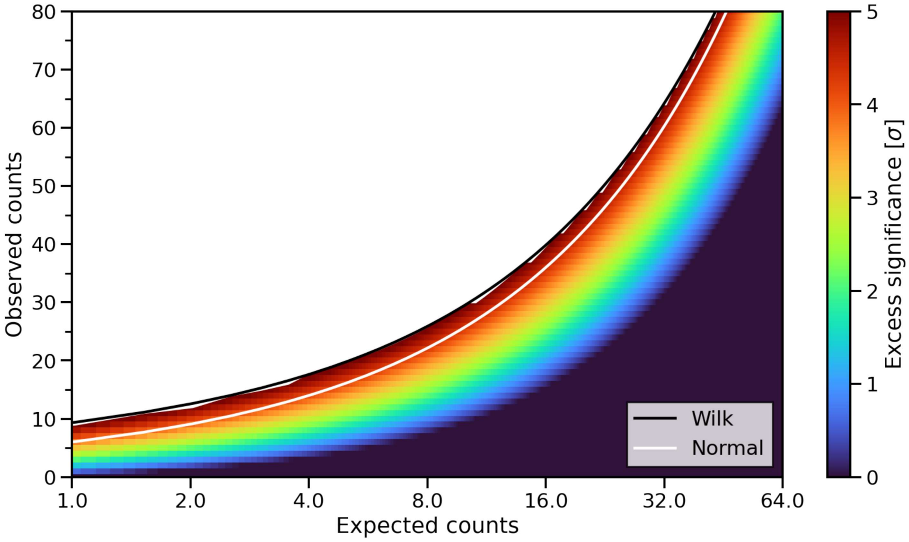


Figure 4.4: Threshold values according to Wilks theorem approximate well the true value, even when the expected background count is small. The x-axis and y-axis represent different Poisson processes mean rates and observed count values, respectively. Different colors are used to represent the significance of excess counts, computed according to the true value of Eq. 4.1.7 and in units of standard deviations. The white region contains those observations with significance $> 5\sigma$, a typical threshold value. A black solid line is used to represent the Wilk's threshold, while a white solid line represent the effect of the normal approximation on the detectable events. The dark blue region of the plot contains observations lacking in counts when compared to the expected value. For these points, the actual (negative) significance is not plotted.

approximation of Eq. 4.1.7 is given by:

$$\zeta_{i,\mu} > \frac{T^2}{2} \tag{4.11}$$

The threshold condition of Eq. 4.11 is fast to compute and provides a good approximation even at modest values of λh for T values higher than 5, see Fig. 4.4. Note that the Wilks threshold condition can be used outside the CUSUM framework. In particular, maximizing Eq. 4.1.7 over μ and applying the Wilks theorem, we obtain the significance test:

$$n \log\left(\frac{n}{b}\right) - (n - b) > \frac{T^2}{2} \tag{4.12}$$

where n and b are the number of counts observed and expected from background, respectively, and T is a threshold value in units of standard deviations.

4.1.8 FOCuS for Poisson data

It is possible to adapt the FOCuS algorithm for Poisson distributed count data [117]. The Poisson version of FOCuS was developed by Kester Ward, Idris Eckley and Paul Fearnhead of the department of statistics of the University of Lancaster which kindly accepted our invitation to collaborate on this subject.
Consider again the Poisson CUSUM test statistics, this time as a function of the intesity multiplier parameter μ:

$$\zeta_i(\mu) = \max\left(0, \zeta_{i-1,\mu} + x_i \log\mu - \lambda_i(\mu - 1)\right) \tag{4.13}$$

Solutions to Eq. 4.13 are piece-wise functions but—crucially for the operations of FOCuS Poisson—they are not piece-wise quadratics, as it was the case for the solutions $Z_i(\mu)$ to Eq. 4.5. A branch c of $\zeta_i(\mu)$ is in fact a logarithmic curve whose general expression is:

$$c(\mu) = a\log(\mu) + b(\mu - 1). \tag{4.14}$$

We will refer to a generic branch of $\zeta_i(\mu)$ with the more general term of *curve*, in contrast with the term quadratic which we used to identify branches of $Z_i(\mu)$.
The test statistic sequentially computed by FOCuS-Poisson is obtained maximizing solutions to Eq. 4.13 over the intensity multiplier parameter μ:

$$\max_{\mu>1} \zeta_i(\mu) \tag{4.15}$$

Since we are interested in anomalous excesses in observations, only those values of the intensity multiplier exceeding 1 are considered. Similar considerations hold for a version of the algorithm searching for anomalous lacks and operating over the intensity multiplier domain $0 \leq \mu < 1$.
The memory state of FOCuS-Poisson is composed by chronologically ordered tuples; each of the tuples being associated with a particular curve branch of $\zeta_i(\mu)$. Given this relation we refer to the memory state tuples and the curves they associate to interchangeably. A tuple associated with the curve $c(\mu)$ as in Eq. 4.14 has three components, (a, b, c). This contrasts with the 2-tuples managed by the normal version of the algorithm. The a component of a tuple is a non-negative integer carrying a cumulative record of the counts observed since tuple creation. Similarly, the tuple component b is a negative real number which carries a record of the cumulative background counts as expected since the tuple inception. Ultimately, the need for an extra tuple component stems from the fact that the shape of a Poisson distribution with a given mean rate parameter depends on the mean rate parameter itself. This fact implies that it

is not possible to summarize the information content of backgrounds estimates and observations into a single, standardized variable. Finally, the c components is a negative integer serving as a step-index 'clock' to keep track of a tuple age. The sign choice of the tuple components is arbitrary and different choices are possible as long as the functions `update`, `ymax` and `dominates` are adjusted accordingly.
Speaking of these function, moving from normal to Poisson a number of fundamental changes are needed. In order to update a tuple (a, b, t) we do 1. sum the most recent count observation to the a tuple component; 2. subtract the most recent background estimate to the b component and 3. diminish the clock component t by a unit.

```
input  : a 3-tuple c = $(a, b, t)$, a background estimate $\lambda_i$ and a count
         observation $x_i \sim P(\lambda_i)$.
output: the updated tuple c $\leftarrow (a + x_i, b - \lambda_i, t - 1)$.
```

Algorithm 7: The function, `update(q,` x_i, λ_i`)` performs an in-place update over `q` and returns it.

A tuple is deemed obsolete when it gets dominated by a tuple which predates it. When this happens, that particular tuple and all the tuples which follow may be removed from the memory state. The algorithm computational efficiency benefits greatly from curves *pruning* since less computations will be needed in subsequent iterations. We say that a curve c dominates a second curve k, whenever $c(\mu) > k(\mu)\ \forall \mu > 1$. Dealing with the normal version of the algorithm, to assess if a quadratic dominated another, all that was needed was to compute and compare the quadratics derivative at zero and the quadratic non-zero root. Both these quantities had simple analytic forms. No analytic form exists for the non-unit roots of a logarithmic curve such as c from Eq. 4.14. Estimating the root value numerically is computationally expensive, hence undesirable. One can overcome this difficulty noting that the value of the non-unit root of $c(\mu)$ depends only on the ratio $\frac{a}{b}$. This is evident by the fact that multiplying the a and b component only rescales the curve values and does not affect the location of the root. Hence, in order to assess if a curve c dominates a second curve k it is enough to compute and compare the curves derivative at $\mu = 1$, $a + b$, and the components ratio, a/b.

```
input  : Two 3-tuples c = $(a, b, t)$ and k = $(d, e, s)$
output: False if $a + b < c + d$ or $\frac{a}{b} > \frac{d}{e}$ else True.
```

Algorithm 8: The function `dominate(c,k)` checks if `c` dominates `k`.

During the i-th step, a new curve $(x_i, -\lambda_i, -1)$ may be added to the memory state if it is not dominated by any other curves in the memory state or, in case of an empty memory state, if $x_i > \lambda_i$. Once the memory state has been updated the algorithm still needs to maximize the curves in its memory state to compute its test statistic. To do so it is enough to compute and compare the maximum for each curve in the memory state, i.e. to compute the quantities:

$$\max_{\mu>1} \zeta_i(\mu) = \max_{c_i \in C_i} a_i \log\Big(-\frac{a_i}{b_i}\Big) - (a_i + b_i) \tag{4.16}$$

where a_i and b_i are the components ot the curve c_i and C_i is the memory state at the i-th algorithm step. Note that this operations requires the computationally expensive evaluation of $\log(-a_i/b_i)$ for each of the tuples in the memory state.

```
input  : A 3-tuples $c = (a, b, t)$.
output: Value $a_i \log(-\frac{a_i}{b_i}) - (a_i + b_i)$.
```

Algorithm 9: The function `ymax(c)` returns the maximum of c.

Once these fundamental distinctions have been made clear, the FOCuS logic flow is essentially unchanged moving from normal to Poisson statistics. A pseudocode implementation is given in Algo. 10. Like its normal counterpart (cf. Algo 6) this implementation exploits a dummy-head curve. The head curve cannot be pruned out of the memory state and its maximum is never evaluated since it bears no statistical meaning. Yet this curve has a notable property: it dominates curves negative over the intensity multiplier domain $\mu > 1$. A visual representation of the algorithm operation is proposed in C.4 and discussed in Sec. C. As with the normal version of the algorithm, the amount of computational and memory resources needed to complete an iteration of the algorithm is proportional to the number of curves in the memory state. In Fig. 4.5 results from a simulation aimed at comparing the average number of memory state components between FOCuS normal and FOCuS Poisson are shown. In this simulation, the average number of curves in FOCuS Poisson memory state shows asymptotic behaviour similar to FOCuS normal, whose theoretical number of memory state components after t iterations is expected to be proportional $\log t$ (see Sec. 4.1.5). A minimal functional implementation of FOCuS Poisson is proposed in B.2.

4.1.9 Other FOCuS implementations

Implementations with constant amortized update. It is possible to further improve the implementations described by Algo. 6 and Algo. 10. At each

```
init:
an ordered list of 3-tuples, C = $[(1, -1, 0)]$;
global_max ← 0;
time_offset ← 0;
while not triggered do
    $i \leftarrow i + 1$;
    set c to the first element in C;
    C.append($(0, 0, 0)$);
    while next(c) is not null do
        k ← update(next(c),$x_i, \lambda_i$);
        if dominate(c,k) then
            delete all elements after c from C;
            break;
        else
            if ymax(k) > global_max then
                global_max ← ymax(k);
                time_offset ← k.t;
            end
        if global_max > $T^2$ then
            triggered ← True
        end
    end
end
```

Algorithm 10: FOCuS for Poisson data

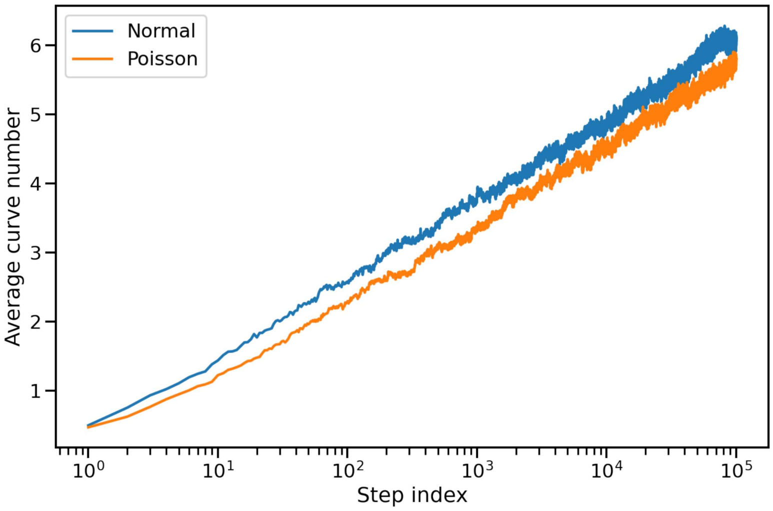


Figure 4.5: FOCuS average number of components versus algorithm iteration step index. A blue line is used to represent the number of components of FOCuS normal, averaged over 10^3 datasets composed of 10^4 random numbers sampled from normal distribution $N(0,1)$. Similarly, an orange line represents the number of components of FOCuS Poisson, averaged over 10^3 datasets composed of 10^4 random realizations sampled from a Poisson distribution with mean rate $\lambda = 2.5$, $P(2.5)$.

algorithm iteration, FOCuS accomplishes two essential tasks: 1. it updates the memory states according to the most recent observation and background estimate; and 2. it maximizes over the memory state to compute the FOCuS test statistics. In our pseudo-code implementation these tasks are intertwined and each requires a number of computations proportional to the number of components in the memory state. It is possible to develop implementations of FOCuS in which the update and maximization steps are decoupled from each other and performed sequentially. Moreover, the computational cost of the update step may be constrained. For FOCuS Normal, implementations bringing the amortized computational cost of the update step down to $O(1)$ exist in the literature [98]. Similar solutions for FOCuS Poisson are currently under development [117]. The key fact making linear cost update steps possible is that the difference between the components of two tuples in the FOCuS memory state at a given time is unchanged through the update step, making individual tuple updates unnecessary for all tuples but one.

Approximated implementations Even if the computational requirements of the update step can be constrained, FOCuS still needs to maximize the

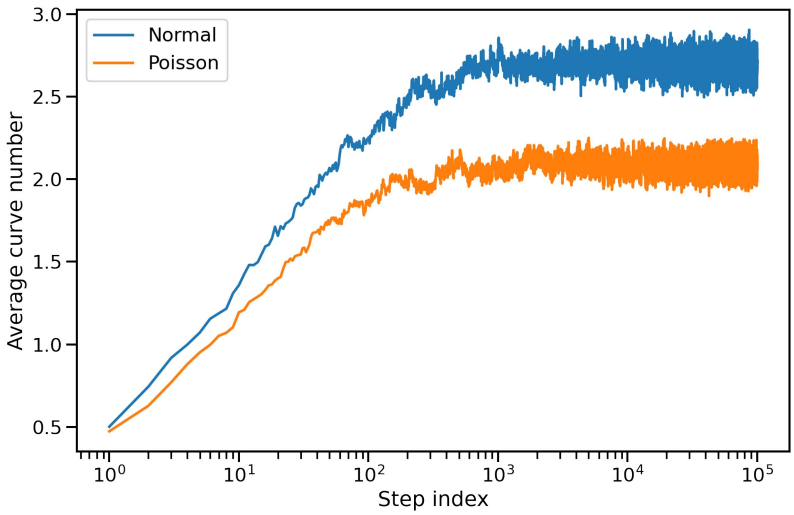


Figure 4.6: FOCuS average number of components versus algorithm iteration step index when mu_{min} cut is used. A blue line is used to represent the number of components of FOCuS normal with $mu_{min} = 0.1$, averaged over 10^3 datasets composed of 10^4 random numbers sampled from normal distribution $N(0, 1)$. Similarly, an orange line represents the number of components of FOCuS Poisson with $mu_{min} = 1.1$, averaged over 10^3 datasets composed of 10^4 random realizations sampled from a Poisson distribution with mean rate $\lambda = 2.5$, $P(2.5)$.

functions represented by the tuples in the memory state. For FOCuS Poisson this is particularly problematic since maximizing the curves of Eq. 4.14 requires evaluating multiple, computationally expensive natural logarithms. Different strategies can be adopted to constrain the number of tuples in the memory state. The most obvious solution is to remove old tuples from the memory state after a certain number of iterations from the tuple inception. It is also possible to remove a tuple from the memory state whenever its non-null (non-unity in the case of FOCuS Poisson) root gets smaller than a threshold value μ_{min}. We refer to these strategies as t_{max} cut and μ_{min} cut, respectively. Both these solutions result in an eventual saturation of the expected number of memory state tuples, effectively constraining the asymptotic computational and memory cost of FOCuS. See for example the results of Fig. 4.6 and compare with Fig. 4.5.

If on one hand it is indeed possible to limit the computational cost of the update step without affecting FOCuS sensibility, on the other hand, the same is not true for techniques such as t_{max} cut and μ_{min} cut. This is not always bad news. For example, the net effect of μ_{min} cut is to make the algorithm not sensible to anomalies whose mean intensity is under a threshold determined by the μ_{min} value. This is a desirable feature since 1. in real applications, the background estimates are affected by uncertainties; and 2. many known fixed, gamma-ray astrophysical sources (e.g. Crab nebula) appear as faint, nearly constant bright sources with rise-times comparable to GRB duration (see for example pg. 79 of [68]). The ordering of the memory state makes implementing μ_{min} and t_{max} cut techniques straightforward. It is enough to test the oldest tuple in the memory state for its non-trivial root or t component against the μ_{min} or t_{max} threshold, prior to each FOCuS iteration. This operation requires a constant amount of computational resources for a reasonable memory state implementation (e.g, linked lists). With FOCuS Poisson, a μ_{min} cut can be achieved without directly estimating the non-unity root of a curve. In fact, this value depends only on the ratio between the a and b curve component and a tuple (a, b, t) may be cut whenever:

$$\frac{a}{b} = \frac{1 - \mu_{min}}{\log(\mu_{min})}$$

Note that the quantity on the RHS may be computed once and for all; there is no need to compute $\log(\mu_{min})$ at each algorithm step.

Another compromise may be taken to further improve FOCuS computational efficiency at expense of the algorithm sensibility. At each algorithm iteration one can test for a maximum only a subset of curves in the memory state. For FOCuS Poisson for example, one could actually maximize only those tuples (a, b, t) whose t parameter equals a power of two, a condition which can be tested efficiently. We will refer to this technique as log-grid maximization. A visual representation of the operation of log-grid FOCuS Poisson is proposed in C.5.

4.2 Background Estimate

The discussion of TAs of this chapter has been unconcerned so far with the problem of how background count rates are actually estimated. In particular, in Sec. 4.1.1, it was assumed that a reasonable estimate of the number of counts expected from background λ_i was available at any given time-step index i. In this section we will lift this hypothesis and consider the problem of assessing an estimate of the background against which compare the detector count observations.

Background modulations The physical sources of background in a LEO, near-equatorial orbits and their expected contributions to the HERMES scientific background have been discussed already in Sec. 2.5. The orbital background is not constant. For example roughly periodic modulations exist on timescales equal to the duration of an orbit or the duration of a day due to features of the orbital geomagnetic environment, the daily orbit precession and the activation resulting from SAA passages. In Low-Earth Orbit, the entity of the background modulation depends strongly on the orbital inclination parameter. Near-equatorial orbits span a restricted range of geomagnetic latitudes and grazes the high radiation environment of the SAA only in its outermost regions. This results in minimal background modulation, mostly due to the detector orientation relative to Earth and the residual variation of the geomagnetic field [33]. On the other hand, the background count rate of a spacecraft crossing the SAA in depth (cf. Fig. 3.9b) is affected by the nuclear activation induced by energetic protons on the spacecraft detector and service module. In Fig. 4.8, the percentual variation in background count rate observed by BeppoSAX-PDS along 15 consecutive orbits is plotted, folded with respect to the orbital phase. The BeppoSAX orbit had a low inclination parameter (4 degrees) hence background modulation was small, typically under 10% of the daily mean value outside the SAA. The Fermi orbit, on the other hand, has inclination 20 degrees and the GBM background count rate reaches excursion over 40% after deep SAA passages. In Fig. 4.7a and 4.7b, the background count rates outside SAA and over one day of Fermi-GBM and BeppoSAX-PDS operations are plotted.

Moving average techniques The approach of many GRB monitor experiments (see the discussion of Sec. 1.6) to automatically assess an estimate of the count rate expected from background sources can be reduced to a simple moving average (SMA). In simple moving average, past data are summarized by the unweighted mean of a number n of previous data points. Using the same notation of Sec. 4.1.1, the general expression for a moving average of length n, at time-index t is:

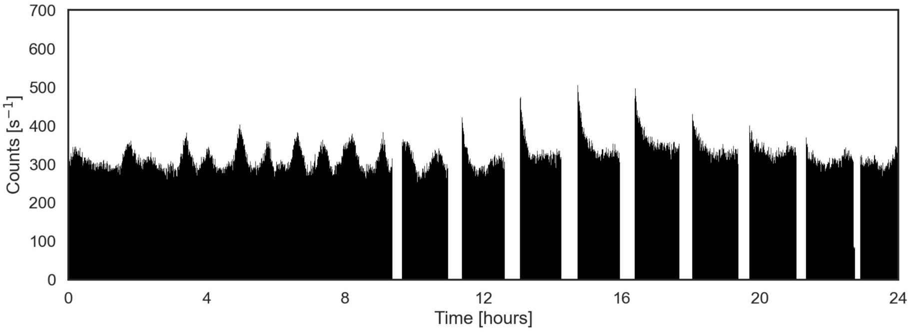
$$M_{t,n} = \frac{x_{t,n}}{n} \tag{4.17}$$

A SMA can be computed efficiently through the recursion:

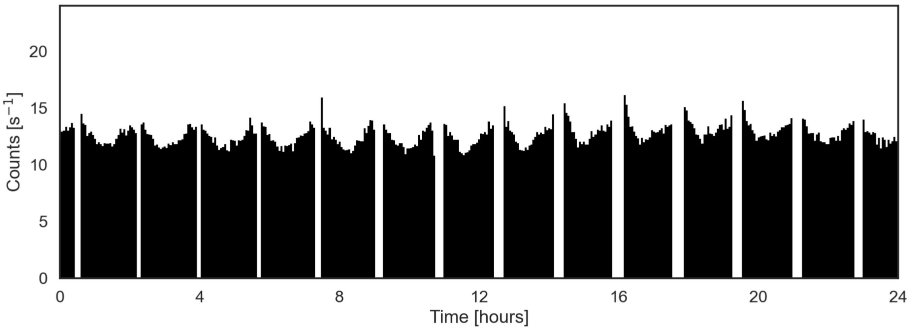
$$M_{t,n} = M_{t-1} + \frac{x_t - x_{t-n+1}}{n} \tag{4.18}$$

A SMA can be viewed as a low-pass filter with a cut-off frequency inversely proportional to n, making SMA suitable for assessing background estimates in GRB detection, provided that n is larger than the GRB rise time. The computation of a SMA can be delayed to reduce ‘pollution’ from anomalous data. For example, this is the approach of Fermi GBM which estimates the background count rate over a period nominally set to 17 s and excluding the most recent 4 s of observations.

Simple moving average techniques are generally not able to handle significant



(a) Fermi GBM NaI-6 detector background count rate during 2011 April 20th in band 50-300 keV. Fermi orbits Earth at altitude ~ 550 km and inclination ~ 25 degrees. Data from NASA HEASARC GBM daily data database [58].



(b) BeppoSAX-PDS background count rate during 1998 December 31th in band 20 - 200 keV. BeppoSAX orbited Earth at altitude ~ 600 km and inclination ~ 4 degrees. The archival BeppoSAX-PDS data are available for download at NASA HEASARC [57].

Figure 4.7: Trends in background count rates. Count rates during SAA passages are not plotted.

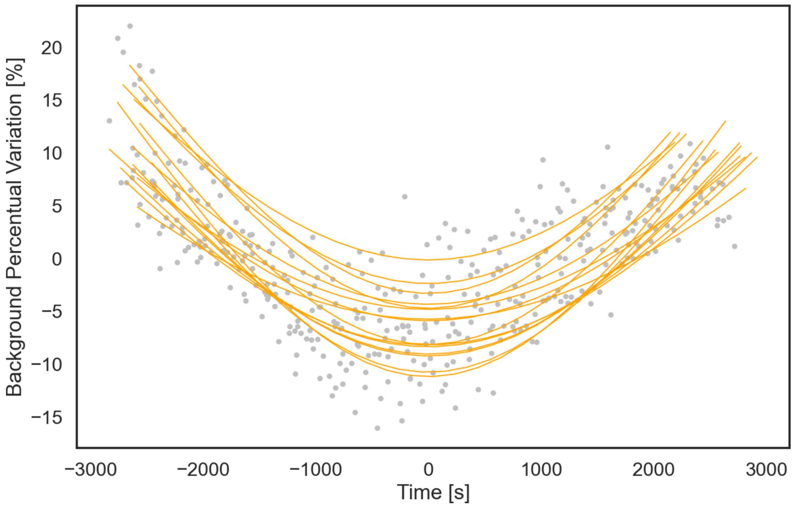


Figure 4.8: Percentual orbital variation relative to the daily mean and best quadratic fits for 15 consecutive BeppoSAX-PDS orbits (see Fig. 4.7b).

trends in data. In presence of trends, better suited moving average techniques do exist. Exponential smoothing techniques for example assign exponentially decreasing weights to the observations, as the observations get older [6]. At a given time-index t, the single exponential smoothing (SES) value S_t can be computed through:

$$S_t = \alpha x_t + (1 - \alpha) S_{t-1} \tag{4.19}$$

where $0 < \alpha \leq 1$ is the smoothing constant parameter. The SES weights assigned to each observation decrease geometrically and their sum is equal to 1. The choice of the α parameter depends on the features of the data and is generally achieved optimizing for the mean squared error (MSE) over a grid of parameters on a test dataset. When α is close to 0 more weight is given to older observations, resulting in slower dampening and smoother S_t curves. Conversely, α value near unity will result in quicker dampening and more jagged S_t curves. See for example Fig. 4.9. Computing the SES variable S_t requires the setting of an initial value S_0. No unequivocal recipes exist for this purpose and possible initialization methods includes setting S_0 to x_0, to an a priori estimate of the process target or to the average of a number of initial observations [6].

In presence of strong trends the performance of exponential smoothing techniques may benefit from the introduction of a second constant, β. The resulting technique is known as Double Exponential Smoothing (DES) and requires the

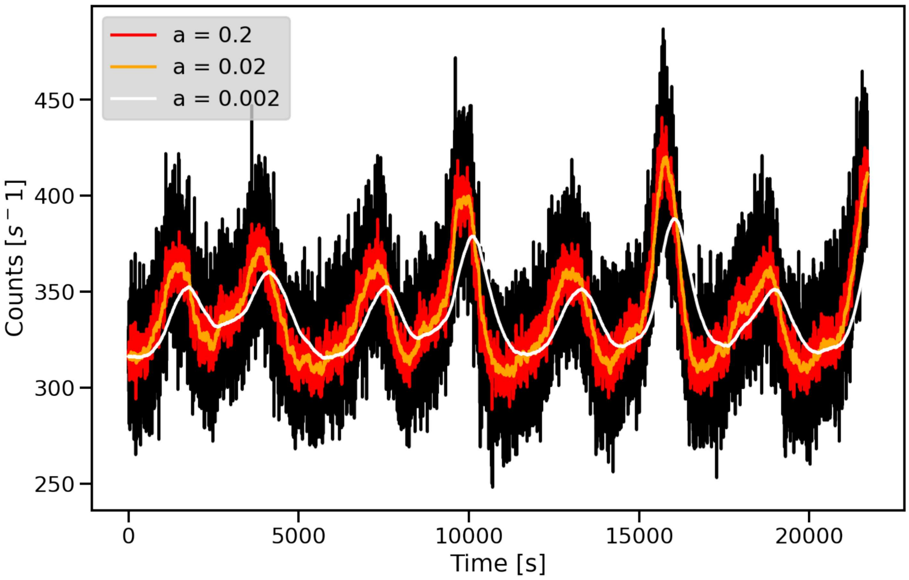


Figure 4.9: Different single exponential smoothing approximations of Fermi-GBM count data.

computations of two quantities, S_t and b_t, according to:

$$\begin{aligned} S_t &= \alpha y_t + (1-\alpha)(S_{t-1} + b_{t-1}) \\ b_t &= \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1} \end{aligned} \tag{4.20}$$

where $0 < \alpha \leq 1$ and $0 < \beta \leq 1$. The first smoothing equation adjusts S_t for the trend estimate observed during the previous iteration b_{t-1}. Again, the best parameters choice depends on the feature of the data themselves and is generally achieved through optimization techniques over archival data. The same consideration on the initialization of α we made for SES holds true for DES. Common initialization techniques for the trend parameter include setting b_0 to $x_2 - x_1$ or an average of the differences between initial subsequent pairs of observations. Linear extrapolation allows for a natural forecasting techniques based on DES. At step index t, the m-step ahead forecast is given by:

$$F_{t+m} = S_t + mb_t \tag{4.21}$$

See Fig. 4.10 for an example application of double exponential smoothing forecast to data from the Fermi-GBM dataset.

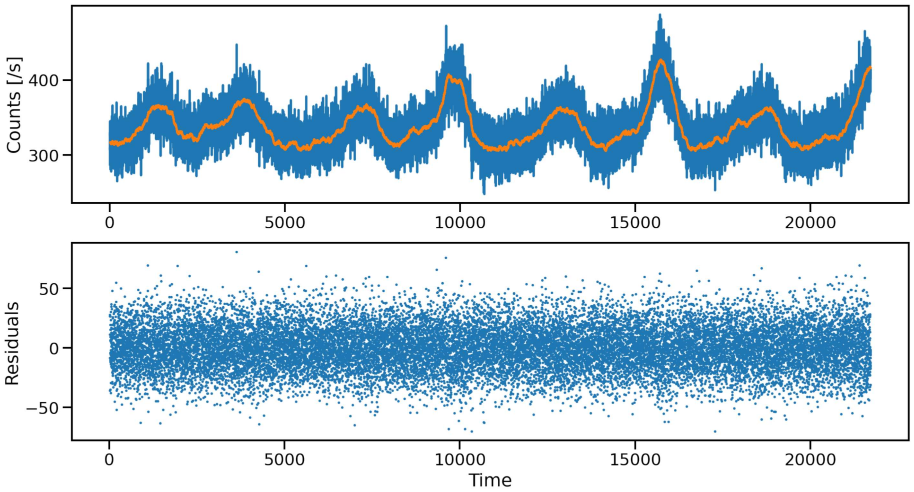


Figure 4.10: Double exponential smoothing forecasts (orange solid line) and residuals over 6 hours of Fermi-GBM count data (blue line). Data from 6 hours of observations acquired during 20 April 2019 from the GBM NaI-0 detector in band 50 - 300 keV. Double exponential smoothing parameters $\alpha = 0.02$ and $\beta = 0.01$, forecast step-ahead $m = 4$s.

4.3 Trigger algorithms tests

The performances of four different algorithms have been tested and compared. These tests addressed three performance metrics: detection efficiency, computational efficiency and false positive rates. The algorithms we tested are:

1. An exhaustive, true-background TA (labeled 'True') designed to provide a standard reference for the detection efficiency. This algorithm sequentially computes the significance for all possible signal sections, compatibly with a fundamental binning choice, until a threshold condition is met. For a discussion of the algorithm logic, the reader may refer to Sec. 4.1.3. Significance is computed according to Eq. 4.1. In significance computation, the background count-rate is not estimated from the data but provided by the user. For this reason, the 'true' algorithm will be only tested on synthetic data, for which the exact background count-rate is known. The algorithm is controlled by two parameters, a threshold value T and a background value B. An optional parameter h_{max} can be specified to force significance computations up to a limit time-scale.

2. A true-background implementation of FOCuS Poisson (labeled 'FOCuS'). FOCuS only tests the significance for those signal sections which may actually result in a trigger. This is achieved storing information from previous observations in a dynamic memory state, see the discussion of Sec. 4.1.8. As with the 'true' TA, this algorithm requires the user to

provide a background count-rate. The algorithm is controlled by two parameters, a threshold value T and a background value B. An optional μ_{min} parameter can be specified (cf. Sec. 4.1.9).

3. Approximate FOCuS Poisson with single exponential smoothing background estimate (labeled 'FOCuS-AES'). The core algorithmic logic of this TA is an approximate version of FOCuS Poisson implementing both μ_{min} cut and log-grid maximization. The μ_{min} cut limits the growth of the FOCuS memory state. Log-grid maximization implies that a curve is evaluated for a maximum only if its age parameters equals a power of 2. Both these techniques were discussed in Sec. 4.1.9. At start-up, FOCuS-μDES stays in 'sleep' mode for an adjustable amount of iterations. In sleep mode, counts are monitored to assess an initial estimate of the background count-rate and the trigger logic is not operated. Once out of sleep mode, the DES main variable S_0 is initialized to an average of the counts observed up to that moment, and the trigger operations start. At each iteration, the background count-rate is forecasted according to Eq. 4.19, excluding from the estimate an adjustable amount of recent counts. In total, the algorithm is operated by 5 parameters: a threshold value T; the SES smoothing parameters α; a forecast delay parameter defining the duration of the sleep mode `sleep`, a parameter m setting the amount of recent counts to exclude from background estimate; and a μ_{min} value. An extra parameter can be specified to override default initialization and provide a value for S_0.

4. A conventional trigger algorithm with SMA background estimate (labeled 'Parametric'). This algorithm was designed to emulate the behaviour of the Fermi-GBM TA in the energy band 50-300 keV, see Sec. 1.6 and Sec. 4.1.2. The count-rate expected from the background is estimated according to Eq. 4.18, excluding an adjustable amount of recent observations. The algorithm operations start only after enough background counts have been collected. Nominally, the algorithm checks 9 logarithmically equispaced timescales. For a fundamental binning $\tau = 0.016$ s these timescales are equivalent to 0.016 s, 0.032 s, 0.064 s, 0.128 s, 0.256 s, 0.512 s, 1.024 s, 2.048 s and 4.096 s. For all but the shortest timescale, checks are scheduled with phase offset equal to half the accumulation length (e.g., a timescale with characteristic length 4, is checked 2 times over 4 subsequent iterations). The significance scores are computed according to the Wilks equation Eq. 4.12 and only if an excess in counts is observed in first place (i.e. if $n > b$). The choice of the Wilks condition for significance computation was driven by two facts: 1. to our knowledge, the exact

recipe implemented by the Fermi flight-software to compute significance scores is not described in the literature; 2. the use of the same formula employed by FOCuS-AES provides desirable comparability when testing computational performances. The algorithm is operated by 3 parameters: a threshold T; a background duration parameter n setting the characteristic length of the SMA estimate; a forecast delay parameter m setting the amount of recent counts to exclude from background estimate.

Many of the tests we performed were executed over synthetic data. A software package called SynthBurst was specifically developed to generate these data. In essence, SynthBurst is a Monte Carlo tool for creating lists of time-energy tagged photon events resembling those produced by the observations of real-world experiment for GRB observation (e.g., Fermi-GBM). SynthBurst synthetic light-curves are highly customizable by the user which can set a wide array of parameters controlling features such as the background mean intensity and time-profile, the light-curve duration and so on. Most importantly, SynthBurst allows to generate photon lists as if they were generated by a GRB source. A GRB source is modelled after real world observations of Fermi-GBM and every GRB in the Fermi catalogue can be used as a template. The number of source photons generated according to a given template is selected by the user. This allows the user to generate source events with different brightness values. Moreover, background and source photons can be generated dynamically ('in steps') allowing for example to generate multiple lightcurves with equal background events but different source intensities. Even if in the next section we will focus only on the temporal properties of a given lightcurves, SynthBurst allows the user to provide analytic spectra for both background and source photons. The response of different detectors can be simulated providing the instrument standard ARF and RMF matrices [3].
The code of SynthBurst and all the algorithm presently tested has been made available online[1].

4.3.1 Detection performances

The detection performances of different TA were tested and compared over two synthetic GRB lightcurve datasets generated with SynthBurst. Each dataset is composed of 1440 lightcurves, all with equal duration and mean background count-rate. The background count-rate is constant for each lightcurve duration and equal to 350 s^{-1}, a value representative of the background flux observed by

---

[1] `https://www.dropbox.com/sh/ca09n1e2dhgou3g/AAB03SeQj_SyMzXdGuo7Fbswa?dl=0`
If you encounter problems using this repository, please feel free to contact me at peppedilillo@gmail.com

the Fermi GBM NaI detectors in the energy band $(50, 300)$ keV. Each lightcurve hosts a simulated GRB event. Across one dataset, the intensity of the GRB event was modulated over 30 different steps. For each source intensity step, different lightcurves were generated for a total of 48 repetitions. For all synthetic lightcurves we generated, plots have been added to the online repository of this thesis. The simulated GRBs hosted by the lightcurves in the first dataset were modelled after the Fermi-GBM observations of the short burst GRB180703949 (see Fig. 4.12a), while the simulated GRB events of the second dataset were modelled after the Fermi-GBM observations of the long burst GRB120707800 (see Fig. 4.12b). Both the burst models were selected for their very high photon flux. It is impossible to justify this choice without discussing how SynthBurst works under the hood. To generate a synthetic lightcurve resembling a real model event, SynthBurst first subtracts a background estimate from the real event observations. The goal of this operation is to obtain a histogram representative of the model's temporal evolution, the *silhouette*. The model's silhouette is then normalized to obtain the actual discrete probability distribution used by the Monte Carlo inverse transform sampling responsible for generating the synthetic burst events. After background subtraction, a shot noise component due to background statistical fluctuation still persists in the model silhouette. Using model events with very high signal-to-noise ratio allows to minimize the spurious contributions of background shot noise to the model silhouette, hence to the final synthetic data.

For all lightcurves, time series of counts were formed with a GBM-like fundamental binning equal 16 ms. Trigger algorithms were set up with the following parameters:

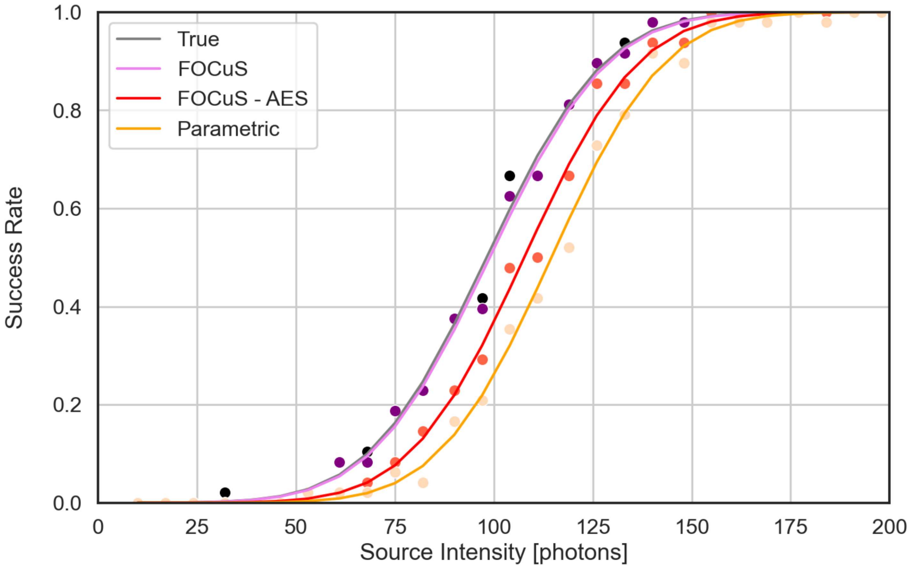
1. True: threshold $T = 5.5\sigma$ and background value $B = 350\ \mathrm{s}^{-1}$.
2. FOCuS: threshold $T = 5.5\sigma$ and background value $B = 350\ \mathrm{s}^{-1}$.
3. FOCuS-AES: threshold $T = 5.5\sigma$, SES smoothing parameter $\alpha = 0.02 \cdot 0.016$, `sleep` mode equivalent duration 16.992 s, forecast delay `m` equivalent duration 4 s and μ_{min} value equal 1.3.
4. Parametric: threshold $T = 5.5\sigma$, background equivalent duration equal 16.992 s and forecast delay `m` equivalent duration 4 s. Note that these parameters were set in such a way that the parametric TA would emulate the well-understood Fermi-GBM trigger algorithm.

Each algorithm was run on both datasets. A trigger algorithm is said to have achieved a correct detection if it met its threshold condition over a section of the lightcurve hosting the simulated source. In Fig. 4.11 the fraction between

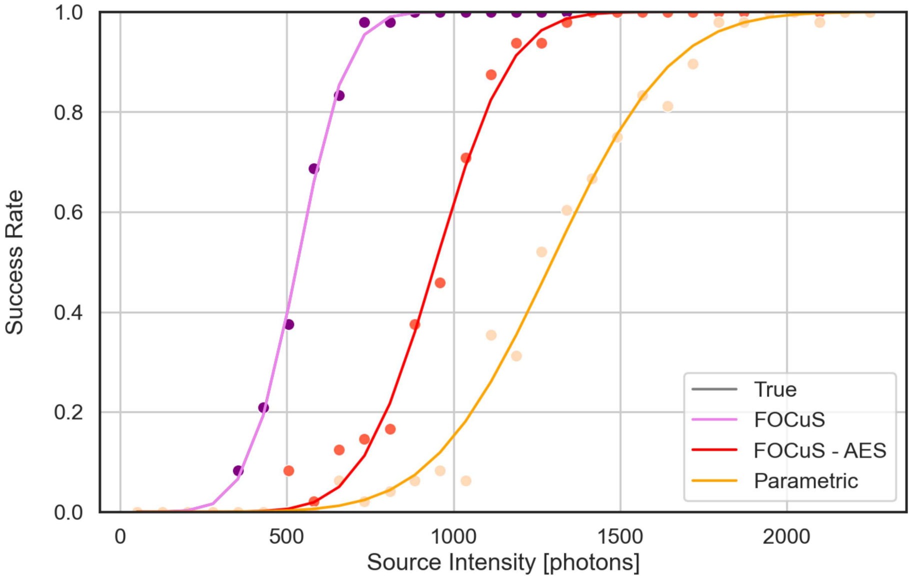
the number of correct detections and the total number of simulated lightcurve value is plotted against the source intensity level, expressed as the number of simulated source photons and for each dataset.
The performances of the true and FOCuS TAs were identical over the long bursts dataset. FOCuS performed slightly worse than true over the short bursts dataset. The discrepancies observed over the long dataset were ultimately due to the different way in which the two algorithms assess count excess significance. FOCuS - AES had better success rates than the GBM-like Parametric TA over both dataset with significant larger margin over the long dataset. Over the long burst dataset, FOCuS-AES was able to detect 855 of the 1111 events which triggered the ideal exhaustive search algorithm (81%), while the number of events observed by Parametric was 625 (63%). Over the short burst dataset, FOCuS-AES was able to detect 770 of the 834 events detected by the exhaustive search algorithm (92%), while Parametric found 724 (86%) The higher detection efficiency of FOCuS-AES over the long dataset is easy to interpret. FOCuS-AES is able to evaluate count significance over timescales exceeding the maximum timescale tested by the parametric TA, which equals 4.096 s. This allows FOCuS-AES to evaluate the simulated burst significance over intervals better fitting the length of the burst itself. On the other hand, over both tested datasets, the detection performances of FOCuS-AES were worse than those of the true and FOCuS algorithm. This fact is motivated by three reasons. First, the true and FOCuS TAs had access to the true background count-rate value, while FOCuS-AES had to assess a background estimate from the same data it was testing. Secondarily, the approximation introduced by the μ_{min} cut techniques makes it possible for changepoint which would have otherwise resulted in a trigger to be discarded prematurely. Finally, log-grid maximization can result in a trigger being missed due to the curve maxima not being actually computed. No false detections were observed in these tests.

4.3.2 Computational performances

The computational performances of the FOCuS-AES and parametric TAs were tested over 1000 random generated count series. Each count time series consisted of 10^5 samples drawn from a Poisson distribution with rate parameter $\lambda = 350 \cdot 0.016 = 5.6$. All but the threshold TA parameters were set as discussed in Sec. 4.3.1. Threshold parameters were set to an unreasonably large value to avoid eventual false positives to interrupt the simulation prematurely. A box plot of the time needed to the FOCuS-AES and Parametric TAs to run over a single count time series is presented in Fig. 4.13. On average, FOCuS-

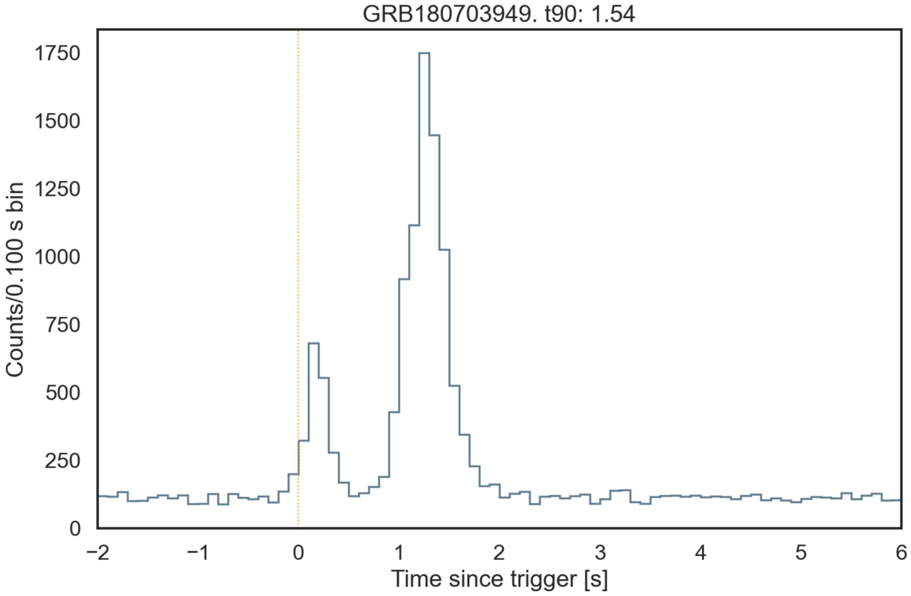


(a) Detection performances over a dataset of synthetic lightcurves hosting a source event modelled after the short Fermi-GBM burst GRB180703949.

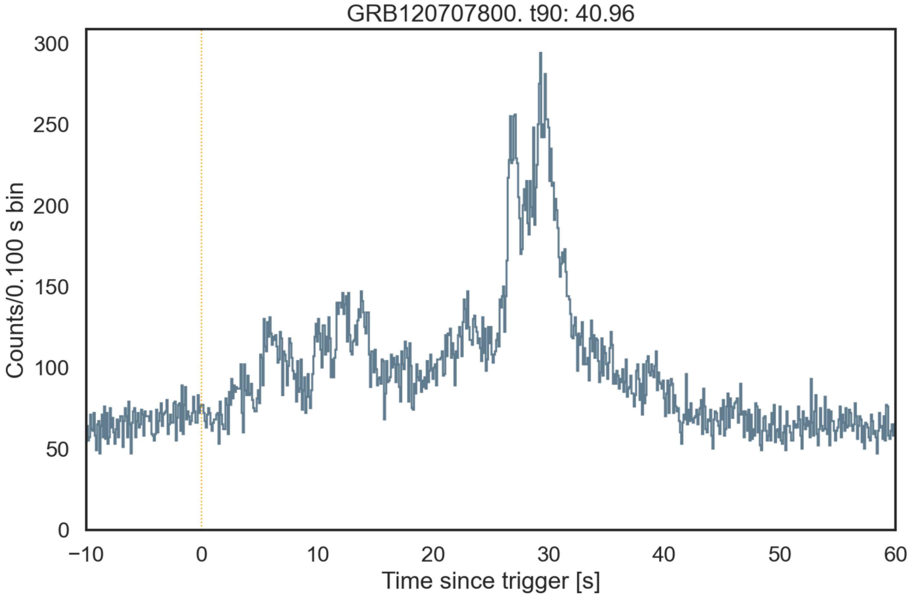


(b) Detection performances over a dataset of synthetic lightcurves hosting a source event modelled after the long Fermi-GBM burst GRB120707800. Results from FOCuS and true TA were equal and are superimposed to each other.

Figure 4.11: Detection performances of different trigger algorithms over two datasets of synthetic GRB lightcurves. The fraction between the number of correct detections and the total number of simulated lightcurves value is plotted (points) against the source intensity level, expressed in terms of simulated source photons. For each trigger algorithm best fit to the error function is displayed (solid lines).



(a) Histogram of photon counts obtained merging the time-tagged event observations of GRB180703949 by Fermi GBM triggered detectors NaI-0, NaI-1 and NaI-3 in band (50, 300) keV. Time is plotted in seconds since the Fermi trigger time.



(b) Histogram of photon counts obtained merging the observations of GRB120707800 by Fermi GBM triggered detectors NaI-8, NaI-11 in band (50, 300) keV. Time is plotted in units of seconds since the Fermi trigger time.

Figure 4.12: Gamma-ray burst models used as templates for generating lightcurves to test the sensibility of different trigger algorithms.

AES took 22% less time to run over a count time series than the Parametric TA. Similar margins were observed for the first and third quartiles.
The algorithms we tested were implemented in Python 3.6 and tested on a desktop computer equipped with a 3.9 GHz processor. Both Parametric and TA implementations were developed with no external dependencies outside the Python standard library.

The details of an implementation play a fundamental role in computational performances tests. Whenever possible we tried to preserve comparability between FOCuS-AES and Parametric. Hence, both algorithms have been set in such a way to refresh their background estimate at the same rate and to compute significances by equivalent routines. A number of optimization precautions have been taken. For Parametric we made sure that the background SMA was computed through the recursive equation Eq. 4.18 and that the significance values were assessed only if an excess in counts was actually observed. Regarding FOCuS-AES, most of the curve arithmetic functions were embedded in the FOCuS basic step to reduce the overhead costs associated to function calls. For both programs, computationally expensive arithmetic operation (e.g., square roots, divisions) were avoided when possible. Within the limit of the implementation language, we tried to use optimal data structures. FOCuS-AES does not require access to any of the elements in the count buffer but those at the buffer ends, hence a deque data structure was employed [4]. On the other hand, since Parametric requires access to inner elements of the observation buffer a standard list was employed. We experimented with using a double buffer with mixed data structures for storing count observations (a list for the foreground section of the buffer and a deque for the background) but no benefits to the computational performances were observed.
The performances of Parametric can be improved restructuring the foreground section of the count buffer in such a way that a cumulative record of the observed counts is stored instead of the counts observed over each time bin. In conjunction with techniques to manage integer wrap-around when a large number of count is collected, this would avoid summing over many buffer elements at each algorithm iteration. The usage of this technique has been documented in literature [42]. We did not take this precaution. However we report that in our tests and on average, Parametric spent an amount of computational time $< 10\%$ summing buffer counts. Most of the computational time (~ 40 %) is spent in significance computations. On the other hand we report on the sub-optimal nature of the data structure used to implement the curve memory state of FOCuS-AES. Given the fact that FOCuS access the data cell in its memory state in a consistent order, a linked list would have made for an ideal candi-

date. However, an efficient linked list implementation is currently not provided within the Python standard library and user-defined solutions have been found to perform no better than a standard list.

Python is a dynamically typed, interpreted language, with complex memory management routines [7]. Ultimately, it is a poor choice for testing the performances of algorithms designed to be fast. This choice however was forced by the lack of a reliable C implementation of the Parametric TA, a problem we plan to fix in the near future.
A C implementation of FOCuS-AES—or more precisely, an implementation of FOCuS Poisson with double exponential smoothing background estimate—does in fact exist. This implementation was developed specifically for HERMES miniaturized on-board computer during a brief research period at Eberhard Karls University of Tübingen. Code for this implementation (provisionally named 'Caduceus', as the staff of the mythological Hermes) is available in the thesis repository. Using the same 3.9 GHz CPU which executed the tests of Fig. 4.13, we report on user time requirements under 0.3 s to run a parametrization equivalent to FOCuS-AES over one million counts i.e., computing times better than the Python implementation by a factor of ~ 100.
Beside common sense, no particular optimization precautions were taken with this implementation. We believe that a large margin for improvement still exists in what regards developing a hyper-efficient implementation of FOCuS. For example, the fact that the number of curves in approximated versions of FOCuS is probabilistically bounded can be exploited to implement the algorithm memory state using a circular, fixed-size array, hence avoiding the need for garbage data collection (obsolete curves). Given the facts that a curve is essentially described by one float and two integers and that even exact implementation requires no more than 15 data cells after 10^6 iterations [98], the memory requirements of such an optimization would be minimal. Furthermore, for the normal version of FOCuS, the quadratic update step can be achieved with $O(1)$ amortized computational costs [98]. The application of similar solutions to FOCuS Poisson has yet to be investigated.

Due to their exact nature, the FOCuS and True TAs have larger computational cost than FOCuS-AES and Parametric. To compare their performances we generated 8 datasets, each composed of 100 time series of counts with different lengths 100, 200, 400, 800, 1600, 3200, 6400 and 12800. All counts were random generated sampling a Poisson distribution with parameter $\lambda = 350 \cdot 0.016 = 5.6$. FOCuS and True were run with a very large threshold on each dataset and the time needed by each TA to run over every time series was recorded. The results

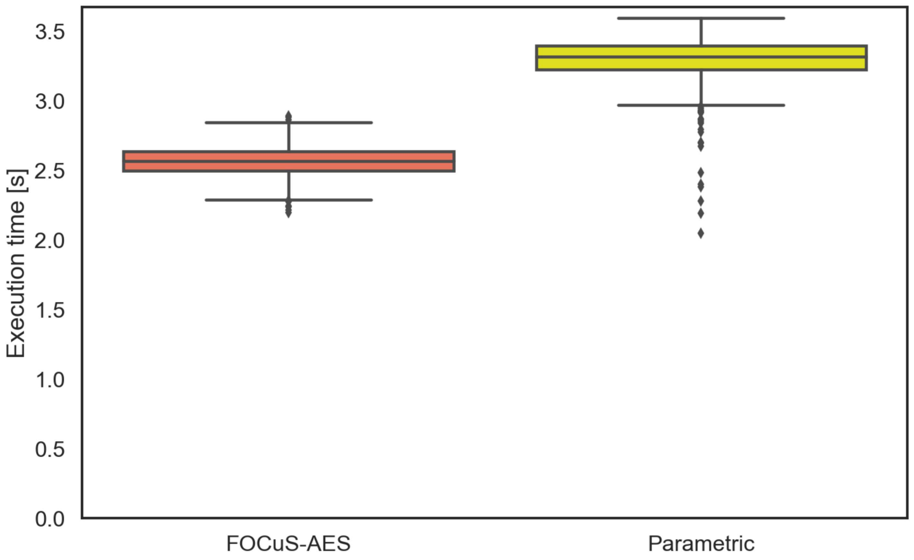


Figure 4.13: Box plot of the time needed to run over a time series of 10^5 random generated counts sampled from a Poisson distribution with rate parameter $\lambda = 5.6$. Results obtained using a 3.9 GHz processor and Python implementations of both algorithms.

of this test are presented in 4.14. Even if FOCuS and True have almost equal success rates (see Sec. 4.3.1), FOCuS proved to be order of magnitudes faster than True over all datasets. This result is expected, since the computational complexity of the True TA grows quadratically with the number of observations, while the growth in FOCuS computational cost is slower (cf. the discussion of Sec. 4.1.3 and Sec. 4.1.8).

4.3.3 Average run length

For a given TA, the average run length (ARL) is defined as the average number of algorithm iterations before a trigger when no anomaly exists in the data. The ARL quantifies how susceptible to type-I errors an algorithm is. The more sensible an algorithm, the smaller its ARL is. The ARL of the Parametric and FOCuS-AES TAs has been studied. For this purpose, a stream (a count time series of arbitrary length) of Poisson distributed count data with mean rate $\lambda = 5.6$ was implemented. The FOCuS-AES and Parametric TAs were launched over 1000 streams of data. The same parameters settings outlined in Sec. 4.3.1 were used, the exception being the threshold value which was set to 5.0σ for both algorithms. For each data stream, the stopping iterations of both algorithms was recorded. A histogram of the run-lengths observed in these tests is plotted in Fig. 4.15. The ARL observed for the Parametric TA was $1.3 \cdot 10^6$.

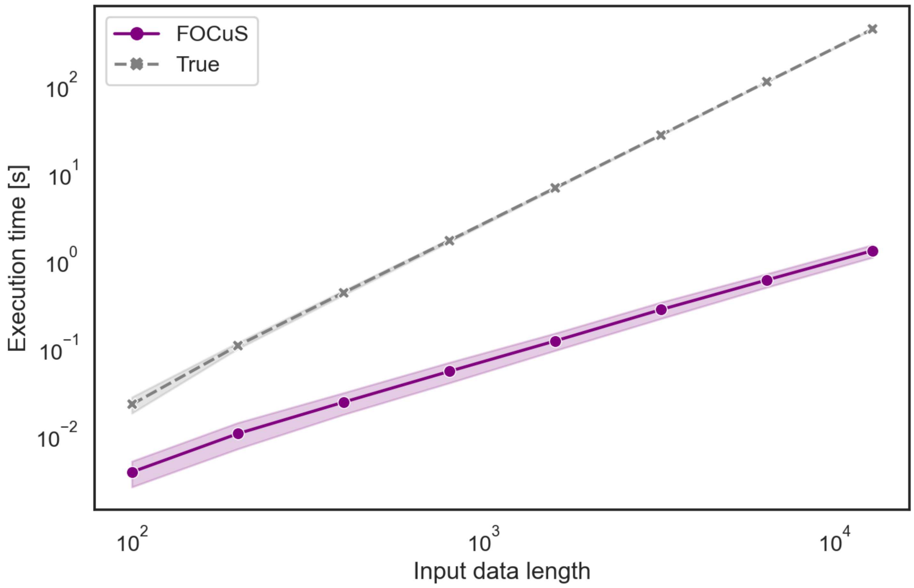


Figure 4.14: Average time (dots,crosses) and standard deviations (bands) observed running FOCuS and True TAs with infinite threshold over a count time series of different length. Results were repeated over 100 different random generated time series per input data lenght. Time series were random generated sampling a Poisson distribution with rate parameter $\lambda = 5.6$. Results obtained using a 3.9 GHz processor and Python implementations of both algorithms.

A value 38% smaller and equal to $9.3 \cdot 10^5$ was observed for FOCuS-AES.

Statistical fluctuations in Poisson distributed data can cause false triggers, the rate of which can be reduced lowering the TA sensibility e.g., raising the threshold parameter of a TA. Similar precautions however may not protect from false triggers due to a lacking background estimate or other phenomena which can not be explained by Poisson statistics (see for example the discussions of [42] and [44]). To investigate these points, testing over real data is crucial. For this purpose, we selected 22 hours observations from one of the Fermi-GBM NaI detectors, see Fig. 4.16. Over this observation period no events triggered the Fermi-GBM flight software. Counts from this dataset were formed with fundamental binning 0.016 s. Data were preprocessed removing 5 s of observations before and after each detector turn-off at SAA entering. This precaution allowed to remove gradients in count data which were observed at these times, presumably due to the finite time needed to the instrument to become fully operative. The Parametric and FOCuS-AES TAs were launched with the same settings discussed in Sec. 4.3.1 and the observed maximum in excess significance recorded at each algorithm iteration. In Fig. 4.17 the per-iteration maxima observed by the FOCuS-AES and Parametric TAs are plotted against each other. The observed maximum significance in count excess was 3.7σ for FOCuS-AES

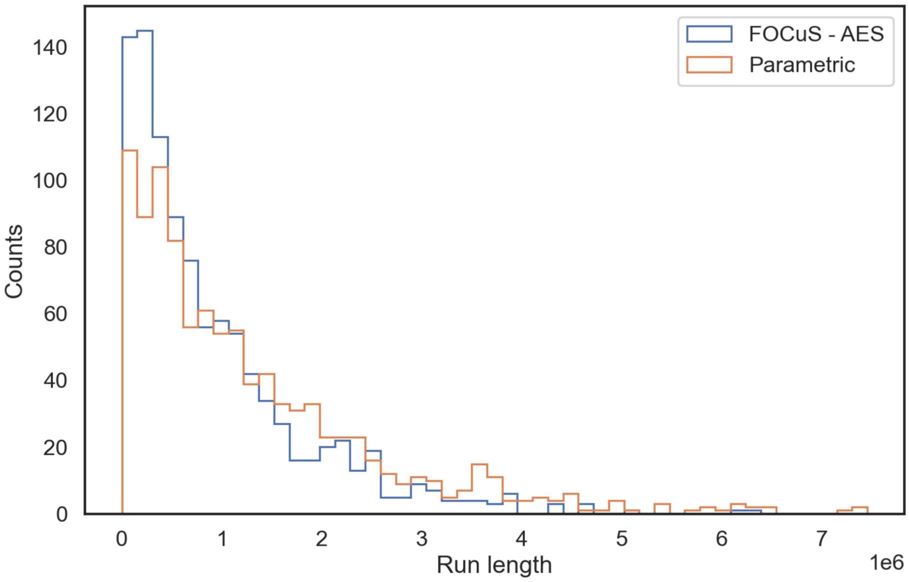


Figure 4.15: Run lengths of FOCuS-AES and Parametric TAs over 1000 Poisson count data streams with mean rate $\lambda = 5.6$. Both algorithms thresholds were set to 5.0σ. All other parameters were set as described in 4.3.1.

and 3.6 for Parametric. Given the amount of tested count data (over 4 millions) and the ARL values we reported for a homogeneous Poisson process at threshold 5σ, these values seem small and suggest the existence of some mechanism dampening the count data tail probability distribution. The nature of this mechanism is unclear.

The last two hours of observations were excluded since during this time a GRB event (GRB190220981) was observed by Fermi-GBM NaI-0 detector. Both algorithms were able to detect this event with significance equal 6.4σ for FOCuS-AES and 6.1σ for Parametric.

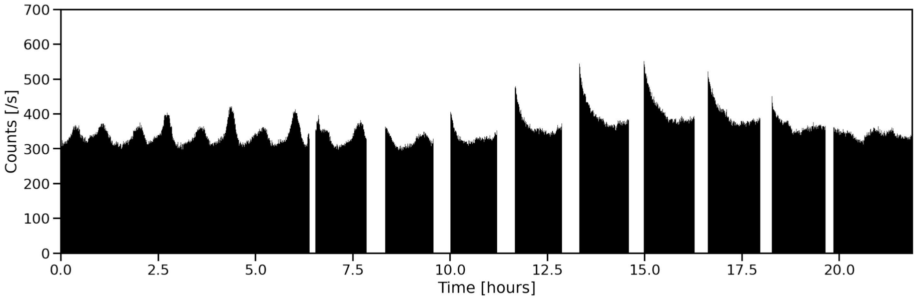


Figure 4.16: Fermi GBM NaI-0 detector background count rate during the first ~ 22 hours of 2011 April 20th in band 50-300 keV.

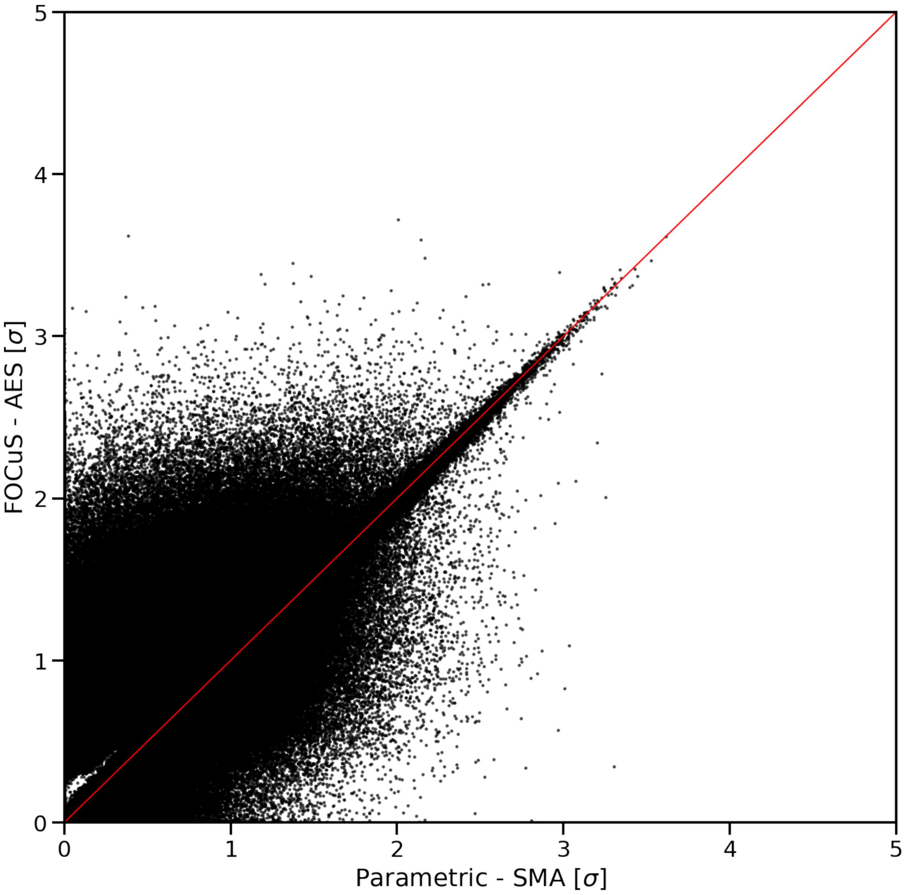


Figure 4.17: FOCuS-AES versus Parametric maximum significance per iteration as observed over the dataset of Fig. 4.16.

Conclusions

Two original studies relevant to the hardware and software development of the HERMES mission have been presented in this thesis.

The first work concerned an investigation of GAGG:Ce scintillation crystal phosphorescence emission. The goal was to determine whether the delayed luminescence caused by the interaction between the scintillator and the energetic particles of the near-Earth radiation environment could pose a threat to the well-functioning of the detector in its lifetime in space. For this reason, we conducted an irradiation campaign in which a GAGG:Ce sample was irradiated with 70 MeV protons at dose levels representative of those expected from space operations. Starting from the experimental data we gathered, we developed an original model of GAGG:Ce afteglow emission. I used this model, in conjunction with the IRENE AE9/AP9 models of the near-Earth radiation environment, to predict that the equivalent leakage current induced by GAGG:Ce afterglow on a HERMES SDD detector cell will not exceed 1.0 pA, for a baseline orbit at altitude 550 km and inclination 10 degrees. Since the HERMES low-noise front-end electronics guarantees nominal performances up to ~ 100 pA of leakage current, GAGG:Ce afterglow emission should not endanger the operations of the HERMES TP/SP detectors.

The second work inquired algorithms for detecting gamma-ray bursts. Several algorithmic strategies were discussed, including the application of FOCuS, a novel changepoint detection technique developed by researchers of the University of Lancaster Department of Statistics, with whom we collaborated. Exact implementations of FOCuS enables the design of an efficient trigger algorithm whose sensibility matches that of a logic checking over all possible timescale parameters. The use of similar changepoint detection techniques has been advocated in the past [103]. Yet, this is the first time that similar strategies are thoroughly evaluated for applications to GRB detection.
I generated two synthetic datasets of 1440 synthetic lightcurves hosting GRB

events modelled after real observations of a short and a long burst. Over these datasets I found almost identical detection success rates between exact implementations of FOCuS and an ideal exhaustive search algorithm. In these tests both algorithms had access to the true background value. In a real scenario however, a background estimate must be inferred from the same data which are tested for a burst. I inquired the usage of non-conventional moving average techniques for assessing background estimates and implemented an approximated version of FOCuS with background estimates obtained by means of single exponential smoothing (FOCuS-AES). I tested the performances of FOCuS-AES on the aforementioned synthetic GRB datasets against that of benchmark trigger algorithm modeled after the logic employed by Fermi-GBM (Parametric). Over the long burst dataset, FOCuS-AES has proven able to detect 855 of the 1111 events which triggered the ideal exhaustive search algorithm (81%), while the number of events observed by Parametric was 625 (63%). Over the short burst dataset, FOCuS-AES was able to detect 770 of the 834 events detected by the exhaustive search algorithm (92%), while Parametric found 724 (86%).

No false detections were observed during these tests. I performed specific tests to address the average run length of the FOCuS-AES and Parametric TA prior to a false positive at 5σ threshold. In these tests the average run lengths of FOCuS-AES was 38% smaller than Parametric. Additionally, I launched the FOCuS-AES and Parametric algorithms over one day of Fermi-GBM data from a single NaI detector. Within 5σ thresholds, no false positives were observed. Both algorithm were able to correctly detect the only GRB event in this dataset with significance levels exceeding 6σ.

Finally I tested the computational performances of FOCuS-AES and Parametric. FOCuS-AES took on average 22% less time than Parametric to run over one thousand time series of 10^5 randomly generated count data. The extent of this result is limited by the choice of the implementation language, which in turn was driven by the lack of a low-level, reference implementation of the Parametric algorithm. A baseline C implementation of FOCuS-AES apt for operations on the HERMES-TP/SP on-board computer exists, which I developed during a research period at Eberhard Karls University of Tübingen.

Appendix A

Detailed afterglow model derivation

To build a viable model of the radiation induced afterglow we made some simplifying assumptions motivated by the fact that our observations are largely unable to account for the details of the physical processes at play.

1. Since we started observing the luminescence one minute after the end of each irradiation, and we sampled the PMT anode current once every second, our measurements are not sensitive to features in the signals that change at time scales shorter than few seconds. We expect that at the time scales accessible through our data the characteristic times of de-excitation of the metastable states do not form a continuum, so we assume there exist a discrete set of mean lifetimes of the occupied metastable states that we call 'trap species'. The dimension of this set, N_t, is not known beforehand: it has to be determined by data analysis.

2. The electrons in the ionization cloud produced by the interaction of a proton recombine or get trapped in a time span that is much shorter than the minimum time scale our experiment is sensitive to. For this reason we simplify the description of the capture process considering it to be instantaneous, and by assigning a given, yet unknown, discrete probability distribution for the electrons to be captured in the trap species. In principle this probability distribution could change for the ionization clouds of different protons because of statistical fluctuations in ionization, uneven trap distributions, and so on.

3. We assume that the distributions of the trap species are uniform, or at least vary slowly, within the crystal volume.

4. We assume that the densities of the trap species are much higher than the largest density of electrons in the conduction band, so that the capture probabilities have at most a very week dependence on the number of occupied traps, which we initially neglect.

5. We ignore retrapping of the electrons emitted by the metastable state. Indeed, as it will be verified by the results of the afterglow data analysis, the probability for the electrons to be trapped in the relevant metastable states is so small that our measurements are not sensitive enough to be affected by this phenomenon.

6. We ignore the temporal structure of the beam because it has a negligible impact: the correction factor on the emission intensity is less than 1% for mean lifetimes larger than 2.5 s.

A further simplification stems from the large number of ionization electron-hole pairs liberated by a single proton (several millions), and from the large number of protons striking the crystal within the time resolution of our measurements: these large numbers allow us to mathematically threat the problem using continuous variables and differential equations, by employing averaged quantities, releasing ourselves from keeping track of the fluctuations involved in the real physical processes. This is also true when considering light excitation of the crystal.

In general, we can consider the capture process of an electron by a trap, or a recombination center, governed by an elementary probability π which may differ according to the trap or recombination center species. Let us call d their densities, and assign indices r to recombination centers, k to the trap species with mean lifetimes so short that they are not observable in the experiment, and i to the trap species we are interested in. It can be shown that if the densities are constant in the volume of the crystal, besides edge effects, the capture probability of the i-labelled trap species can be written

$$p_i = \frac{\pi_i d_i}{\sum_r \pi_r d_r + \sum_k \pi_k d_k + \sum_i \pi_i d_i} \tag{A.1}$$

In a good scintillator most of the electrons recombine radiatively, so $p_i \ll 1$. This is the reason why we neglect recapture of the charge carriers emitted by one of the traps into the traps we are interested to. In the case in which the densities of recombination centers and traps vary slowly within the volume of the crystal we replace Eq. A.1 with its average over the volume of the crystal affected by ionization.

As a way to further simplify the calculation, we relate our averaged quantities to a single incident particle since, from a practical point of view, they all produce the same average ionization. This choice allows for a simple rescaling of the results by the ratio of the average ionization energies when we apply our model to the prediction of the afterglow expected in the space radiation environment of the HERMES nano-satellites.

Since we initially consider the capture probabilities to be constant, we can treat the electron capture by the different trap species independently. Focusing our attention to the i-labelled species, then, the number N_i of occupied traps varies during a given constant stimulation (either irradiation or illumination) according to the following differential equation:

$$\frac{d}{dt}N_i(t) = n_i\,\phi - \frac{N_i(t)}{\tau_i}$$

The first term on the RHS accounts for the rate of electrons being trapped during irradiation, which is given by the product of the radiation flux integrated over the exposed surface of the crystal, ϕ, with the average number of electrons trapped for particle of incident radiation, n_i. This last quantity is directly linked with measurement data, and it may be used to calculate approximate average capture probabilities of the trap species by dividing its value by the average number of ionization electron-hole pairs produced by an incident particle. For simplicity we will often refer to n_i as the 'capture rate' of the traps.

The second term accounts for the electron emission from the traps, which happens with a rate that is proportional to the reciprocal of their mean lifetime in this particular metastable state.

The general solution to the equation above, referred to the j-th stimulation in a set, is

$$N_i(t) = N_{i,j}^0 e^{-\frac{t}{\tau_i}} + n_i\phi_j\tau_i\left(1 - e^{-\frac{t}{\tau_i}}\right) \tag{A.2}$$

where the time origin is at the beginning of the stimulation, and $N_{i,j}^0$ represents the number of traps already occupied at $t = 0$. Since the system is linear, we can treat separately the initial condition and the effects of all the m stimulation steps, thus the change in occupied traps at the end of the j-th stimulation, after time Δt_j, due only to the excitation it provided, is

$$\Delta N_{i,j} = n_i\tau_i\phi_j\left(1 - e^{-\frac{\Delta t_j}{\tau_i}}\right) \tag{A.3}$$

By introducing t_j, the time at which the j-th stimulation starts, and $t_{>j} = t - t_j - \Delta t_j$, the time referred to the end of the j-th stimulation, the number of occupied states of the trap species we are considering evolves with time, valid outside the stimulation steps, in the following way:

$$N_i(t) = N_i^0\, e^{-\frac{t}{\tau_i}} + \sum_{j=1}^{m} \theta(t_{>j})\, \Delta N_{i,j}\, e^{-\frac{t_{>j}}{\tau_i}} \tag{A.4}$$

where $\theta(x)$ is the Heaviside step function.

The afterglow model developed so far can be used to describe the emission of the crystal after a low intensity stimulation if radiation damage can be neglected. Under intense excitation conditions, some of the trap species may fill up rapidly, so their capture probabilities will decrease, and other traps may become more efficient in trapping the electrons. Indeed, in Eq. A.1 the parameters d_k and d_i should refer to the densities of the unoccupied traps.

The complete picture complicates if we attempt to account for the effects produced by radiation damage, which can create new defects or alter existing ones. In extending the model to consider these situations we are strongly constrained by the available data since the afterglow emission from the crystal was not characterized before the campaign, and we performed only six irradiation steps with large changes in the proton flux. These limitations force us to introduce only a simple modification to the assumptions we made: we allow the capture probability p_i to change with time during a stimulation. Assuming the change to be sufficiently small, we can expand the capture rate in a Taylor series of time, keeping only the linear term

$$n_i(t) \approx n_{i,0} + \frac{d}{dt} n_i(t)\, t$$

We consider first the case in which radiation damage is absent or negligible. The main cause of change in the capture rates is the reduction in available empty traps due to trapping of the free charge carriers. As a consequence, neglecting the small number of decaying metastable states, we may write the approximation

$$n_i(t) = n_i^0 + k_i\, \phi\, t$$

where k_i is a negative parameter. After the end of the stimulation the traps emit electrons so that at the beginning of the following stimulation the capture rate has partially recovered following the exponential decay of the metastable states.

Besides the mechanism just described, the change in the capture rate may be brought about by an increase in the electrons available for capture due to the reduction of trapping by another species, however this effect is quite small, since $p_i \ll 1$, and it can be observed only if the previous mechanism is negligible for the trap species under consideration.

In this approximation the time evolution of the number of occupied traps of the i-th species during the j-th stimulation is found by solving the system of equations

$$\begin{cases} \frac{d}{dt} N_i(t) = n_{i,j}(t)\phi_j - \frac{N_i(t)}{\tau_i} \\ n_{i,j}(t) = n^0_{i,j} + k_i\, \phi_j\, t \end{cases} \tag{A.5}$$

where $n^0_{i,j}$ is the value of the capture rate at the beginning of the j-th stimulation. The solution is

$$N_{i,j}(t) = N^0_{i,j}\, e^{-t/\tau_i} + \tau_i \phi_j n^0_{i,j} \left(1 - e^{-t/\tau_i}\right) + k_i \tau_i^2 \phi_j^2 \left[\frac{t}{\tau_i} - \left(1 - e^{-t/\tau_i}\right)\right]$$

where $N^0_{i,j}$ is the number of traps already occupied at the beginning of the stimulation. The first two terms on the RHS are equal to the result of the previous model, Eq. A.2, while the last term accounts for the variation in the capture rate of the trap species under consideration.

To write a complete formula for the evolution of the number of occupied traps we have to explicitly write the dependence of the trap rate on the previous stimulation steps in the second equation of system A.5. By introducing the fluence $\Phi_j = \phi_j \Delta t_j$ received by the crystal in the j-th stimulation (defining $\Phi_0 = 0$), the time interval $\Delta t_{j-1,j}$ between the end of the $(j-1)$-th stimulation and the beginning of the j-th one (with $\Delta t_{0,1} = t_1$), the sum of these intervals

$$\Delta t_{k,j} = \sum_{l=k}^{j-1} \Delta t_{l,l+1}\ ,$$

and by replacing the parameter k_i with $\Delta n_i / \Phi_{tot} = \Delta n_i / \sum_j \Phi_j$, to ease the interpretation of the results, we obtain

$$n^0_{i,j} = n^0_i + \frac{\Delta n_i}{\Phi_{tot}} \sum_{k=0}^{j-1} \Phi_k e^{-\frac{\Delta t_{k,j}}{\tau_i}}\ .$$

At the end of the j-th stimulation, then, the number of newly occupied traps is given by

$$\Delta N^{\phi}_{i,j} = \tau_i \phi_j \left\{ \left[n^0_i + \frac{\Delta n_i}{\Phi_{tot}} \left(\sum_{k=0}^{j-1} \Phi_k e^{-\frac{\Delta t_{k,j}}{\tau_i}} - \tau_i \phi_j \right) \right] \left(1 - e^{-\frac{\Delta t_j}{\tau_i}}\right) + \frac{\Delta n_i}{\Phi_{tot}} \Phi_j \right\} \tag{A.6}$$

and the time evolution of the occupied traps of the i-th species is still given by Eq. A.4 if we replace $\Delta N_{i,j}$ with $\Delta N^{\phi}_{i,j}$. We remark the fact that the effect becomes cumulative for trap species with mean lifetimes large with respect to the complete duration of the crystal stimulation.

We consider now the case in which the occupied traps can be neglected, due to a sufficiently low flux, but the radiation damage affects afterglow emission profile. The interaction of the radiation with the material making up the scintillator causes an increase in the number of defects present in the crystal, and possibly changes their densities in different ways. As a consequence, see Eq. A.1, the capture probabilities of the different trap species changes: they increase in the species that feature an augmented density, and they decrease slightly in the species whose densities are not affected by radiation damage. The phenomenon is cumulative, and at small irradiation levels the trap densities change linearly with the fluence received by the crystal. During irradiation, the capture rates are expected to change linearly with time because the radiation flux is constant, and the electron trapping is again described by the system of equations A.5, but now the time evolution of the capture rate is replaced by

$$n_{i,j}(t) = n_i^0 + \frac{\Delta n_i}{\Phi_{tot}} \left(\sum_{k=0}^{j-1} \Phi_k + \phi_j t \right)$$

where we used the same normalization of the rate of change in n_i. The number of newly occupied traps of the i-th species at the end of the j-th irradiation is now

$$\Delta N_{i,j}^{\Phi} = \tau_i \phi_j \left\{ \left[n_i^0 + \frac{\Delta n_i}{\Phi_{tot}} \left(\sum_{k=0}^{j-1} \Phi_k - \tau_i \phi_j \right) \right] \left(1 - e^{-\frac{\Delta t_j}{\tau_i}} \right) + \frac{\Delta n_i}{\Phi_{tot}} \Phi_j \right\} \quad \text{(A.7)}$$

By replacing $\Delta N_{i,j}$ with this quantities in Eq. A.4 we get the time evolution of the occupied traps of the i-th species during the irradiation campaign, valid outside the irradiation steps.

Having modeled the electron capture in the traps we are now able to calculate the afterglow emission of the crystal, which is due to recombination of the charge carriers liberated in the deexcitation of the metastable states. For each trap species we choose the appropriate model, by using Eq. A.3, A.6 or A.7 in Eq. A.4, driven by the information contained in the data. By neglecting retrapping, the photon flux emitted by the crystal is

$$\phi_q(t) = -\sum_{i=1}^{N_t} \frac{d}{dt} N_i(t) \, . \quad \text{(A.8)}$$

We observe two facts. In situations like the one we are analysing, in which large steps in fluence are obtained by corresponding increases in the flux, it may be difficult to distinguish between radiation damage and a reduced availability of empty traps, because most of the effect is due to a single irradiation and both scenarios are compatible by adequately scaling the parameter Δn_i. In this case help may come from the observation of a trap species with large increase in the capture rates that is not balanced by the reductions in the capture by other species, since this is only compatible with radiation damage (also the capture rates of both unobserved trap species and recombination centers have to decrease slightly).

The second observation concerns the trap species with smallest mean lifetime. The models we formulated take into account only the trap species with observable mean lifetime, while we know that there certainly are traps that decay faster than the detection limit of the experiment. The emission tails of these traps influence the shape of the first portion of the measured light curve, so the parameters we get for the species with smallest mean lifetime by fitting the model to the data are offset from the real values. This fact is unavoidable, because we don't have a description of what happens before the first data point, and it hinders the interpretation of the results for this trap species. The phenomenon propagates to the other trap species, with an effect that vanishes rapidly as the mean lifetime increases.

Appendix B

FOCuS minimal implementations

B.1 FOCuS normal

```python
'''
a minimal, functional implementation of FOCuS normal.
this code is based on the original OOP implementation
of Kester Ward. it was not intended for performance
sensible applications.
requires python 3.8.

parameters:
    X: observations (sequence or generator)
    threshold: in units st. dev. (numeric constant)
'''

def update(q, x_t):
    return (q[0]-1, q[1]+2*x_t)

def dominates(q, p):
    return not ((q[1] < p[1]) or (q[1]/q[0] > p[1]/p[0])):

def ymax(q):
    return - q[1]**2/(4*q[0])

def qtocp(q):
    return (q[0] and ymax(q) or 0, q[0])

def focus_rstep(qs, x_t, q):
    if qs and not dominates(q, p := update(qs[0], x_t)):
            return [p] + focus_rstep(qs[1:], x_t, p)
    return [(0,0.)]
```

```

def focus(X, threshold):
    qs = [(0,0.)]

    for t, x_t in enumerate(X):
        qs = focus_rstep(qs, x_t, (1,0.))
        global_max, time_offset = max(map(qtocp, qs))

        if global_max > threshold:
            return global_max, t+time_offset+1, t

    return 0., t+1, t
```

B.2 FOCuS Poisson

```
'''
a minimal, functional implementation of FOCuS Poisson.
this code is based on the original OOP implementation
of Kester Ward. it was not intended for performance
sensible applications.
requires python 3.8.

parameters:
    X: observations (a sequence or a generator)
    lambda_t: background countrate (a numeric constant)
    threshold: according to Wilks th. (numeric constant)
               i.e. T**2/2 with T in units of st. devs.
'''

from math import log

def update(c, x_t, lambda_t):
    return (c[0] + x_t, c[1] - lambda_t, c[2] - 1)

def dominates(c, k):
    return not ((c[0]/c[1] > k[0]/k[1]) or (c[0]+c[1] < k[0]+k[1]))

def ymax(c):
    return c[0]*log(-c[0]/c[1]) - (c[0] + c[1])

def ctocp(c):
    return (c[0] and ymax(c) or 0, c[0])

def focus_rstep(cs, x_t, lambda_t, c):
    if cs and not dominates(c, k := update(cs[0], x_t, lambda_t)):
            return [k] + focus_rstep(cs[1:], x_t, lambda_t, k)
    return [(0,0.,0)]

def focus(X, lambda_t, threshold):
    cs = [(0,0.,0)]
```

```

    for t,x_t in enumerate(X):
        cs = focus_rstep(cs, x_t, lambda_t, (1,-1.,0))
        global_max, time_offset = max(map(ctocp, cs))
        if global_max > threshold:
            return global_max, t+time_offset+1, t

    return 0., t+1, t
```

Appendix C

A visual comparison of different algorithmic logics for GRB detection

In Fig. C.1 a tabular representation of a synthetic count time series and its intervals is presented. We refer to this visualization device as a checker plot. A checker plot is made a up of tiles arranged in a table. A tile with index i, h is associated to an unique interval $X_{i,h}$ of the observation time series. Only tiles in the table upper-diagonals are present since tiles in the lower-diagonals would represent intervals starting before the observation series itself. Colors are used to represent interval significance scores computed according to Eq. 4.1. For those interval where a count excess was observed relatively to the number of counts expected from background (i.e. intervals such that $S_{i,h} > 0$) the significance score is also printed in digits. Tiles in a column share the same end count value and represent intervals whose duration increases from top to bottom. Hence the tiles in the i-th columns represents all the recent intervals at time index i. Black tiles are used to represent those intervals for which the significance score is maximum along a column, if it exceeds 0. On the other hand, intervals in the i-th upper diagonal share the first count value x_i and grow in length moving from top-left to bottom-right.

Checker plots are useful tools for visualizing the operations of TAs and comparing computational performances between different strategies. In fact, different TAs will test the significant of different intervals-tiles. By plotting only the tiles whose significance is actually evaluated we can visualize the behaviour of a given TA strategy. Moreover, since most of the computational resources available to a TA will be spent in significance computations, we can achieve a first-order

comparison of the computational performances comparing the number of tiles evaluated by different TAs.
Consider for example the plot Fig. C.3 which represents the operations of a conventional TA, such that of Algo. 1, pg. 76. The only intervals tested have duration in a predefined, logarithmically equispaced set of bin-lengths. Tests happen at regular interval of time, proportional to the bin-length themselves. Since no interval with significance exceeding 5σ was tested, the algorithm operations do not result in a trigger and reach the end of the observation series. The plot of Fig. C.2 represents the operation of an exhaustive search algorithm such that sketched in Algo. 2 at pg. 77. The significance value of all intervals is computed and compared against a 5σ threshold up to the first trigger, after which the algorithm operation stops. Algorithm with ideal sensibilities, such as exhaustive search, will always check the significance score of those intervals represented with a black tile. Black tiles are used to represent those intervals for which the significance score is maximum along a column and positive. The distribution of black tiles in a checker plot is far from uniform; most of the black tiles being arranged in such a way to seemingly follow a number of diagonals. At each iteration, FOCuS dynamically updates the duration of the intervals 'scheduled' for a test, conditionally on the past observation. On a checker plot this behaviour results in an operation patterns tracking the maximum significance diagonals, see Fig. C.4.

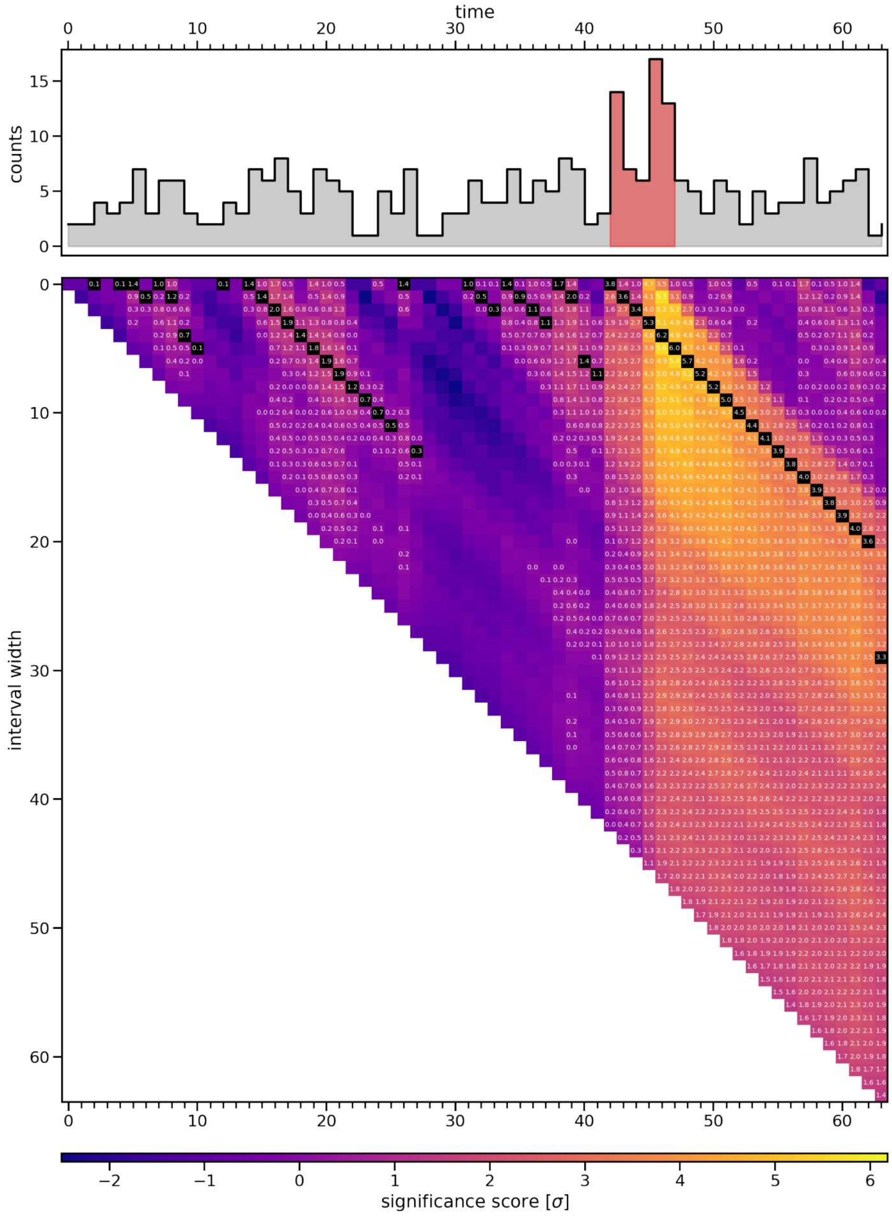


Figure C.1: Checker plot representation of the intervals of an observation series of 64 data points (bottom panel) and of the observation series itself (top panel). Data-points with indeces comprised between 40 and 45 were sampled according to a Poisson distribution with mean rate 9.0 while all the rest were sampled from a Poisson distribution with mean rate 4.5. Colors are used to represent interval significance scores computed according to Eq. 4.1 and background rate 4.5. For those interval such that $S_{i,h} > 0$) the significance score is also printed in digits. Black tiles are used to represent those intervals for which the significance score is both maximum across a column and positive. See Sec. C.

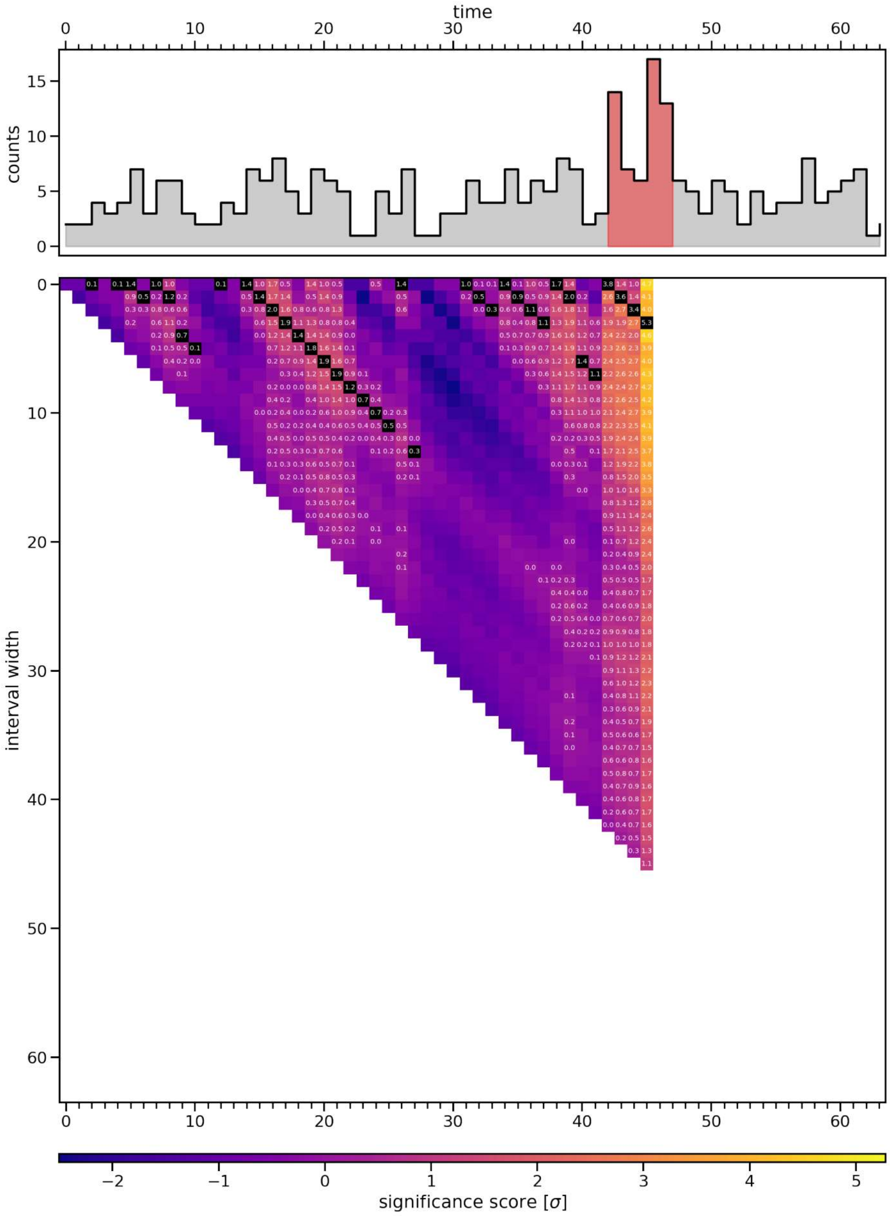


Figure C.2: Checker plot representation of the operations of an exhaustive search TA. Interval significances computed according to Eq. 4.1 and true background rate 4.5.

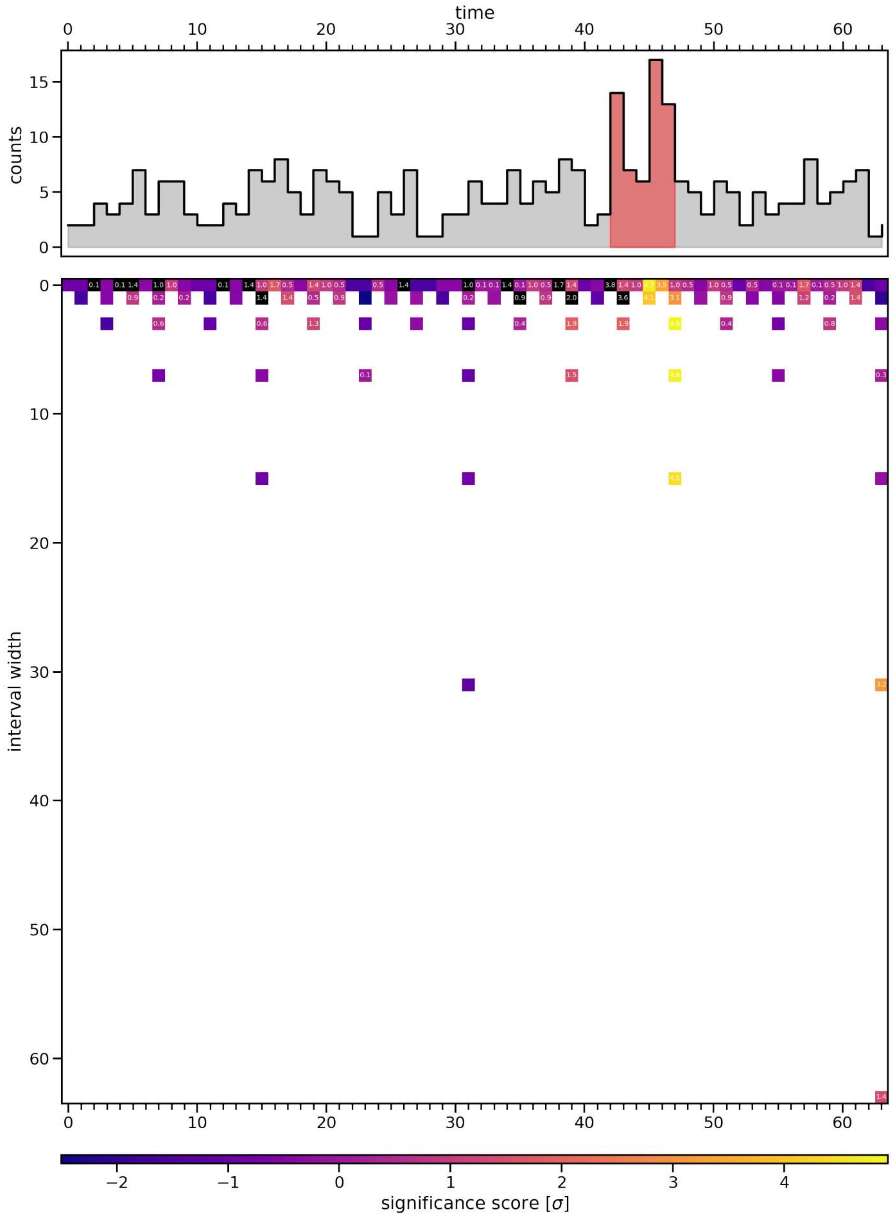


Figure C.3: Checker plot representation of the operations of a conventional TA. Interval significances computed according to Eq. 4.1 and true background rate 4.5.

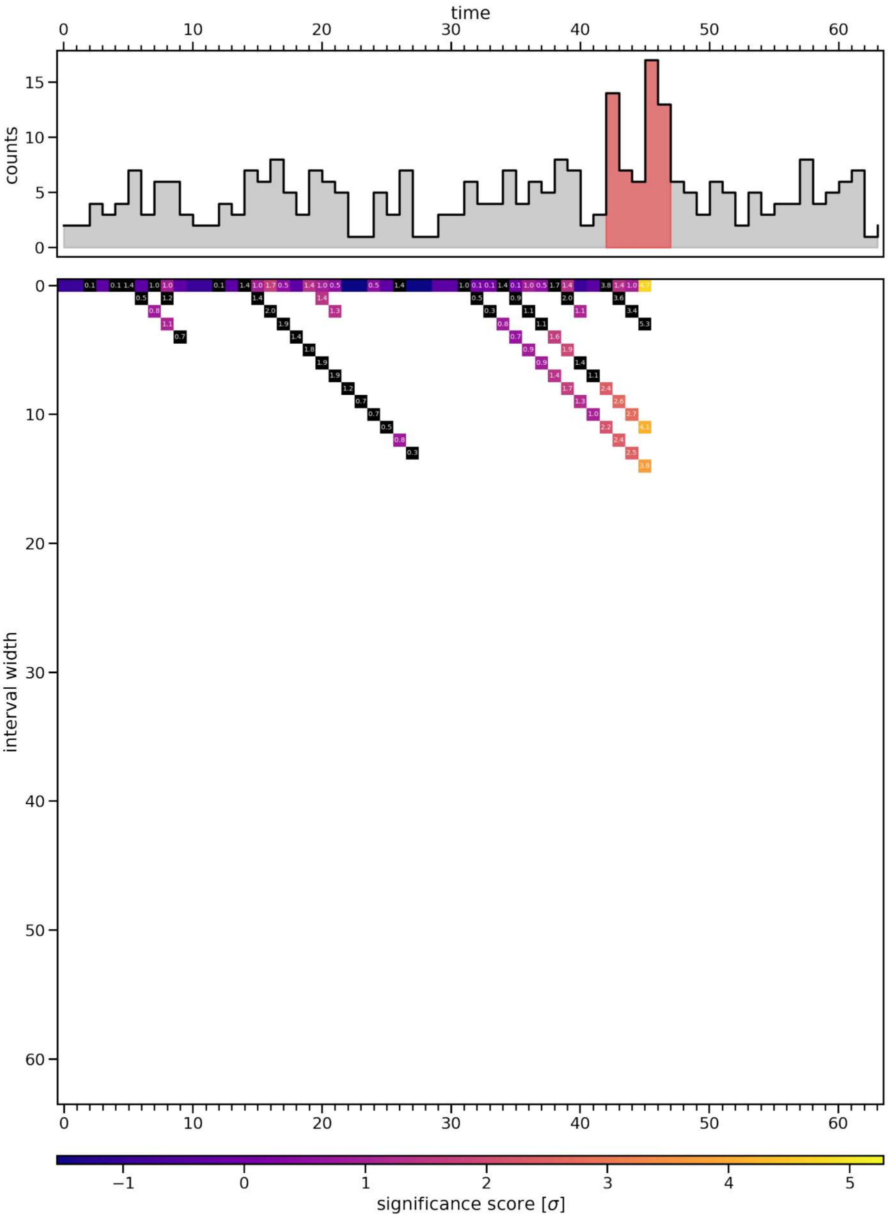


Figure C.4: Checker plot representation of the operations of FOCuS Poisson with known background rate 4.5.

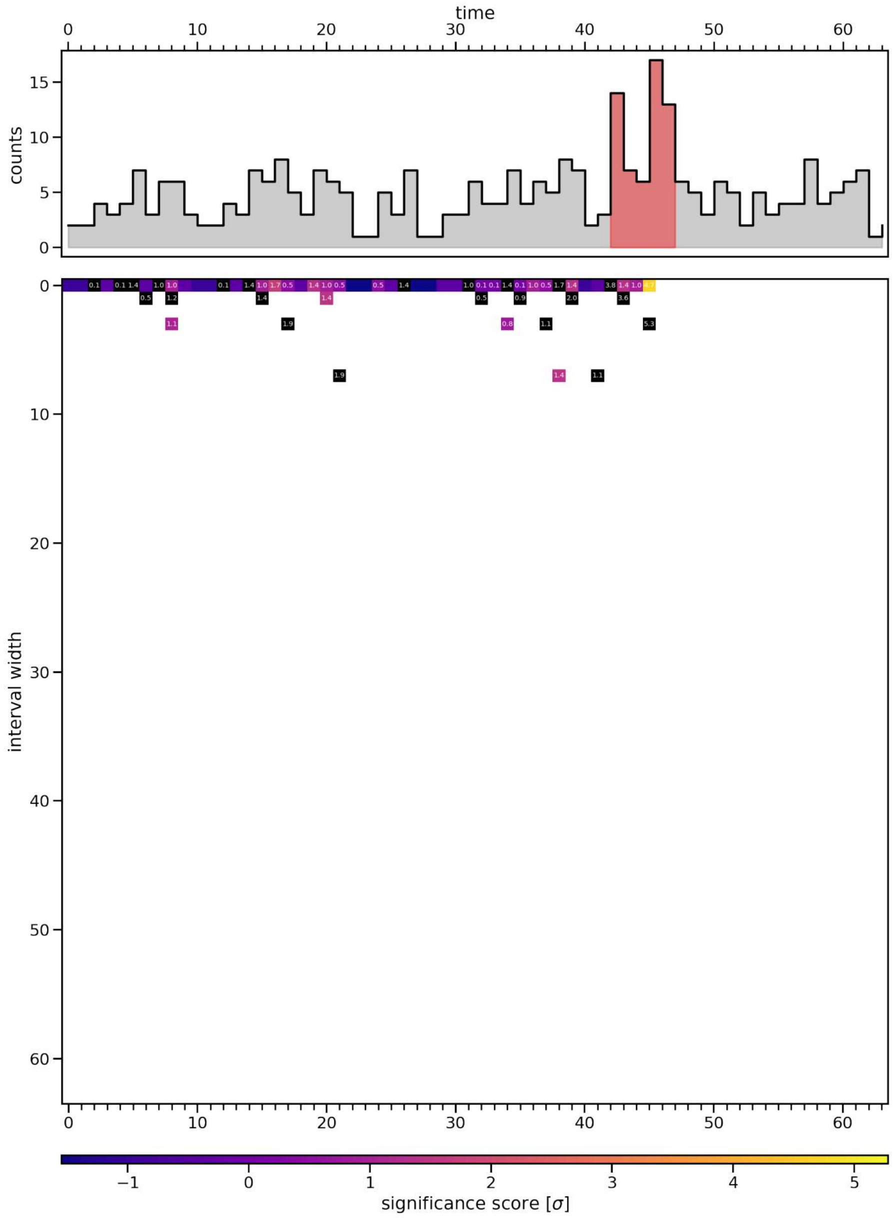


Figure C.5: Checker plot representation of the operations of FOCuS Poisson with known background rate 4.5, checking only curve over a logarithmically spaced grid.

Appendix D

Bibliography

[1] Batse 4b gamma-ray burst catalog. `https://gammaray.nsstc.nasa.gov/batse/grb/duration/`. [Online; accessed June 2021].

[2] C&a gagg scintillation properties. `http://www.c-and-a.jp/assets/img/products/101180228_GAGG.pdf`. Accessed: 2020-01-29.

[3] The calibration requirements for spectral analysis (definition of rmf and arf file formats). `https://web.archive.org/web/20210318081945/https://heasarc.gsfc.nasa.gov/docs/heasarc/caldb/docs/memos/cal_gen_92_002/cal_gen_92_002.html`. Accessed: 2022-01-14.

[4] collections — container datatypes). `https://docs.python.org/3/library/collections.html`. Accessed: 2022-01-14.

[5] Ecss-e-st-10-04c – space environment. `https://ecss.nl/standard/ecss-e-st-10-04c-space-environment/`. Accessed: 2020-09-25.

[6] What is exponential smoothing? `https://web.archive.org/web/20211113234158/https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc43.htm`. Accessed: 2022-01-14.

[7] Why python is slow: Looking under the hood). `https://web.archive.org/web/20210331133101/http://jakevdp.github.io/blog/2014/05/09/why-python-is-slow/`. Accessed: 2022-01-14.

[8] B. P. Abbott et al. Multi-messenger Observations of a Binary Neutron Star Merger. *Astrophys. J. Lett.*, 848(2):L12, 2017.

[9] Benjamin P Abbott, R Abbott, TD Abbott, S Abraham, F Acernese, K Ackley, C Adams, VB Adya, C Affeldt, M Agathos, et al. Prospects for

observing and localizing gravitational-wave transients with advanced ligo, advanced virgo and kagra. *Living reviews in relativity*, 23(1):1–69, 2020.

[10] Bannanje Sripathi Acharya, Iván Agudo, Imen Al Samarai, R Alfaro, J Alfaro, C Alispach, R Alves Batista, JP Amans, E Amato, G Ambrosi, et al. Science with the cherenkov telescope array. 2017.

[11] Markus Ackermann, Marco Ajello, K Asano, Magnus Axelsson, Luca Baldini, Jean Ballet, Guido Barbiellini, D Bastieri, K Bechtol, R Bellazzini, et al. The first fermi-lat gamma-ray burst catalog. *The Astrophysical Journal Supplement Series*, 209(1):11, 2013.

[12] European Space Agency. Esa athena mission summary. `https://sci.esa.int/web/athena/-/59896-mission-summary`, 2021. [Online; accessed April 2022].

[13] M Aguilar, J Alcaraz, J Allaby, B Alpat, G Ambrosi, H Anderhub, L Ao, A Arefiev, P Azzarello, E Babucci, et al. The alpha magnetic spectrometer (ams) on the international space station: Part i–results from the test flight on the space shuttle. *Physics Reports*, 366(6):331–405, 2002.

[14] Mahdi Ahangarianabhari, Daniele Macera, Giuseppe Bertuccio, Piero Malcovati, and Marco Grassi. Vega: A low-power front-end asic for large area multi-linear x-ray silicon drift detectors: Design and experimental characterization. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 770:155–163, 2015.

[15] Marco Ajello, J Greiner, G Sato, DR Willis, G Kanbach, AW Strong, R Diehl, G Hasinger, N Gehrels, CB Markwardt, et al. Cosmic x-ray background and earth albedo spectra with swift bat. *The Astrophysical Journal*, 689(2):666, 2008.

[16] V. Alenkov et al. Irradiation studies of a multi-doped $Gd_3Al_2Ga_3O_{12}$ scintillator. *Nucl. Instrum. Meth. A*, 916:226–229, 2019.

[17] Lorenzo Amati, Paul T O'Brien, Diego Götz, Enrico Bozzo, and Andrea Santangelo. The theseus space mission: updated design, profile and expected performances. In *Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray*, volume 11444, page 114442J. International Society for Optics and Photonics, 2021.

[18] Sami W Asmar and Steven Matousek. Mars cube one (marco) shifting the paradigm in relay deep space operation. In *14th International Conference on Space Operations*, page 2483, 2016.

[19] Yoichi Aso, Yuta Michimura, Kentaro Somiya, Masaki Ando, Osamu Miyakawa, Takanori Sekiguchi, Daisuke Tatsumi, Hiroaki Yamamoto, Kagra Collaboration, et al. Interferometer design of the kagra gravitational wave detector. *Physical Review D*, 88(4):043007, 2013.

[20] Astrophysics Science Division at NASA/ GSFC. Batse all-sky plot of gamma-ray burst locations. `https://heasarc.gsfc.nasa.gov/docs/cgro/cgro/batse_src.html`, 2012. [Online; accessed June 2021].

[21] D. Band et al. BATSE observations of gamma-ray burst spectra. 1. Spectral diversity. *Astrophys. J.*, 413:281–292, 1993.

[22] David L. Band. Gamma-ray burst spectral evolution through crosscorrelations of discriminator light curves. *Astrophys. J.*, 486:928, 1997.

[23] X Barcons, D Barret, A Decourchelle, J-W den Herder, T Dotani, AC Fabian, R Fraga-Encinas, H Kunieda, D Lumb, G Matt, et al. Athena (advanced telescope for high energy astrophysics) assessment study report for esa cosmic vision 2015-2025. *arXiv preprint arXiv:1207.2745*, 2012.

[24] Didier Barret, Thien Lam Trong, Jan-Willem Den Herder, Luigi Piro, Xavier Barcons, Juhani Huovelin, Richard Kelley, J Miguel Mas-Hesse, Kazuhisa Mitsuda, Stéphane Paltani, et al. The athena x-ray integral field unit (x-ifu). In *Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray*, volume 9905, pages 714–754. SPIE, 2016.

[25] E. Berger. Short-Duration Gamma-Ray Bursts. *Annual Review of Astronomy and Astrophysics*, 52:43–105, August 2014.

[26] E Berger, SR Kulkarni, DA Frail, and AM Soderberg. A radio survey of type ib and ic supernovae: Searching for engine-driven supernovae. *The Astrophysical Journal*, 599(1):408, 2003.

[27] P Narayana Bhat, Charles A Meegan, Andreas von Kienlin, William S Paciesas, Michael S Briggs, J Michael Burgess, Eric Burns, Vandiver Chaplin, William H Cleveland, Andrew C Collazzi, et al. The third fermi gbm gamma-ray burst catalog: the first six years. *The Astrophysical Journal Supplement Series*, 223(2):28, 2016.

[28] Ž. Bošnjak, D. Götz, L. Bouchet, S. Schanne, and B. Cordier. The spectral catalogue of INTEGRAL gamma-ray bursts: results of the joint IBIS/SPI spectral analysis. *Astron. Astrophys.*, 561:A25, 2014.

[29] M. S. Briggs et al. Observations of grb 990123 by the Compton gamma-ray observatory. *Astrophys. J.*, 524:82, 1999.

[30] Michael S. Briggs, William S. Paciesas, Geoffrey N. Pendleton, Charles A. Meegan, Gerald J. Fishman, John M. Horack, Martin Brock, Chryssa Kouveliotou, Dieter H. Hartmann, and Jon Hakkila. Batse observations of the large scale isotropy of gamma-ray bursts. *Astrophys. J.*, 459:40, 1996.

[31] R. Campana et al. A compact and modular X and gamma-ray detector with a CsI scintillator and double-readout Silicon Drift Detectors. *Proc. SPIE Int. Soc. Opt. Eng.*, 9905:99056I, 2016.

[32] Riccardo Campana, Marco Feroci, Ettore Del Monte, Teresa Mineo, Niels Lund, and George W Fraser. Background simulations for the large area detector onboard loft. *Experimental Astronomy*, 36(3):451–477, 2013.

[33] Riccardo Campana, Fabio Fuschino, Yuri Evangelista, Giuseppe Dilillo, and Fabrizio Fiore. The hermes-tp/sp background and response simulations. In *Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray*, volume 11444, page 114444U. International Society for Optics and Photonics, 2020.

[34] Riccardo Campana, Mauro Orlandini, Ettore Del Monte, Marco Feroci, and Filippo Frontera. The radiation environment in a low earth orbit:the case of BeppoSAX. *Experimental Astronomy*, 37(3):599–613, Nov 2014.

[35] S. Campana et al. The association of GRB 060218 with a supernova and the evolution of the shock wave. *Nature*, 442:1008–1010, August 2006.

[36] S. Campana et al. The shock break-out of grb 060218/sn 2006aj. *Nature*, 442:1008–1010, 2006.

[37] Andrea Colagrossi, Jacopo Prinetto, Stefano Silvestrini, and Michèle R Lavagna. Sky visibility analysis for astrophysical data return maximization in hermes constellation. *Journal of Astronomical Telescopes, Instruments, and Systems*, 6(4):048001, 2020.

[38] E. Costa et al. Discovery of an X-ray afterglow associated with the gamma-ray burst of 28 February 1997. *Nature*, 387:783–785, 1997.

[39] S. Dichiara et al. Average power density spectrum of long GRBs detected with BeppoSAX/GRBM and with Fermi/GBM. *Monthly Notices of the Royal Astronomical Society*, 431:3608, 2013.

[40] Giuseppe Dilillo, Nicola Zampa, Riccardo Campana, Fabio Fuschino, Giovanni Pauletta, Irina Rashevskaya, Filippo Ambrosino, Marco Baruzzo, Diego Cauz, Daniela Cirrincione, et al. Space applications of gagg: Ce

scintillators: a study of afterglow emission by proton irradiation. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 513:33–43, 2022.

[41] Yuri Evangelista, Fabrizio Fiore, Fabio Fuschino, Riccardo Campana, Francesco Ceraudo, Evgeny Demenev, Alejandro Guzman, Claudio Labanti, Giovanni La Rosa, Mauro Fiorini, et al. The scientific payload onboard the hermes-tp and hermes-sp cubesat missions. In *Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray*, volume 11444, page 114441T. International Society for Optics and Photonics, 2020.

[42] Ed Fenimore, David Palmer, Mark Galassi, Tanya Tavenner, Scott Barthelmy, Neil Gehrels, Ann Parsons, and Jack Tueller. The Trigger algorithm for the Burst Alert Telescope on Swift. *AIP Conf. Proc.*, 662(1):491, 2003.

[43] EE Fenimore, David Palmer, Mark Galassi, Tanya Tavenner, Scott Barthelmy, Neil Gehrels, Ann Parsons, and Jack Tueller. The trigger algorithm for the burst alert telescope on swift. In *AIP Conference Proceedings*, volume 662, pages 491–493. American Institute of Physics, 2003.

[44] Marco Feroci, Filippo Frontera, Enrico Costa, Daniele Dal Fiume, Lorenzo Amati, L Bruca, Maria Nerina Cinti, Alessandro Coletta, P Collina, C Guidorzi, et al. In-flight performances of the bepposax gamma-ray burst monitor. In *EUV, X-Ray, and Gamma-Ray Instrumentation for Astronomy VIII*, volume 3114, pages 186–197. International Society for Optics and Photonics, 1997.

[45] Fabrizio Fiore, Luciano Burderi, Michelle Lavagna, Roberto Bertacin, Yuri Evangelista, Riccardo Campana, Fabio Fuschino, Paolo Lunghi, Angel Monge, Barbara Negri, et al. The hermes-technologic and scientific pathfinder. In *Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray*, volume 11444, page 114441R. International Society for Optics and Photonics, 2020.

[46] W. Fong, E. Berger, and D. B. Fox. Hubble Space Telescope Observations of Short Gamma-Ray Burst Host Galaxies: Morphologies, Offsets, and Local Environments. *The Astrophysical Journal*, 708:9–25, January 2010.

[47] A.S. Fruchter et al. Long γ-ray bursts and core-collapse supernovae have different environments. *Nature*, 441:463–468, May 2006.

[48] Fabio Fuschino, RICCARDO Campana, CLAUDIO Labanti, Yuri Evangelista, MARCO Feroci, L Burderi, Fabrizio Fiore, Filippo Ambrosino,

G Baldazzi, P Bellutti, et al. Hermes: An ultra-wide band x and gamma-ray transient monitor on board a nano-satellite constellation. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 936:199–203, 2019.

[49] Emilio Gatti and Pavel Rehak. Semiconductor drift chamber—an application of a novel charge transport scheme. *Nuclear Instruments and Methods in Physics Research*, 225(3):608–614, 1984.

[50] Neil Gehrels, Guido Chincarini, P ea Giommi, KO Mason, JA Nousek, AA Wells, NE White, SD Barthelmy, DN Burrows, LR Cominsky, et al. The swift gamma-ray burst mission. *The Astrophysical Journal*, 611(2):1005, 2004.

[51] B. Gendre, Q. T. Joyce, N. B. Orange, G. Stratta, J. L. Atteia, and M. Boër. Can we quickly flag Ultra-long Gamma-Ray Bursts? *Mon. Not. Roy. Astron. Soc.*, 486(2):2471–2476, 2019.

[52] Gregory Ginet, T. O'Brien, S. Huston, W. Johnston, Timothy Guild, R. Friedel, C. Lindstrom, Christopher Roth, P. Whelan, R. Quinn, D. Madden, Steven Morley, and Yi-Jiun Su. Ae9, ap9 and spm: New models for specifying the trapped energetic particle and space plasma environment. *Space Science Reviews*, 179, 11 2013.

[53] V. Zach Golkhou, Nathaniel R. Butler, and Owen M. Littlejohns. The Energy-Dependence of GRB Minimum Variability Timescales. *Astrophys. J.*, 811(2):93, 2015.

[54] DE Gruber, JL Matteson, LE Peterson, and GV Jung. The spectrum of diffuse cosmic hard x-rays measured with heao 1. *The Astrophysical Journal*, 520(1):124, 1999.

[55] A Guzman, S Pliego, J Bayer, Y Evangelista, G La Rosa, G Sottile, S Curzel, R Campana, F Fiore, F Fuschino, et al. The payload data handling unit (pdhu) on-board the hermes-tp and hermes-sp cubesat missions. In *Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray*, volume 11444, page 1144450. International Society for Optics and Photonics, 2020.

[56] A Hajela, R Margutti, J Bright, KD Alexander, W Fong, E Berger, R Chornock, DL Coppejans, P Blanchard, VA Villar, et al. Continued chandra observations of gw170817 at 3.3 years since merger. *GRB Coordinates Network*, 29375:1, 2021.

[57] NASA HEASARC. Bepposax data archive. `https://heasarc.gsfc.nasa.gov/docs/sax/archive/dataarchive_files.html`, 2021. [Online; accessed April 2022].

[58] NASA HEASARC. Fermigdays - fermi gbm daily data. `https://heasarc.gsfc.nasa.gov/W3Browse/fermi/fermigdays.html`, 2021. [Online; accessed April 2022].

[59] Jens Hjorth et al. A Very energetic supernova associated with the gamma-ray burst of 29 March 2003. *Nature*, 423:847–850, 2003.

[60] Bernhard Hofmann-Wellenhof, Herbert Lichtenegger, and Elmar Wasle. *GNSS–global navigation satellite systems: GPS, GLONASS, Galileo, and more.* Springer Science & Business Media, 2007.

[61] D.J. Huntley. An explanation of the power-law decay of luminescence. *Journal of Physics Condensed Matter*, 18(4):1359–1365, Feb 2006.

[62] Kunihito Ioka, Kenta Hotokezaka, and Tsvi Piran. Are Ultra-Long Gamma-Ray Bursts Caused by Blue Supergiant Collapsars, Newborn Magnetars, or White Dwarf Tidal Disruption Events? *Astrophys. J.*, 833(1):110, 2016.

[63] Kei Kamada, Takayuki Yanagida, Takanori Endo, Kousuke Tsutumi, Yoshiyuki Usuki, Martin Nikl, Yutaka Fujimoto, Akihiro Fukabori, and Akira Yoshikawa. 2inch diameter single crystal growth and scintillation properties of ce:gd3al2ga3o12. *Journal of Crystal Growth*, 352(1):88 – 90, 2012. The Proceedings of the 18th American Conference on Crystal Growth and Epitaxy.

[64] Hamamatsu Photonics K.K. *Photomultiplier Tubes, Basics and Applications.* Hamamatsu Photonics K.K., 3 edition, 2007.

[65] Ray W. Klebesadel, Ian B. Strong, and Roy A. Olson. Observations of Gamma-Ray Bursts of Cosmic Origin. *Astrophys. J. Lett.*, 182:L85–L88, 1973.

[66] Glenn F. Knoll. *Radiation Detection and Measurement.* John Wiley and Sons, 4 edition, 2010.

[67] Shiho Kobayashi, Tsvi Piran, and Re'em Sari. Can internal shocks produce the variability in GRBs? *The Astrophysical Journal*, 490:92–98, 1997.

[68] Jefferson Michael Kommers. *Faint gamma-ray bursts and other high-energy transients detected with BATSE.* PhD thesis, Massachusetts Institute of Technology, 1999.

[69] Chryssa Kouveliotou, Charles A. Meegan, Gerald J. Fishman, Narayana P. Bhyat, Michael S. Briggs, Thomas M. Koshut, William S. Paciesas, and Geoffrey N. Pendleton. Identification of two classes of gamma-ray bursts. *Astrophys. J. Lett.*, 413:L101–104, 1993.

[70] Pawan Kumar and Bing Zhang. The physics of gamma-ray bursts \& relativistic jets. *Phys. Rept.*, 561:1–109, 2014.

[71] Davide Lazzati and Gabriele Ghisellini. Constraints on the bulk Lorentz factor in the internal shock scenario for gamma-ray bursts. *Monthly Notices of the Royal Astronomical Society*, 309:13, 1999.

[72] A. J. Levan et al. A new population of ultra-long duration gamma-ray bursts. *Astrophys. J.*, 781:13, 2013.

[73] Hui Li and Edward E. Fenimore. Log-normal distributions in gamma-ray burst time histories. *Astrophys. J. Lett.*, 469:L115, 1996.

[74] T. P. Li and Y. Q. Ma. Analysis methods for results in gamma-ray astronomy. *Astrophys. J.*, 272:317–324, September 1983.

[75] James M. Lucas. Counted data cusum's. *Technometrics*, 27(2):129–144, 1985.

[76] M.T. Lucchini, V. Babin, P. Bohacek, S. Gundacker, K. Kamada, M. Nikl, A. Petrosyan, A. Yoshikawa, and E. Auffray. Effect of mg2+ ions co-doping on timing performance and radiation tolerance of cerium doped gd3al2ga3o12 crystals. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 816:176 – 183, 2016.

[77] A. MacFadyen and S. E. Woosley. Collapsars: Gamma-ray bursts and explosions in 'failed supernovae'. *The Astrophysical Journal*, 524:262, 1999.

[78] M Marisaldi, C Labanti, H Soltau, C Fiorini, A Longoni, and F Perotti. X-and gamma-ray detection with a silicon drift detector coupled to a csi (tl) scintillator operated with pulse shape discrimination technique. *IEEE Transactions on Nuclear Science*, 52(5):1842–1848, 2005.

[79] Kassandra McLean, Ed E. Fenimore, David Palmer, S. Barthelmy, N. Gehrels, H. Krimm, C. Markwardt, and A. Parsons. Setting the triggering threshold on Swift. *AIP Conf. Proc.*, 727(1):667–670, 2004.

[80] C. A. Meegan, G. J. Fishman, R. B. Wilson, W. S. Paciesas, G. N. Pendleton, J. M. Horack, M. N. Brock, and C. Kouveiotou. Spatial distribution of gamma-ray bursts observed by BATSE. *Nature*, 355:143–145, 1992.

[81] Charles Meegan et al. The Fermi Gamma-Ray Burst Monitor. *Astrophys. J.*, 702:791–804, 2009.

[82] M. R. Metzger, S. G. Djorgovski, S. R. Kulkarni, C. C. Steidel, K. L. Adelberger, D. A. Frail, E. Costa, and F. Frontera. Spectral constraints ont he redshift of the optical counterpart to the gamma-ray burst of 8 May 1997. *Nature*, 387:878–880, 1997.

[83] Tsunefumi Mizuno, T Kamae, G Godfrey, T Handa, DJ Thompson, D Lauben, Y Fukazawa, and M Ozaki. Cosmic-ray background flux model based on a gamma-ray large area space telescope balloon flight engineering model. *The Astrophysical Journal*, 614(2):1113, 2004.

[84] K. P. Mooley, A. T. Deller, O. Gottlieb, E. Nakar, G. Hallinan, S. Bourke, D. A. Frail, A. Horesh, A. Corsi, and K. Hotokezaka. Superluminal motion of a relativistic jet in the neutron-star merger GW170817. *Nature*, 561(7723):355–359, 2018.

[85] Igor V Moskalenko and AW Strong. Production and propagation of cosmic-ray positrons and electrons. *The Astrophysical Journal*, 493(2):694, 1998.

[86] E. Nakar and Tsvi Piran. GRBs light curves - Another clue on the inner engine. *Astrophys. J. Lett.*, 572:L139–L142, 2002.

[87] Ramesh Narayan, Tsvi Piran, and Pawan Kumar. Accretion models of gamma-ray bursts. *The Astrophysical Journal*, 557:949, 2001.

[88] J. P. Norris, G. H. Share, D. C. Messina, B. R. Dennis, U. D. Desai, T. L. Cline, S. M. Matz, and E. L. Chupp. Spectral Evolution of Pulse Structures in Gamma-Ray Bursts. *The Astrophysical Journal*, 301:213, February 1986.

[89] G. Oganesyan, L. Nava, et al. Detection of Low-energy Breaks in Gamma-Ray Burst Prompt Emission Spectra. *The Astrophysical Journal*, 846:137, 2017.

[90] Gor Oganesyan, Lara Nava, et al. Characterization of Gamma-Ray Burst prompt emission spectra down to soft X-rays. *Astronomy & Astrophysics*, 616:A138, 2018.

[91] William S. Paciesas et al. The Fermi GBM Gamma-Ray Burst Catalog: The First Two Years. *Astrophys. J. Suppl.*, 199:18, 2012.

[92] William S Paciesas, Charles A Meegan, Geoffrey N Pendleton, Michael S Briggs, Chryssa Kouveliotou, Thomas M Koshut, John Patrick Lestrade, Michael L McCollough, Jerome J Brainerd, Jon Hakkila, et al. The fourth batse gamma-ray burst catalog (revised). *The Astrophysical Journal Supplement Series*, 122(2):465, 1999.

[93] Bohdan Paczynski. Gamma-ray bursters at cosmological distances. *Astrophys. J. Lett.*, 308:L43–L46, 1986.

[94] Ewan S Page. Continuous inspection schemes. *Biometrika*, 41(1/2):100–115, 1954.

[95] Anthony L Piro and Juna A Kollmeier. Evidence for cocoon emission from the early light curve of sss17a. *The Astrophysical Journal*, 855(2):103, 2018.

[96] R. D. Preece, M. S. Briggs, R. S. Mallozzi, G. N. Pendleton, W. S. Paciesas, and D. L. Band. The BATSE Gamma-Ray Burst Spectral Catalog. 1. High time resolution spectroscopy of bright bursts using high energy resolution data. *Astrophys. J. Suppl.*, 126:19, 2000.

[97] Yi-Ping Qin, Guang-Zhong Xie, Sui-Jian Xue, Xue-Tang Zheng, and Dong-Cheng Mei. The hardness-duration correlation in the two classes of gamma-ray bursts. *Publ. Astron. Soc. Jap.*, 52:759, 2000.

[98] G Romano, I Eckley, P Fearnhead, and G Rigaill. Fast online changepoint detection with a constant per-iteration computational cost. *arXiv preprint*, 2021.

[99] M Sakano, T Nakamori, Shuichi Gunji, J Katagiri, S Kimura, S Otake, and H Kitamura. Estimating the radiative activation characteristics of a gd3al2ga3o12:ce scintillator in low earth orbit. *Journal of Instrumentation*, 9:P10003, 10 2014.

[100] Andrea Sanna, Luciano Burderi, Tiziana Di Salvo, Fabrizio Fiore, Alessandro Riggio, Angelo Gambino, Michèle Lavagna, Roberto Bertacin, Yuri Evangelista, Riccardo Campana, et al. Timing techniques applied to distributed modular high-energy astronomy: the hermes project. In *Space Telescopes and Instrumentation 2020: Ultraviolet to Gamma Ray*, volume 11444, page 114444X. International Society for Optics and Photonics, 2020.

[101] T Sanuki, H Matsumoto, M Nozaki, K Abe, K Anraku, Y Asaoka, M Fujikawa, M Imori, T Maeno, Y Makida, et al. Precise measurements of cosmic-ray hydrogen and helium spectra with bess. *Advances in Space Research*, 27(4):761–766, 2001.

[102] Jeffrey D. Scargle. Studies in astronomical time series analysis: 5. Bayesian blocks, a new method to analyze structure in photon counting data. *Astrophys. J.*, 504:405, 1998.

[103] Jeffrey D. Scargle, Jay P. Norris, Brad Jackson, and James Chiang. Studies in Astronomical Time Series Analysis. VI. Bayesian Block Representations. *Astrophys. J.*, 764:167, 2013.

[104] M Tavani, G Barbiellini, A Argan, F Boffelli, A Bulgarelli, P Caraveo, PW Cattaneo, AW Chen, V Cocco, E Costa, et al. The agile mission. *Astronomy & Astrophysics*, 502(3):995–1013, 2009.

[105] Planet Team. The planet constellation. `https://www.planet.com/our-constellations/`. [Online; accessed April 2022].

[106] F. Tommasino, M. Rovituso, S. Fabiano, S. Piffer, C. Manea, S. Lorentini, S. Lanzone, Z. Wang, M. Pasini, W. J. Burger, C. La Tessa, E. Scifoni, M. Schwarz, and M. Durante. Proton beam characterization in the experimental room of the Trento Proton Therapy facility. *Nuclear Instruments and Methods in Physics Research A*, 869:15–20, Oct 2017.

[107] E. Troja and others. The outflow structure of GW170817 from late-time broad-band observations. *Monthly Notices of the Royal Astronomical Society*, 478(1):L18–L23, 2018.

[108] CS Unnikrishnan. Indigo and ligo-india: Scope and plans for gravitational wave research and precision metrology in india. *International Journal of Modern Physics D*, 22(01):1341010, 2013.

[109] V. V. Usov. Millisecond pulsars with extremely strong magnetic fields as a cosmological source of gamma-ray bursts. *Nature*, 357:472–474, 1992.

[110] S. Valenti et al. The discovery of the electromagnetic counterpart of GW170817: kilonova AT 2017gfo/DLT17ck. *The Astrophysical Journal Letters,*, 848(2):L24, 2017.

[111] Giacomo Vianello. The Significance of an Excess in a Counting Experiment: Assessing the Impact of Systematic Uncertainties and the Case with a Gaussian Background. *Astrophys. J. Suppl.*, 236(1):17, May 2018.

[112] A von Kienlin, CA Meegan, WS Paciesas, PN Bhat, E Bissaldi, MS Briggs, E Burns, WH Cleveland, MH Gibby, MM Giles, et al. The fourth Fermi-GBM gamma-ray burst catalog: a decade of data. *The Astrophysical Journal*, 893(1):46, 2020.

[113] Andreas Von Kienlin, Charles A Meegan, William S Paciesas, PN Bhat, Elisabetta Bissaldi, Michael S Briggs, J Michael Burgess, David Byrne, Vandiver Chaplin, William Cleveland, et al. The second fermi gbm gamma-ray burst catalog: the first four years. *The Astrophysical Journal Supplement Series*, 211(1):13, 2014.

[114] Jakub Řípa, Giuseppe Dilillo, Riccardo Campana, and Gábor Galgóczi. A comparison of trapped particle models in low Earth orbit. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 11444 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 114443P, December 2020.

[115] Katharine C. Walker, Bradley E. Schaefer, and E. E. Fenimore. Gamma-ray bursts have millisecond variability. 10 1998.

[116] K.C. Walker et al. Gamma-Ray Bursts Have Millisecond Variability. *The Astrophysical Journal*, 537:264–269, July 2000.

[117] K. Ward, I. Eckley, and P. Fearnhead. Fast online changepoint detection for high-velocity count data. unpublished, 2021.

[118] Samuel S Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The annals of mathematical statistics*, 9(1):60–62, 1938.

[119] Christoph Winkler, G Di Cocco, N Gehrels, A Giménez, S Grebenev, W Hermsen, JM Mas-Hesse, F Lebrun, N Lund, GGC Palumbo, et al. The integral mission. *Astronomy & Astrophysics*, 411(1):L1–L6, 2003.

[120] S. E. Woosley. Gamma-ray bursts from stellar mass accretion disks around black holes. *The Astrophysical Journal*, 405:273–277, March 1993.

[121] Takayuki Yanagida, Kei Kamada, Yutaka Fujimoto, Hideki Yagi, and Takagimi Yanagitani. Comparative study of ceramic and single crystal ce: Gagg scintillator. *Optical Materials*, 35(12):2480–2485, 2013.

[122] M. Yoneyama, J. Kataoka, M. Arimoto, T. Masuda, M. Yoshino, K. Kamada, A. Yoshikawa, H. Sato, and Y. Usuki. Evaluation of GAGG:ce scintillators for future space applications. *Journal of Instrumentation*, 13(02):P02023–P02023, feb 2018.

[123] B.B. Zhang et al. A Comprehensive Analysis of Fermi Gamma-ray Burst Data. I. Spectral Components and the Possible Physical Origins of LAT/GBM GRBs. *The Astrophysical Journal*, 730:141, April 2011.

[124] Weiqun Zhang, SE Woosley, and AI MacFadyen. Relativistic jets in collapsars. *The Astrophysical Journal*, 586(1):356, 2003.