Statistical Investigations of the Emission Processes in Gamma-ray Bursts

Doctoral Thesis in Physics
Stockholm, Sweden 2019
Doctoral Thesis in Physics

Statistical Investigations of the Emission Processes in Gamma-ray Bursts

Particle and Astroparticle Physics, Department of Physics
Royal Institute of Technology, SE-106 91 Stockholm, Sweden

Stockholm, Sweden 2019
Akademisk avhandling som med tillstånd av Kungliga Tekniska högskolan i Stockholm framlägges till offentlig granskning för avläggande av teknologie doktorsexamen den 28 januari 2020 kl. 13.00 i sal FB42, AlbaNova Universitetscentrum, Roslagstullsbacken 21, Stockholm.

Avhandlingen försvaras på engelska.

TRITA–SCI-FOU 2019:61
© Zeynep Acuner, December 2019
Printed by Universitetsservice US-AB
Abstract

Physical emission mechanisms responsible for gamma-ray bursts (GRBs) remain elusive to this day, 50 years after their discovery. Although there are well studied physical models, their power to explain the observed data is a matter of debate. In this thesis, the main focus is the statistical studies of the different physical models given the available data from the Fermi Gamma-Ray Space Observatory to make better comparisons between these models as well as ascertaining how well they can explain the available observations so far. To this end, theoretically predicted thermal and non-thermal GRB spectra are investigated. This investigation entails both finding groupings in the catalog data (clustering) and then simulating the expected physical emission processes to test how they would look like in the current data acquisition, processing and fitting procedures. Finally, a Bayesian model comparison is performed in a sub-sample of these bursts to quantify the preference of different models by the data. In conclusion, it is found that around one third of all bursts include intervals where the emission is from a photosphere which is non-dissipative. This means that during these intervals, the emission is either emitted close to the saturation radius or in a flow which is laminar. The results further indicate that dissipation below the photosphere is responsible for the spectral shape in a majority of GRB spectra. It is consequently argued that the dominant emission mechanism during the prompt emission phase in GRBs is thermal emission from the jet photosphere at distance of around $10^{12}$ cm from the central engine. A small percentage of the bursts are better explained with a non-thermal generating process such as the synchrotron emission.
Sammanfattning

Contents

Abstract iii
Sammanfattning v
Contents vii
List of publications ix
Author’s contribution to the attached papers xi
Acknowledgments xiii

1 Introduction 1
   1.1 Temporal characteristics and the progenitors . . . . . . . . . . . . 2
   1.2 Spectral characteristics and energetics . . . . . . . . . . . . . . . 2
   1.3 Gravitational waves . . . . . . . . . . . . . . . . . . . . . . . . . 5
   1.4 Highest energy photons from GRBs . . . . . . . . . . . . . . . . . 5

2 Fermi - GBM : the instrument and the data analysis 7
   2.1 Gamma-ray Burst Monitor . . . . . . . . . . . . . . . . . . . . . . 7
   2.2 GBM data products . . . . . . . . . . . . . . . . . . . . . . . . . . 8
   2.3 Data reduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
   2.4 Data analysis and forward folding . . . . . . . . . . . . . . . . . . 10
   2.5 Empirical models . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

3 Physical emission processes 13
   3.1 The Fireball model . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
   3.2 Photospheric emission . . . . . . . . . . . . . . . . . . . . . . . . . 14
      3.2.1 Coasting phase . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
      3.2.2 Accelerating phase . . . . . . . . . . . . . . . . . . . . . . . . . 16
   3.3 Synchrotron emission . . . . . . . . . . . . . . . . . . . . . . . . . . 17
   3.4 Dissipation mechanisms . . . . . . . . . . . . . . . . . . . . . . . . 20
      3.4.1 External and internal shocks . . . . . . . . . . . . . . . . . . . . 20
      3.4.2 Sub-photospheric dissipation . . . . . . . . . . . . . . . . . . . . 21
3.5 Alternatives to the Fireball model ........................................ 23

4 Statistical methods ......................................................... 25
  4.1 Overview ........................................................................... 25
  4.2 Paper 1: Unsupervised machine learning via clustering .......... 25
  4.3 Bayesian Data Analysis ..................................................... 30
    4.3.1 Paper 2: Bayesian parameter estimation ..................... 30
    4.3.2 Paper 3: Bayesian model comparison ......................... 33
    4.3.3 Information Criteria .................................................. 38

5 Discussion ........................................................................... 41
  5.1 Clusters in GRB catalogs .................................................. 41
  5.2 Spectral shape due to subphotospheric dissipation ............... 44
  5.3 Model comparison ........................................................... 44
  5.4 Catalogue parameter values .............................................. 45
  5.5 Synchrotron model fits ..................................................... 47
  5.6 Polarization ................................................................. 48
  5.7 Time variability ............................................................. 48
  5.8 The general view ........................................................... 49

6 Outlook .............................................................................. 53

List of figures ........................................................................ 55

Bibliography ......................................................................... 59
List of publications

Publications included in this thesis

Paper 1


Paper 2


Paper 3

To be submitted to ApJ.
Publications not included in this thesis

Paper 4

Paper 5
Author’s contribution to the attached papers

Paper 1

The formulation of the problem and the method is done by me. The data reduction and analysis as well as the statistical analysis were done by me. The manuscript was mainly written by me with some smaller contribution from Felix Ryde. The figures are done mainly by me with the assistance of Felix Ryde.

Paper 2

The project was conceived from discussions between me and Felix Ryde. The data simulation and analysis as well as the statistical analysis were done by me. The manuscript was written by me and Felix Ryde. The figures are done mainly by me with the assistance of Felix Ryde. Hoi-Fung Yu provided assistance on the use of the catalogue by Yu, Dereli-Bégué, and Ryde (2019). The interpretation was a result of discussions among the authors.

Paper 3

The formulation of the problem was done by me and the development of the project was cooperative between me and Felix Ryde. The data reduction and analysis as well as the statistical analysis were done by me. The formulation of the statistical method was done by me with some advice given by Daniel Mortlock. The manuscript was written by me and Felix Ryde. The figures are done by me, Felix Ryde and Björn Ahlgren. The interpretation of the results was done collectively.
Acknowledgments

I would like to thank my supervisor, Felix Ryde, for his continuous support throughout my PhD. I greatly value all the things I have learned through his help and advising.

I am grateful for my family’s full support in every interest I have taken, including Astrophysics. They have always valued logic and wisdom which gave me the best possible mindset for doing science.

I rejoice all friends and colleagues, here and abroad, that have been a part of this interesting and challenging time of my life.
Chapter 1

Introduction

Gamma-ray bursts (GRBs) are arguably the most majestic shows that are staged by our Universe. Temporarily outshining all the light in gamma rays, a GRB dominates the photon realm of the Universe, albeit for a short time. GRBs deliver huge energy outbursts that last from a few seconds up to a few hundred seconds. Earth, being our safe cradle in a very violent universe, shields us from gamma-rays. Hence these bright flashes of light are best viewed from outside the Earth’s atmosphere and this is how they were first discovered in the late 1960s by the U.S. Vela satellites. These satellites were placed in the orbit to observe possible nuclear tests that might be carried out by the Soviet Union after the agreement of the two countries over the Nuclear Test Ban Treaty in 1963. Starting from 1967, they kept observing these short-lived flashes of gamma-rays but it wasn’t until 1973 that the findings were declassified and published by Klebesadel, Strong, and Olson (1973), which was the first report of these mysterious gamma-ray flashes of cosmic origin.

Launched into the orbit around the Earth, several satellites have observed the sky for these incredible events that take place almost daily. The first major instrument to explore GRBs was the Burst and Transient Source Experiment (BATSE) (Fishman, 1992) which was launched in 1991 on board the Compton Gamma Ray Observatory (CGRO) (Gehrels, Chipman, and Kniffen, 1994). Throughout 9 years BATSE detected more than 2700 bursts showing an isotropic sky distribution. Hinting to a cosmological origin, the actual cosmic distance to GRBs, the redshifts, were first calculated thanks to the data from Beppo-SAX (Boella et al., 1997). Beppo-SAX was able to observe the afterglows of the GRBs for the first time in 1997. Along with the isotropical distribution of GRBs, this solidified the GRB’s cosmological origins. Later, HETE–2 (Ricker, 1997) provided the evidence for the relationship of the GRBs to supernovae. Currently operational satellites include Integral (Jensen et al., 2003), Swift (Burrows et al., 2005) and the Fermi Gamma-ray Space Telescope (The Fermi-LAT collaboration, 2019), launched in 2002, 2004 and 2008 respectively.
1.1 Temporal characteristics and the progenitors

A few characteristics of the spectra and the lightcurves have been fundamental in describing the different populations that might exist among the observed GRBs so far.

The observed lightcurves are quite distinct for each burst (see Figure 1.1), but the time interval that 90 per cent of the fluence is accumulated, called $T_{90}$, separates distinctly into two groups. Short bursts with $T_{90} \lesssim 2$ seconds and long bursts $T_{90} \gtrsim 2$ seconds were first identified by Kouveliotou et al. (1993) and to this date is the preliminary classification used for describing GRBs.

Short bursts are thought to be originating from the merging of two compact objects, such as a neutron star-neutron star or a neutron star and a black hole merger. Long bursts, on the other hand, are tied to the end stages of stars and the formation of a black hole through a stellar collapse, the so called core collapse supernovae. The idea that stars going through core collapse are associated with GRBs came from the joint observations of GRBs with supernovae, for the first time in 1998 (Galama et al., 1998). Since then, more joint observations have verified this relationship.

The lightcurves observed are generally in the form of a fast rising slow decaying count rate in time, with single or multiple peaks. The number and the time duration of GRB pulses are highly variable. A broad log-normal distribution can be observed for pulse profile parameters, within and across bursts (Piran, 2004).

Another time series related property that has been extensively discussed is the time variability in the lightcurves of the GRBs. Different variability timescales indicate different regions of emission, which has been fundamental in making sense of the morphology of the events.

1.2 Spectral characteristics and energetics

Observed fluences from GRBs, in combination with their cosmological origin, reveal the energetics of these events being of the order of $10^{51-52}$ ergs/second, which is comparable to supernovae in terms of energy release (Piran, 2004). GRBs are observed both in gamma-rays and lower energy photons at later times, mainly in X-rays but also in optical and radio wavelengths. This later, lower energy emission is called the afterglow and it has been instrumental for the assessment of redshifts for determining the energetics as well as for determining the burst environment, which in turn aids in pointing out the host galaxy.

However, most of the energy emitted from GRBs is in the form of gamma-rays that is mainly released within a few seconds. This is called the prompt phase of GRBs. The gamma-ray spectra observed during the prompt phase are typically non-thermal and well described by a smoothly broken powerlaw function. The empirical Band function (Band, Matteson, and al., 1993) is widely used to derive physical descriptions from GRB spectra (Figure 1.2). It has three parameters:

\footnote{Figure taken from https://www.slac.stanford.edu/exp/glast}
1.2. Spectral characteristics and energetics

Figure 1.1. Diversity of the light curves and variability time scales for different BATSE bursts. Figure taken from Fishman and Meegan [1995].
Chapter 1. Introduction

Chapter 1. Introduction

Figure 1.2. The Band function. Superimposed are its index parameters and the energy ranges from GBM and LAT instruments.

Figure 1.3. Distributions of Band function for $\alpha$ and $E_{pk}$ parameters from the Fermi GBM catalog.

the low energy powerlaw index ($\alpha$), the high energy powerlaw index ($\beta$) and the break energy $E_{pk}$ (see Section 2.5 for the full function). The power law indices are defined in photon flux $N_E \propto E^{\alpha}$. Note that the plots are usually presented as $\nu F_{\nu} \equiv E F_E$ where $F_E$ is the energy flux. The slopes in the $\nu F_{\nu}$ plots become $\alpha + 2$. The parameter $E_{pk}$ is defined as the peak energy in the $\nu F_{\nu}$ plots.

Figure 1.3 shows the up to date low energy photon index and $E_{pk}$ distribution from the Fermi GBM catalog obtained from the Band function. The distribution of $\alpha$ values clearly peak at non-thermal indices, at around -0.7. This is the reason for the attempt to explain GRB spectra primarily with synchrotron emission (see Section 3.3). However, very hard $\alpha$ values, above -2/3, have been difficult to reconcile with any type of synchrotron emission due to the spectral narrowness of these bursts. This point will be discussed further in the thesis.
1.3 Gravitational waves

17th of August 2017 marks the beginning of a new era in Astrophysics due to the first ever detection of a gravitational wave signal from merging neutron stars observed by both LIGO and Virgo instruments (Abbott et al., 2017). This event was special because it was a joint detection of gravitational waves and an electromagnetic counterpart, namely a short gamma-ray burst, both coming from the vicinity of the galaxy NGC 4993. This observation decisively confirmed the compact object mergers as short GRB progenitors. There is still some doubt about black hole - neutron star or black hole - black hole mergers’ ability to create any electromagnetic counterpart that can be observed. With 11 confirmed merger observations from LIGO and future missions being planned to focus on the electromagnetic counterparts of compact object mergers such as Nimble (Barclay et al., 2019) and BurstCube (Racusin et al., 2017), it is only a matter of time before more about these objects are discovered.

1.4 Highest energy photons from GRBs

Recently, photons above 100 GeV were observed for two bright gamma-ray bursts. These observations were done by the Major Atmospheric Gamma Imaging Cherenkov (MAGIC) Telescope (Baixeras, 2003) and the High Energy Stereoscopic System (HESS) array of telescopes (Hinton and HESS Collaboration, 2004). These detections made it possible to model the synchrotron and synchrotron self-Compton components in the GRB afterglow spectra. The specific spectral shape produced by these two components were predicted however not observed so far due to the intricacies of the observational methods. The detection of TeV photons unlocked a novel aspect of GRBs which provides a wealth of new information to be discovered (MAGIC Collaboration et al., 2019).
Chapter 2

Fermi - GBM: the instrument and the data analysis

Fermi Gamma-Ray Space Telescope (FGST) has been observing high-energy gamma-rays since its launch on 11th of June 2008. The Fermi satellite has two instruments on board, the Large Area Telescope (LAT) (Atwood, Abdo, and al., 2009) and the Gamma-ray burst Monitor (GBM) (Meegan et al., 2009). The main instrument is the pair conversion telescope LAT, which operates in the energy range of $\sim 20$ MeV to 300 GeV. This thesis makes use of the data from the complementary instrument GBM which is sensitive to both X-rays and gamma-rays, working in the energy range of 8 keV to 40 MeV. In this section, I give more detailed information about GBM and how the GBM data is processed and analysed.

2.1 Gamma-ray Burst Monitor

Gamma-ray Burst Monitor (GBM) is built to carry out spectral and time series analysis of gamma-ray emission jointly with LAT. It consists of twelve sodium iodide (NaI) and two bismuth germanate (BGO) scintillators. The NaI detectors observe the energy range $\sim 8$ keV to $\sim 1$ MeV while the BGO detectors cover the higher energy range of $\sim 150$ keV to $\sim 30$ MeV. The energy resolution is 15 per cent at 100 keV and 10 per cent at 1 MeV. These energy ranges help to cover the gap between the commonly available hard X-ray data and the high energy observations obtained from the LAT. A detection for GRBs require a distinct change of the count rate in at least two of the NaI detectors which have an on board threshold of $\sim 0.7$ photons cm$^{-2}$s$^{-1}$ with 2.6 micro-seconds of dead time per event. With this configuration, the GBM triggers for $\sim 250$ GRBs per year and to date there have been 2684 detections classified as GRBs. A secondary task of GBM is to localize the bursts detected both by itself and the LAT. The on-board burst location error for GBM is less than 15 degrees (Goldstein et al., 2012). For strong bursts, this
enables a re-orientation of the spacecraft which are then observed by the LAT for an extended stretch of time.

Besides GRBs, GBM triggers for solar flares, terrestrial gamma flashes (TGFs) and soft gamma-ray repeaters (SGRs). The background data that the GBM collects are used for the studies of hard X-ray pulsars.

### 2.2 GBM data products

GBM collects three types of science data. CSPEC data provides continuous high spectral resolution with a time resolution of 4.096 seconds, consisting of 128 energy channels with adjustable channel boundaries. CTIME data also provides continuous high time resolution with a nominal resolution of 0.256 seconds. This resolution changes during bursts to 1.024 seconds and can be adjusted between 1.024 and 32.768 seconds. CTIME consists of 8 energy channels with adjustable channel boundaries. TTE data are collected as time-tagged events with a time resolution of 0.064 seconds during bursts and an adjustable range of 0.064 to 1.024 seconds. It allows for 2 micro-second time tags for 300 seconds after the trigger and 500,000
2.3. Data reduction

In this thesis, XSPEC (Arnaud, 1996) is used for Paper 1 and the Multi-Mission Maximum Likelihood 3ML framework (Vianello et al., 2017) is used for Papers 2 and 3 for the data reduction and analysis procedure.

The data analysis presented in this thesis makes use of the TTE files to create spectra from each detector. All data has been taken from the HEASARC\(^1\) website which is a NASA repository for high-level, high-energy satellite data. From these, two or three NaI and one BGO spectrum is utilised to proceed with the modelling. The detectors with the least angle to the line of observation of the bursts (smallest source viewing angles) are chosen, ie. less than 60 degrees.

![Light Curve](image)

**Figure 2.2.** GRB100707 lightcurves for one NaI and one BGO detector, labeled with the background selections, the background fit with the polynomial order of 1 and the source selection. The source selection is the orange region and it is the time bin for which the photon spectrum is analyzed.

---

\(^1\)https://heasarc.gsfc.nasa.gov/W3Browse/fermi/fermigrbst.html
Following the selection of the detectors, source and background intervals are selected. Source time intervals are selected such that the analysed spectra are time resolved (either peak flux time or intervals as described in Yu, Dereli-Bégue, and Ryde (2019). The background intervals are obtained from the GBM catalogue unless otherwise stated. These are then fitted with a low order polynomial determined by a likelihood test and integrated over time to estimate the total count rate from the background in each channel and the errors on these background counts.

![Figure 2.2](image1)
![Figure 2.3](image2)

**Figure 2.2** shows one such selection made for a NaI and a BGO lightcurve from GRB100707 and **Figure 2.3** shows the count spectrum of the selected region with two different models described below. The residuals show that the first model is preferred by the data.

### 2.4 Data analysis and forward folding

Theoretical deliberations in the GRB community mainly take into account the energy spectra of the bursts. The discussion and the interpretations mostly focus on the energy spectral peak and the low energy index. To extract an energy spectrum from the observed counts per energy channel, the following method is utilized. The first step is to obtain the count spectrum itself, which is the sum of the burst flux convolved with the GBM response and the background. Here, the response matrix is simply the relationship between the photon’s true energy and the observed count’s apparent energy. The expression of counts and the fluxes as vectors gives the Detector Response Matrix (DRM), as mentioned in Section 2.2. The DRM is not a square matrix and the equation linking counts to the energy spectrum cannot be solved by inverting this matrix.

---

2 The GRB name denotes the date it was detected in year, month and day respectively ie, GRB-yymmdd.
2.5. Empirical models

This is why the method of forward folding is used for obtaining GBM energy spectra. In the forward folding method, the model flux is folded through the response and the model count spectrum is obtained. The model counts are then compared to the observed counts. By varying the spectral parameters, a new model flux vector is calculated that replicates the observed counts better. When the two sets of counts are sufficiently close, i.e., the minimization of the chi-squared gives a sufficiently low value\(^3\), this iterative process is ended\(^4\). If there are more than one count spectra to be fitted with the same flux model, this is done through the same process as a joint fit.

2.5 Empirical models

There are two empirical models that are used throughout the GRB community and in the papers this thesis is compiled from. The Band function as defined in Band, Matteson, and al. (1993) is a smoothly broken powerlaw that can mimic the spectral shape of the gamma-ray emission coming from GRBs, in the energy range of 8 keV to 1 MeV. The function follows as,

\[
N_{E}^{\text{Band}} = A \begin{cases} 
\left( \frac{E}{100\text{keV}} \right)^{\alpha} e^{\left[ \frac{-(\alpha+2)E}{E_{pk}} \right]} : E < E_c \\
\left( \frac{E}{100\text{keV}} \right)^{\beta} e^{(\beta-\alpha) \left[ \frac{E_{c}}{100\text{keV}} \right]^{\alpha-\beta}} : E > E_c
\end{cases}
\]  

(2.1)

where \(A\) is the normalization factor at 100 keV with units of \(s^{-1}cm^{-2}keV^{-1}\), \(\alpha\) and \(\beta\) are low and high energy powerlaw photon indices respectively and \(E_c = \frac{\alpha-\beta}{\alpha+2}E_{pk}\). Note that \(E_{pk}\) is the peak energy in \(\nu F_{\nu}\) space and \(E_c\) is the characteristic energy, both in the units of keV.

The second model is called the cut-off powerlaw (CPL) or the Comptonized model. CPL is a powerlaw with a high-energy exponential cut-off. The relation is,

\[
N_{E}^{\text{CPL}} = A \left( \frac{E}{100\text{keV}} \right)^{\alpha} e^{\left[ \frac{-(\alpha+2)E}{E_{pk}} \right]} 
\]  

(2.2)

with the defined parameters same as that of the Band function. Both models are in the units of \(s^{-1}cm^{-2}keV^{-1}\).

\(^3\)A value that is comparable to the number of degrees of freedom is generally assumed to be sufficiently low.

\(^4\)See https://fermi.gsfc.nasa.gov/ssc/data/analysis/
Chapter 3

Physical emission processes

In this chapter, I will discuss different physical phenomena that are thought to be generating GRBs. These include a review of the Fireball model, photospheric and synchrotron emissions as well as the subphotospheric dissipation.

3.1 The Fireball model

In order to account for the observed fluences as well as the time-scales of GRBs originating from cosmological distances, the initial suggestion was that they should start off in very compact regions as fireballs of baryons, electrons, positrons and photons which are expanding relativistically.

In this scenario, the dense ball of matter and radiation is formed due to the stellar core collapse. There is a sudden release of gravitational energy as well as neutrinos, leptons and baryons. A small fraction of these particles goes on to form the high temperature fireball with $kT > \text{MeV}$. When the radiation pressure of this thermalized fireball is large enough so that it cannot be compressed further, it starts adiabatically expanding. This means there is no energy loss to dissipation yet and the thermal photons dominate the available energy in the fireball. During this accelerating phase, the bulk Lorentz factor $\Gamma$ is proportional to the expansion radius $r$. The expansion is due to the thermal pressure and temperature diminishes as $r$ increases. At a certain $r$, called the saturation radius, thermal photons start losing significant energy to the expansion and are no longer able to support the acceleration. Here, the fireball enters the so called coasting phase where $\Gamma$ is constant. At this phase, energy of the fireball is mostly in the form of kinetic energy carried by the electrons (see Figure 3.1).

The dimensionless entropy is an important parameter to showcase this relationship between the thermal energy in the fireball versus the total rest mass energy,

$$\eta = \frac{L}{M c^2}$$  (3.1)
where \( c \) is the speed of light, \( L \) is the total luminosity and \( \dot{M} \) is the mass outflow rate. In the coasting phase, \( \Gamma = \eta \).

The energy density is given by the Stefan-Boltzmann law as \( aT^4 \) where \( a \) is the radiation constant which is equal to \( L/(4\pi r_0^2c) \). The estimation of the temperature in the fireball base with a radius \( r_0 \) can then be made as \( T_0 = (L/4\pi r_0^2c)^{1/4} \).

The above narrative creates the expectation of a photosphere formation at the early times which would be a black-body. However, as mentioned in Section 1, the observed gamma-ray spectrum is generally a smoothly broken powerlaw which is highly non-thermal. Furthermore, the efficiency for the observable gamma-ray production is very low in this case because of the fact that most of the energy is deposited in the form of the kinetic energy of protons.

### 3.2 Photospheric emission

The assumption of the fireball model leads to the reasoning that there should be a major thermal component to the spectra of the GRBs, especially in the early times. The non-thermal spectra observed from GRBs have diminished this view for some while. However, it is possible to observe very narrow thermal components in GRBs with extremely low baryon loading (Beloborodov, 2011).
3.2. Photospheric emission

Photospheric emission stems from the photosphere where the optical depth in the flow is unity and the photons are released. Hence, in order to estimate the radius of the photosphere, the optical depth needs to be defined. The optical depth for a photon along the line of sight of the observer from $r$ to $\infty$ in a dense environment dominated by electron scattering is given as,

$$\tau(r) = \int_r^\infty \frac{n \sigma_T}{2 \Gamma^2} \, dr$$  \hspace{1cm} (3.2)

where $n$ is the electron number density and $\sigma_T$ is the Thomson cross-section. Note that the photon propagates almost radially due to relativistic effects (Lundman, 2013). Assuming the Lorentz factor is constant, this integration gives,

$$\tau(r) = \frac{L \sigma_T}{8 \pi m_p c^3 \Gamma^3} \frac{1}{r}$$  \hspace{1cm} (3.3)

where $m_p$ is the mass of a proton. The photosphere is defined as radius where $\tau = 1$ and is therefore,

$$r_{ph} = \frac{L \sigma_T}{8 \pi m_p c^3 \Gamma^3}$$  \hspace{1cm} (3.4)

Here, the assumption of constant $\Gamma$ requires the flow to be in the coasting phase above the saturation radius (Pe’er, Ryde, and al., 2007; Iyyani, 2015). As can be seen, $r_{ph}$ depends on the variables $L$ and $\Gamma$. Following from Eqn. 3.1, $r_{ph}$ is proportional to $\dot{M}/\Gamma^2$. For the typical value of $L \approx 10^{52}$ ergs/s, the equation becomes,

$$r_{ph} = \frac{10^{52}(\text{erg/s}) \sigma_T}{8 \pi m_p c^3 \Gamma^3}$$  \hspace{1cm} (3.5)

and if $\Gamma \approx 300$, $r_{ph}$ above $r_s$ is around $10^{12}$ cm.

Paczynski (1986) and Goodman (1986) were the first ones to discuss the possibility of a photosphere in GRBs. They considered thermal radiation dominated flows with low baryon loading (Paczynski, 1986; Goodman, 1986) and matter-dominated (Paczynski, 1990) flows producing a multicolor blackbody spectrum. All of these works have concluded that the final photospheric spectrum observed differs from an exact blackbody spectrum due to relativistic effects.

3.2.1 Coasting phase

In the matter-dominated regime it turns out that a Planck spectrum can never be obtained even if there is no heating in the flow. Therefore a thermal component can have at most a low energy spectral index of 0.4, compared to the Rayleigh-Jean limit of 1, i.e. a blackbody. The photospheric thermal component becomes broader due to several effects.

Angle dependency of the photospheric emission contributes to the broadening effect without the need for dissipation. Broadening is also caused by differences
Chapter 3. Physical emission processes

Figure 3.2. The analytical (dashed line) and the expected blackbody emissions from GRBs. Figure taken from Goodman (1986).

...of the Doppler boosts for different photons, especially from higher altitudes to the line of sight of the observer (Abramowicz, Novikov, and Paczynski, 1991; Pe’er, 2008). The last scattering radii and angles of the photons in the photosphere are different which causes the observer to see spectra of different temperatures at the same time. This in turn creates a broadened blackbody shape (Lundman, Pe’er, and Ryde, 2013).

Different dissipation mechanisms can affect the photospheric spectra by making it broader than an analytic blackbody function (Pe’er, Mészáros, and Rees, 2006). These mechanisms and their effects are discussed later in this chapter.

If the photosphere occurs above the saturation radius, a broader spectrum is obtained, called the coasting phase NDP. The analytical function for this spectrum is,

$$N_E = K \left( \frac{E}{E_{\text{pivot}}} \right)^{0.4} e^{-\left( \frac{E}{E_c} \right)^{0.65}}. \quad (3.6)$$

where $N_E$ is the photon flux (unit), $E_{\text{pivot}}$ is the pivot energy, $E_c$ is the cut-off energy and $K$ is the normalization factor.

3.2.2 Accelerating phase

In the radiation dominated regime, the spectrum is affected by the photon angular distribution becoming more and more anisotropic as they approach the photosphere (Beloborodov, 2010).
Beloborodov (2011) showed that when the photosphere occurs far below the saturation radius, the narrowest possible spectrum is obtained. In a thermally accelerated flow, the temperature in the observer frame remains constant which discards all the broadening effects related to the distribution of radii and angles of last scattering. This is due to the exact compensation of the broadening created by the angular distribution of the radiation field in the lab frame by the Lorentz transformation of angles in the local comoving frame for a flow with $\Gamma \propto r$.

This is called an accelerating phase non-dissipative photosphere (NDP) in Paper 2. The analytical function for this spectrum is,

$$N_E = K \left( \frac{E}{E_{\text{pivot}}} \right)^{0.66} e^{-\left( \frac{E}{E_c} \right)}$$  \hspace{1cm} (3.7)

where $N_E$ is the photon flux (unit), $E_{\text{pivot}}$ is the pivot energy and $E_c$ is the cut-off energy.

All of these spectral (Eqs. 3.6 and 3.7) models are numerically calculated in (Lundman, Pe’er, and Ryde, 2013) and then analytically approximated from the table models derived from the simulations. Figure 3.3 shows these two spectra with a Planck spectrum for comparison.

### 3.3 Synchrotron emission

Synchrotron emission is the physical emission process that best explains the external shocks that creates the afterglow (Sari, Piran, and Narayan, 1998). Therefore, it
was suggested that it could directly explain the non-thermal nature of the GRB prompt spectra. However, it turned out that it has particular problems due to the prompt emission efficiency, time variability and steepness of the low energy powerlaw slope ($\alpha$) in the spectra of the observed GRBs. However, it is still argued that synchrotron radiation is responsible for some prompt GRB spectra, if not all (Ravasio et al., 2018; Oganesyan et al., 2019; Burgess et al., 2019).

Energy extracted in the form of photons in a synchrotron scenario depends on two physical parameters, namely the energy and cooling time scale which are in turn dependent on the Lorentz factor of these relativistic electrons ($\gamma_e$). The photons are generated by the electrons that have been accelerated due to the kinetic energy dissipation at a certain radius $r_d$. The peak energy of these synchrotron photons is given as,

$$E_{\text{synch}} = \frac{3}{2} \frac{\hbar q B}{m_e c} \gamma_e^2 \Gamma \frac{\Gamma}{(1 + z)}$$  \hspace{1cm} (3.8)

where $m_e$ is the mass and $q_e$ is the charge of a single electron, $B$ is the strength of the magnetic field, $\sigma_T$ is the Thompson cross-section, $\hbar$ is the reduced Planck function, $z$ is the redshift and finally $c$ is the speed of light in vacuum.

The observed synchrotron flux then can be written as,

$$F_{\text{synch}} = \frac{\sigma_T c \Gamma^2 \gamma_e^2 B^2 N_e}{24 \pi^2 d_L^2}$$  \hspace{1cm} (3.9)

where $N_e$ is the number of radiating electrons and $d_L$ is the luminosity distance. The cooling time scale of the accelerated electrons emitting synchrotron photons is,

$$t_{\text{cool}} = \frac{6\pi m_e c}{\sigma_T B^2 \Gamma \gamma_e (1 + Y)}$$  \hspace{1cm} (3.10)

where $Y$ is the Compton $Y$ parameter that is defined as the number of scatterings times the energy gain per scattering and hence gives the energy gain due to Compton scatterings. The cooling time scale should be compared to the dynamical time scale to assess the efficiency of the cooling of the energetic electrons. The dynamical time-scale is given as,

$$t_{\text{dyn}} = \frac{R}{2 \Gamma^2 c}$$  \hspace{1cm} (3.11)

Therefore, the fast cooling regime is defined as $t_{\text{cool}} < t_{\text{dyn}}$ and the slow cooling regime is defined as $t_{\text{cool}} > t_{\text{dyn}}$ (Tavani, 1996).

In the fast cooling scenario, the electrons with modest energies are cooling rapidly so it is an efficient mechanism. In the slow cooling scenario, only the high energy tail of the electrons are able to cool which decreases the energy output of the synchrotron radiation. In principle, during the GRB prompt emission, any synchrotron related radiation should be coming from the fast (or moderately fast) cooling electrons in order to avoid an efficiency problem. Slow cooling takes place at a later time, possibly with the external shocks, creating the GRB afterglow.
3.3. *Synchrotron emission*

When electrons in the shock are accelerated to a minimum energy of $\gamma_{\text{min}}$ with a power-law distribution of $dN/d\gamma$ proportional to $\gamma_{\text{min}}^{-p}$, the emission spectrum is also a powerlaw with a low energy cut-off at $\nu_{\text{min}} \propto B\gamma_{\text{min}}^2$. Slope of the powerlaw below $\gamma_{\text{min}}$ is the characteristic slope of the synchrotron emission. Figure 3.4 shows these characteristic indices for slow, fast and marginally fast cooling (where $\gamma_{\text{min}} \approx \gamma_{\text{cool}}$) synchrotron emissions (Yu et al., 2015).

The expected spectral low energy slopes ($\alpha$) for slow and fast cooling synchrotron emission are $\sim -2/3$ and $-3/2$, respectively. In Figure 3.5, these values are plotted over the $\alpha$ distribution from the GBM catalog.

Electrons could absorb back the synchrotron photons they emitted if they are not highly energetic. This is called the synchrotron self-absorption and could affect the spectrum by leading to a steep cut-off at the lower energies. For typical parameters, this does not occur at high energies so it mostly affects the GRB afterglow.
3.4 Dissipation mechanisms

3.4.1 External and internal shocks

One way to form a non-thermal spectrum without efficiency problems is the creation of an external shock that would turn the kinetic energy deposited in the protons to random energy. These shocks occur in optically thin regions and are collisionless. Commonly observed in supernova remnants and active galactic nuclei, this was the first assumption for GRBs as well (Rees and Meszaros, [1992]). As the fireball expands through the interstellar medium, it would interact with existing material creating blast waves. The particles would then go through the Fermi process and are accelerated to ultra-relativitic energies. This would cause the highly energetic electrons to produce non-thermal radiation via synchrotron and inverse Compton processes. These are called external shocks and are strictly non-thermal processes. When sufficiently less energetic, these external shocks go on to produce the commonly observed GRB afterglows in less energetic wavelengths, ranging from X-rays to optical.

Such a shock is expected at the deceleration radius where the blast wave has swept up enough circumburst material (CMB) to start halting its expansion. This radius \((r_{dec})\) is given as a function of burst energy as,

\[
E \approx \frac{4}{3} \pi n_{ext} \Gamma^2 r_{dec}^3
\]  

(3.12)

where \(n_{ext}\) is the density of the external medium. For values of \(E = 10^{52}\) ergs, \(\Gamma \approx 300\) and \(n_{ext} \approx 1 \text{ cm}^{-2}\), \(r_{dec}\) is approximated at \(10^{16}\) cm.
3.4. Dissipation mechanisms

At this radius, the variability time is $\Delta t \sim \frac{R}{2c\Gamma^2}$ which is associated with the minimum variability time-scale of a lightcurve from an external shock emission with variabilities of the order of 10 seconds (Rees and Meszaros, 1992; Meszáróss, 2006). This is much longer than the typically observed variability time scales which are around 0.1 seconds (Golkhou and Butler, 2014; Golkhou, Butler, and Littlejohns, 2015).

An alternative is given as internal shocks which are formed by the shells of the fireball travelling at different $\Gamma$s (Rees and Meszaros, 1994). When the later emitted shell has a higher $\Gamma$ than the previous shell which is already slowing down, it can catch up with this shell and create a shock wave. This shock which occurs at much smaller radii ($\sim 10^{13}$ cm) can explain the observed variability in GRB lightcurves. The shocks would transfer some energy from protons to electrons, which can later be radiated away as synchrotron emission. However, they are very inefficient and they strain the energy budget (Kobayashi, Piran, and Sari, 1997).

In either type of dissipation mechanism, a clumpy and irregular environment can cause multiple complex peaks and time variabilities whereas a homogeneous environment is expected to produce a single and smooth pulse (Bhat et al., 2012).

3.4.2 Sub-photospheric dissipation

An improvement to considering only external shocks as the main dissipation process is the introduction of the internal shocks that can provide the high variability observed. However, they have a major drawback as they are inefficient to produce the observed energetics of the GRBs.

To solve this problem, Rees and Meszáróss (2005) suggested that dissipation can also occur in optically thick regions of the flow (see also Pe’er, Meszáróss, and Rees (2006)). In this picture, the main responsible mechanism for the dissipation is the Comptonization of the thermal photons. Other dissipative effects include magnetic reconnection, neutron decay and shocks occuring in the optically thick regions which would enhance the thermal energy deposited in the photons.

When internal shocks occur in optically thick dense regions, the emitted photons from the electrons could quickly re-thermalize. If the internal shocks take place at larger radii, it is harder and harder for the photons to thermalize and hence, the observed spectrum gradually becomes more non-thermal.

The peak energy would then be formed by the Comptonized thermal radiation in the photons. Furthermore, due to the optically thick region, dissipation from internal shocks can be thermalised sufficiently enough to create a multicolor black-body. See Figure 3.6 for a comparison between different spectral shapes that can be created by the flow under the sub-photospheric dissipation model.

Whatever the reason for dissipation is, the photospheric component will be quite strong below the saturation radius, due to the fact that the growth of the Lorentz factor compensates the energy loss to the adiabatic expansion. When the flow is in the coasting phase with constant $\Gamma$, the thermal component from the photosphere will get weaker and weaker as the flow expands.
Figure 3.6. Schematic of different processes producing a spectrum in the comoving frame. Thermal component and its comptonized component occur closer to the start of the outflow whereas the synchrotron component comes from farther out. There could be another component at higher energies due to the shock with pair formation as described in the text. Note that an ideal black-body would be narrower than the thermal spectrum shown here. Figure taken from Rees and Mészáros (2005).
When dissipation occurs in parts of the flow, it is unavoidable that relativistic electrons lose some of their energy to photons via Compton scattering. High energy electrons can be dominated by synchrotron losses but a good part of the losses for the average energy photons are via Compton scattering since synchrotron emission is suppressed via self absorption. When the optical depth is larger than unity and with strong shocks that generate high energy photons, pair production can thermalize the flow as they lose energy via Compton scattering as well. This would produce a secondary photosphere outside the original photosphere where the main cooling mechanism is the subphotospheric dissipation.

This model can also effectively produce millisecond time variabilities as opposed to the external shock model. Here, the variability time scale is related to $\Delta t \sim R/2c\Gamma^2$ which gives variabilities of the order of $10^{-2}$ seconds. Hence, in cases with longer time variabilities, a steady central engine is required. The combination of the ability to account for non-thermal spectral shapes and the time series variability, sub-photospheric dissipation is a strong candidate for explaining the observed GRB spectra.

### 3.5 Alternatives to the Fireball model

The Fireball model assumes that the flow is dominated by radiation, especially in the first stages of the flow. However, there could also exist powerful magnetic fields in the flow that would give the most energy to the Poynting flux (Bégué and Burgess, 2016).

The dynamics of a magnetized flow are characterized in terms of the magnetization parameter $\sigma = E_{mag}/Mc^2$, $E_{mag}$ being the initial energy in magnetic fields.

In this type of a flow, internal shocks are expected to be particularly inefficient, up until to the point where the magnetic fields can be dissipated by magnetic reconnection, which requires a larger radius to be effective. Then, in principle, external shocks could provide more dissipation via synchrotron emission, as described in the Internal-Collision-Induced Magnetic Reconnection and Turbulence (ICMART) model (Zhang and Yan, 2011).
Chapter 4

Statistical methods

4.1 Overview

There are three main approaches to data analysis. These are the Classical (Frequentist), Bayesian and Exploratory (EDA) data analysis methods. Although generally considered as distinct schools of statistical thought, all three, especially Bayesianism and EDA can be used as complementary tools (Gelman, 2003). They are useful first to describe and then to get as much information as possible from any set of data with the flexibility of continuous re-assessment of the current state of knowledge and available hypotheses. At the same time, they apply strict tests on the ability of the current models to prescribe and predict future data. Below, we describe the three statistical approaches used for the papers forming this thesis. We do not explicitly discuss the Frequentist method that has been used in Paper 1 and also later on for verification reasons since it is well known and used in the astrophysical community.

4.2 Paper 1: Unsupervised machine learning via clustering

After the deployment of the Burst and Transient Source Experiment (BATSE) onboard Compton Gamma Ray Observatory (CGRO), the data from GRB sources started to aggregate rapidly. However, the physical processes leading to the emission from the prompt phase remained unaccounted for. This led the community to search for structural aspects of the GRB catalogue data to find possible hints for existing hypotheses or unexpected aspects. The overarching goal was to refine or rephrase the models to include the characteristics of the data that has been left out previously. The first study to point at a classification of GRBs was Kouveliotou et al. (1993) and they established two separate groups in $T_{90}$, namely the long ($T_{90} \gtrsim 2$ s) and short bursts ($T_{90} \lesssim 2$ s). In 1998, Mukherjee et al. (1998) and Horváth (1998) established a third class by using both a non-parametric (model-free) clustering
method (hierarchical clustering) and a model based clustering method (Gaussian Mixture Model (GMM)) mainly by analysing both spectral features and $T_{90}$. These classes can be summarized with duration/fluence/spectrum bulk properties as follows: Class 1 with long/bright/intermediate bursts, Class 2 with short/hard/faint bursts and Class 3 with intermediate/intermediate/soft bursts where classes 1 and 2 correspond to the two groups of Kouveliotou et al. (1993). Later on, different papers (Horváth, 2002; Horváth et al., 2006; Horváth et al., 2008; Tarnopolski, 2015; Horváth and Tóth, 2016) reached similar conclusions by making use of both model-free and model based clustering algorithms and different combinations of features from BATSE and Fermi GBM catalogues.

Before narrating the findings of Paper 1 and later papers regarding these findings, I briefly explain the methodology used in clustering, the tools and key points that require attention.

Our sample from the GBM catalogue consists of bursts detected until February 2017. Variables that are used in the clustering are parameters that result from the Band function fit to the spectra ($\alpha$, $\beta$, $E_{pk}$) as well as $T_{90}$ and the fluence during $T_{90}$. The Band function is chosen because although it is an empirical model, it can mimic many different spectral shapes seen in the prompt emission of GRBs. Fluence and $T_{90}$ are traditional choices that are the least instrument dependent and give very good insight to the length and energetics of the observed GRBs. After the necessary modifications and cuts, the final sample size becomes 1151.

After testing our clustering results from both non-parameteric and parametric clustering methods, we choose to use a parametric method called the Gaussian Mixture Model Based Clustering (GMMBC). There are several properties of GMBBC that makes it suitable for our data set. First of all, unlike the model-free approaches, GMMBC gives probabilities to each data point belonging to each of the clusters (soft clustering), making the analysis results much more useful in the perspective of identifying possible physical processes taking place in GRBs whether they appear distinctively or in interplay. Secondly, the normally distributed outputs makes it much more intuitive to understand the clusters by giving well known mean, median and standard deviation estimations. Caveats of this approach is the assumption of Gaussianity for the input parameters which is not always satisfied in the sample parameters that are selected. To optimize our data for clustering, we have taken several steps:

1. Making several cuts to the heavily tailed parameters, such as $\beta$, which are later verified to be optimal by an examination of the quantile-quantile plots (QQ-plots).

2. Transforming the variables which entails gathering outliers closer to the mean of each parameter distribution. This has been done via a Box-Cox transformation.

3. Transforming all negative valued variables into positive valued representatives by the addition of proper constants.
4.2. Paper 1: Unsupervised machine learning via clustering

4. Scaling the sample. This method is used so that we have unit-free variables with a comparable span of values.

5. Performing a Principle Component Analysis (PCA) which takes in a high dimensional and possibly correlated data set and produces an uncorrelated eigenvector where each element represents a mix of the original variables where eigenvalues are assigned in accordance with the amount of variance accounted for in the data for each $^1$. This step is very useful in eliminating most of the possible correlations in the sample which may result in spurious extra clusters that have no physical value.

After these steps, the sample is at its peak condition to be analysed with GMMBC. GMMBC fits a finite number of Gaussian distributions to the data and finds the optimal clustering via an Expectation-Minimization (EM) algorithm. EM is an iterative method which depends on maximum likelihood estimates of the parameters of a distribution. It alternates between calculating an estimate of the log-likelihood obtained from the current parameters and maximizing the expectation of the log-likelihood calculated from the previous step. We use the R (R Core Team, 2018) implementation mClust (Fraley and Raftery, 2002), which uses Bayesian information criteria (BIC) to assess the number and shape of the clusters. For our sample, this gives 5 clusters with different properties. These clusters are verified using Silhouette scores (Rousseeuw, 1987), which is a measure of how tightly grouped all the points in a cluster are and later (not presented in Paper 1), bootstrapping that consists of sampling from the original sample enabling re-sampling and clustering this generated sample to compare to the original clustering results. Multiple repetitions of the previous step should give similar clustering results to ensure the authenticity of the original clusters.

In accordance with the previous findings, we find three main clusters of $T_{90}$, namely the short, intermediate and long bursts. These are plotted in Figure 4.1. The long and intermediate bursts groups divide further into two subgroups mainly due to fluence. This gives five clusters as: long-bright (cluster 2), long-faint (cluster 1), intermediate-bright (cluster 4), intermediate-faint (cluster 3) and short-faint (cluster 5) bursts. Additionally, our $\alpha$ and $\beta$ features allow us to comment on the spectral shapes of these clusters and the additional time variability (denoted as $\Delta t_{min}$) data from Golkhov and Butler (2014) and Golkhov and Butler (2014) gives hints to the possible emission mechanism/s. Clusters can now be described as:

1. Long and faint with narrow spectra and short time variability ($\alpha_{med} = -0.36$),

2. Very long and bright with narrow spectra and very short time variability ($\alpha_{med} = -0.74$),

$^1$The covariance matrix is eigendecomposed or a Singular Value Decomposition (SVD) is performed. The eigenvalues are inspected. The eigenvectors with the lowest eigenvalues are discarded since they provide little additional information about the total variance of the data set. The remaining eigenvectors are then projected onto a new features space.
3. Intermediate length and faint with very broad spectra and very long time variability ($\alpha_{med} = 0.44$),

4. Intermediate length and bright with broad spectra and long time variability ($\alpha_{med} = -1.47$),

5. Short and faint with very narrow spectra and very short time variability ($\alpha_{med} = 0.7$).

Finally, by incorporating knowledge from physical emission processes (photospheric and synchrotron emissions), we combine these five clusters into two main clusters, namely a photospheric and a non-photospheric (synchrotron) group. Clusters 1, 3 and 5 are in the former and clusters 2 and 4 are in the latter group. This classification is done by comparing the $\alpha$, $E_{pk}$, fluence and variability distributions to that of theoretical predictions.

Although, the only study to date which makes use of features related to spectral shapes was Paper 1, some later publications did a similar analysis (Chattopadhyay and Maitra, 2018; Tóth, Rácz, and Horváth, 2019; Tarnopolski, 2019). I would like to briefly discuss these studies. Chattopadhyay and Maitra (2018) carries out a multivariate t-mixture-model-based-cluster (tMMBC) analysis on the BATSE 4Br catalogue with features $T_{50}$ (duration required for accumulating 50% of the total fluence, starting from 25% and ending at 75%), $T_{90}$ (duration required for
accumulating 90% of the total fluence, starting from 5% and ending at 95 %), \( P_{64} \) (peak flux at 64 ms), \( P_{256} \) (peak flux at 256 ms), \( P_{1024} \) (peak flux at 1024 ms), \( F_1 \) (fluence at 20-50 keV), \( F_2 \) (fluence at 50-100 keV), \( F_3 \) (fluence at 100-200 keV), \( F_4 \) (fluence at > 300 keV), \( F_T \) (total fluence, \( F = F_1 + F_2 + F_3 + F_4 \)) and two hardness ratios \( H_{32} = F_3/F_2 \), \( H_{321} = F_3/F_1 + F_2 \). Their number of clusters are determined by BIC and with this, they find 5 clusters. These are three long burst clusters with intermediate (hard and intermediate hard) and bright fluence and two short burst clusters which are faint but differ in hardness (one hard and one soft).

The second study of interest is Tóth, Rácz, and Horváth (2019) who carried out a GMMBC cluster analysis on BATSE 4Br catalogue using \( T_{50}, T_{90}, F_T, H_{32}, H_{321} \) and \( P_{256} \) as features. They end up with 5 clusters which are classified in terms of \( T_{90} \) and \( F_T \) as intermediate-faint, long-faint, intermediate-bright, long-bright and short-faint clusters. They discuss the validity of the 5 clusters and conclude that GMMBC cuts the short and intermediate groups into faint and bright groups spuriously. This is caused by the brightness distribution being asymmetrical and not correlated to the duration parameter. Their final grouping consists of three different classes of GRBs separated due to their duration (short, intermediate and long bursts).

Finally, Tarnopolski (2019) carries out an analysis of the duration-hardness ratio plane in GRBs by using 4 different distributions. They use both \textit{FERMI GBM} and \textit{BATSE} data with features \( T_{90} \) and \( H_{21} = F_2/F_1 \) where \( F_1 \) is the fluence at 50 to 100 keV and \( F_2 \) is the fluence at 100 to 300 keV. They do a model comparison between Gaussian, skew-normal, Student-t and skew-t distributions using both AIC and BIC values. This yields that a skewed t-distribution is the best model for clustering for this data set. They find two clusters divided by \( T_{90} \), short and long bursts.

Above, I presented four different studies that focused on clustering GRB data. These featured different data sets, variables and methods. Although the number of clusters discussed have some variability, the data clearly indicates three different groups which are long bursts, intermediate length bursts and short bursts. Long and intermediate duration bursts are further divided in two groups each as faint and bright groups. This is further supported by our results. According to the PCA analysis that is carried out on the sample, most of the variance in our sample is due to \( E_{pk}, T_{90} \) and fluence. This would mean that the clustering results are mostly influenced by these three parameters. Coming back to Paper 1, the fact that these different clusters give different \( \alpha \) distributions with quite distinct medians indicate a difference in not only between the main variable features but also between the physical emission processes that generated them. Investigating clusters in terms of GRB physics gives further validation to the clustering found by different methods and authors.

All of the methods discussed in this section fall into the branch of EDA. This sort of investigation is useful since one does not need to impose a specific model which can be hard for certain data sets of which very little is known about. The step after EDA would be to construct a proper model or models, a hypothesis space to test out the predictions that were obtained by descriptive data analysis. Using
Bayesian data analysis tools is a very suitable next step to EDA as will be explained in the next two sections.

4.3 Bayesian Data Analysis

Most astrophysical phenomena are not repeatable. Researchers in the field do not have laboratories to re-create the events over and over again to reach a Frequentist style statistical significance. This would entail repeating an experiment and registering the outcome for a lot of trial or having the time to wait for the celestial event to repeat itself. This is generally not plausible, both because the events might not repeat by their very own nature (such as, supernovae and GRBs) and ultimately, a human life span (or a few) is currently not long enough to match the vast timescales that Universe works in. Bayesian data analysis (BDA), on the other hand, treats probabilities as degrees of belief which is suitable and actually necessary for many astrophysical problems. Bayes’ theorem gives a simple way to account for our current information, ie.”beliefs”, with the use of prior distributions. This enables one to make probabilistic assessments even for one time events, directly driven from the first principles of Probability Theory (Jaynes and Bretthorst, [2003]). Furthermore, the posterior probability density distributions obtained after carrying out an analysis provides the researchers with much more than mere point estimates given by the Frequentist methods (such as the posterior probability density distributions and diagnostic plots like corner plots and posterior predictive checks). This, in turn, improves both scientific understanding and the accuracy of the inferences. There are a few other points worth mentioning that makes Bayesian framework a more robust and correct approach to almost any problem. First of all, BDA does not depend on certain distribution assumptions and the problems need not to be analytically describable. Next, the parameters which are not of interest to the analysis (called nuisance parameters) could be easily dealt with through marginalization, in a way that is in built in the Bayes’ theorem. Finally, BDA only deals with the actually observed data, leaving no room for ad-hoc presumptions that need to be made in a Frequentist approach to estimate the future observations that may or may not be observed in the name of interpreting probabilities as frequencies (Trotta, [2008]).

The following two sub-sections are related to the analysis of data under the Bayesian framework.

4.3.1 Paper 2: Bayesian parameter estimation

As mentioned in Sections 4.1 and 4.2, following the exploratory data analysis with predictive data analysis gives one to test their understanding of the data once proper hypotheses are constructed. In Paper 1, we determined different groups of GRBs which could be related to distinct physical emissions mechanisms. However, the parameters used for the clustering were derived from an empirical model. The Band function has no physical meaning and is not assumed to be the generating model for

---

2 Probability as frequency.
the observed spectra. Since the empirical models do not match the incoming photon spectra faultlessly, the output parameter values from the fits are not describing the exact observed spectrum. Although they are useful in describing the spectra, the direct physical interpretation of these parameters could be misleading in terms of finding the right physical emission processes. This was pointed out for synchrotron spectra in Burgess (2014) and is tested for non-dissipative photospheres in Paper 2. To this end, we carry out simulations of the physically expected model NDP (non-dissipative photosphere) in two different phases as explained in Section 3.1. We randomly pick a range of peak energies from the GBM catalogue which are then assigned normalizations to satisfy a certain signal to noise ratio. The photon count spectra are then convolved with the DRM of the detector to create the energy spectra. These are fit with a commonly used empirical function, cut-off powerlaw, to see how these bursts would look like in both GBM catalogue and in our clusters.

To this purpose, we introduce the Bayes’ theorem which is used to obtain the posterior distributions for these fits. When \( M \) is a model and \( D \) is the data, the Bayes’ theorem is given as,

\[
P(M \mid D) = \frac{P(D \mid M)P(M)}{P(D)}
\]

(4.1)

where \( P(D \mid M) \) is the posterior probability density distribution (PDF), \( P(D \mid M) \) is the likelihood of the observed data under the model and \( P(M) \) is the prior on the model. \( P(D) \) is called the marginal likelihood or the Bayesian evidence and generally ignored because it is treated as a normalization parameter for parameter estimation, however it plays a big role in Bayesian model comparison which is the topic of the next section (Section 4.3.2). When \( M \) is parametrised with the model parameters \( \theta \), this becomes,

\[
P(\theta \mid D, M) = \frac{P(D \mid \theta, M)P(\theta \mid M)}{P(D \mid M)}.
\]

(4.2)

Although the relation above is quite simple in itself, the analytical calculation of the posterior can be very tricky if at all possible, mainly due to the denominator. This is why BDA makes use a class of algorithms called Markov Chain Monte Carlo (MCMC). MCMC is a simulation based method that can draw random samples from the posterior distribution of a parameter of interest. ”Monte Carlo” part in MCMC is the random sampling that is carried out at each step and the ”Markov Chain” part refers to the fact that all draws depend on the previous ones. Markov chains are simulated via the Metropolis-Hastings algorithm which is a random walk that converges to the target distribution with the help of an acceptance/rejection rule. MCMC is used in cases where it is not possible to directly sample from the

\[\text{A rule that compares posterior probabilities of the previous and current sampled points. If the posterior probability of the current sample is higher, the sample is accepted. If not, the sample can either be accepted or rejected according to the comparison of the ratio of the posterior probabilities of the newly sampled and previously sampled points to a randomly sampled point from a uniform distribution between 0 and 1. If the ratio is larger, the new point is accepted. Otherwise the process continues with the previously sampled point.}\]
Chapter 4. Statistical methods

posterior, by first sampling from an approximate distribution and then refining this sampling to obtain the target posterior distribution (Gelman et al., 2014). This method is most suitable for parameter estimations with unimodal posteriors and when there is no need to calculate the marginal likelihood explicitly.

For the task of parameter estimation, we choose emcee which is an MCMC Ensemble sampler. Three adjustable parameters in the emcee sampler are adjusted for optimal inferences. These are the number of walkers which give the number of different sequences that do the random walk as an ensemble, the burn-in period which is the first $N$ steps that are discarded to eliminate any bias towards different starting points and the sample size which is the number of draws. In our analysis, we have used the values 50, 500, 250 respectively which were optimal for our particular problem. These numbers are chosen to assure that the posterior probability density distributions have enough samples and that the MCMC chain properly converges.

The likelihood distributions for GRBs are Poisson distributed total counts with Gaussian errors. The priors for Paper 2 are selected as below,

$$
\begin{align*}
E_{\text{CPL}} &\sim \mathcal{U}(1, 10000) \text{ keV} \\
K_{\text{CPL}} &\sim \log \mathcal{U}(10^{-11}, 10) \text{ cm}^{-2}\text{s}^{-1}\text{keV}^{-1} \\
\alpha &\sim \mathcal{U}(-3, 2)
\end{align*}
$$

(4.3)

The priors are selected such that they would enclose all possible CPL model parameter fit results in the GBM catalogue.

The convergence of the MCMC fits are checked for by examining corner plots such as in Figure 4.2 (histograms of the posterior probability density distributions), trace plots (plots of step number versus the sampled point) and the amount of autocorrelation or the autocorrelation time.

To verify the viability of the resulting fits, posterior predictive checks (PPC) are applied. The posterior predictive distribution is given as,

$$
P(D_{\text{rep}} \mid M, D) = \int P(D_{\text{rep}} \mid M, \theta) P(\theta \mid M, D) d\theta
$$

(4.4)

where $D_{\text{rep}}$ is the replicated data from the likelihood under the observed posterior distribution $P(\theta \mid D, M)$. $D_{\text{rep}}$ should be coming from the same generative process as $D$ if the fit is to be considered viable. This is done by selecting different parameter sets from the posterior to build up a model spectrum which is used to estimate source and background counts.

Once replicated counts are obtained, they are compared to the original source plus background counts. This is done by plotting the observed cumulative counts against replicated cumulative counts, in the form of a QQ plot. A one to one relation indicates a perfect fit, while a deviation from 68 per cent quantile for a considerable part of the count plot indicates that the model cannot represent the data accurately (Acuner, Ryde, and Yu, 2019; Burgess et al., 2019). Figure 4.3 exemplifies two PPC QQ plots of NDP and SCS fits to a NaI spectrum of GRB100707.

4.3. Bayesian Data Analysis

There are two aspects to BDA. One is the estimation of parameters as explained in Section 4.3.1. A sampler like MCMC can constrain the spread of our estimated parameter values such that we can have an a-posteriori distribution of each parameter that we can use for our physical interpretation of our model given our data. The other is model selection or model comparison. In Section 4.3.1 I mentioned that a model comparison requires the calculation of the Bayesian evidence, $P(D | M)$. Also called the marginal likelihood, this term doesn’t have any dependence on the estimated parameters. As such, the term can be written as the normalization of the posterior over $\theta$,

$$Z = \int d\theta P(D | \theta, M) P(\theta | M). \quad (4.5)$$

The value $Z$ can be calculated for two competing models. Let $M_1$ and $M_2$ be the two models to be compared, with evidences $Z_1$ and $Z_2$ respectively. Then ratio of $Z_1$ and $Z_2$ gives the so called Bayes factor,

$$B_{21} = \frac{Z_2}{Z_1} = \frac{P(M_2 | D)}{P(M_1 | D)} \quad (4.6)$$
which is basically the ratio of the probabilities of the two models given data. The above compact form can be written as the multiplication of the priors and the likelihoods,

$$B_{21} = \frac{P(M_2)}{P(M_1)} \cdot \frac{P(D | M_2)}{P(D | M_1)}. \quad (4.7)$$

If we do not have any reason to assign a significantly higher prior probability to any of the two models, then the prior ratio is simply unity. One can always introduce greater belief in any of the models by simply adjusting this term, which is one of the conveniences of BDA. When $\frac{P(M_2)}{P(M_1)} \approx 1$, we are left with,

$$B_{21} = \frac{P(D | M_2)}{P(D | M_1)}. \quad (4.8)$$

I denote Equation (4.8) as $K$ ($K=B_{21}$) for simplicity. $K$ values are traditionally presented as natural logarithms, $\ln K$. This ratio gives an automatic implementation of the Occam’s razor[5] and penalizes fine tuning. The evidence is larger for a model if more of its parameter space is likely. On the contrary, it is smaller for a model which has larger areas of its parameter space with low likelihood values, even if its likelihood is sharply peaked at some region.

The notion of Bayesian evidence was first introduced in Jeffreys [2003] and later introduced in an astrophysical context in Jeffreys [1957]. Jeffreys [2003] define a general scheme of interpreting the $K$ or $\ln K$ values. The so-called Jeffreys’ scale describes four different regions of certainty (in the belief that one model is better preferred than the other by the data):

1. 1 to 2 - not worth more than a mention
2. 2 to 10 - substantial
3. 10 to 100 - strong

[5] Everything being equal, the simplest model is chosen.
4.3. Bayesian Data Analysis

Kass and Raftery (1995) further refine this and present below criteria with $2\ln K$ values instead of $K$,

1. 0 to 2 - not worth more than a mention
2. 2 to 6 - positive
3. 6 to 10 - strong
4. $>10$ - very strong

It should be noted that these are only very rough guidelines and can be chosen arbitrarily for different applications. A better way to represent the degree of model preference would be to compute the probability that a certain model is preferred by the data over the other. Posterior model probability is given as,

$$
P(M_2 \mid D) = \frac{P(D \mid M_2)}{P(D \mid M_2) + P(D \mid M_1)} \quad (4.9)
$$

$$
P(M_2 \mid D) = \frac{1}{1 + P(D \mid M_1)/P(D \mid M_2)} \quad (4.10)
$$

$$
= \frac{1}{1 + B_{12}} \quad (4.11)
$$

$$
= \frac{1}{1 + e^{\ln(B_{12})}} \quad (4.12)
$$

$$
= \frac{1}{1 + e^{-\ln(B_{21})}} \quad (4.13)
$$

with $B_{21} = K$, we have,

$$
P = \frac{1}{1 + e^{-\ln K}}. \quad (4.14)
$$

In the case of equal priors, the posterior model probability is a sigmoid function.

The calculation of Equation 4.5 requires a high dimensional integration over the whole parameter space. MCMC sampling methods based on the Metropolis-Hastings algorithm do not calculate this integral and hence require extra computations to estimate the evidence which are at least an order of magnitude more expensive than the parameter estimation itself. There exists another class of algorithms called the Nested Sampling (NS) algorithms that are built to estimate the Bayesian evidence itself and do the parameter estimation as a by product. NS is able to estimate the multi dimensional integral in Equation 4.5 by transforming it to a one dimensional integral. This is done by re-writing the integral as a function of the likelihood that is dependent on the prior volume. The evidence then can be approximated using weighted sums of this likelihood. The steps are as follows,

1. At $i = 0$, $N$ live points are drawn from the full prior.
2. At each iteration, the sample with the lowest likelihood is removed from the live point set and is replaced by a point drawn from the prior.

3. Corresponding prior volume contained in this iso-likelihood contour is a random variable given as a linear function of the prior mass.

4. Each iteration shrinks the prior volume by a certain amount which results in nested shells of likelihood.

The main challenge with this method is to draw unbiased samples efficiently from the likelihood constrained prior. We use MultiNest as our NS algorithm, which overcomes this issue by making use of an “ellipsoidal rejection sampling scheme” under the condition that the prior is uniform or can be made uniform under certain transformations. This algorithm works in two main steps:

1. At each iteration, the full set of an ensemble of live points are enclosed within a set of ellipsoids.

2. Expectation-maximization is used to minimise the sum of the volumes of the ellipsoids with a constraint on the total volume.

In MultiNest, at each iteration, new points are drawn from the intersection of the ellipsoids. If the new point is selected, its probability is simply the inverse of the number of ellipsoids it lies in.

The use of ellipsoidal decomposition allows for substantial flexibility in the geometry of the posterior exploration which can deal with elongated, curving degeneracies and the distinct modes can be independently evolved (Feroz, Hobson, and Bridges, 2009).

Since the evidences are dependent on the ratio of the priors, we take care to choose all prior ranges comparably. Figure 4.4 shows the distribution of the parameter estimations obtained from these priors for NDP and SCS models. We test for different ranges of priors to test for the effect of the priors to find the optimal range which encompasses all the true parameter values but not unnecessarily too wide, since this becomes gradually very computationally expensive.

Priors are selected via an iterative process. This is required since there are no catalogue values for NDP and SCS models to reference our priors on. Paper 3 is based on the sample from Yu, Dereli-Bégué, and Ryde (2019) that contains 37 bursts. For all of these bursts we first use maximum likelihood estimation (MLE) to find out point estimates for each parameter in each model. By pooling all MLE results we have our initial crude prior distributions for each parameter. Then we fit the same time bins with the MCMC sampler emcee by providing the MLE based parameter distributions as priors. Emcee posteriors give proper prior distributions that were standardized for each parameter of the each model such that only three sets of priors would be used in the final MultiNest runs. This is important to carry out a proper model comparison analysis.
4.3. Bayesian Data Analysis

We choose the priors in Paper 3 to be as follows, for three different (two physical and one empirical) models,

Non-dissipative photosphere:

\[
\begin{align*}
E_{\text{NDP}} &\sim U(0.1, 500) \text{ keV} \\
K_{\text{NDP}} &\sim \log U(10^{-2}, 2 \times 10^2) \text{ cm}^{-2} \text{s}^{-1} \text{keV}^{-1}
\end{align*}
\] (4.15)

Slow cooled synchrotron:

\[
\begin{align*}
E_{\text{SCS}} &\sim U(1, 2000) \text{ keV} \\
K_{\text{SCS}} &\sim \log U(10^{-2}, 10^2) \text{ cm}^{-2} \text{s}^{-1} \text{keV}^{-1}
\end{align*}
\] (4.16)

Cut-off powerlaw:

\[
\begin{align*}
E_{\text{CPL}} &\sim U(10, 1500) \text{ keV} \\
K_{\text{CPL}} &\sim \log U(10^{-2}, 5 \times 10^2) \text{ cm}^{-2} \text{s}^{-1} \text{keV}^{-1} \\
\alpha &\sim U(-2, 1.5)
\end{align*}
\] (4.17)

Number of live points for the MultiNest runs are selected to be 600 which gives good sampling in the given prior space with manageable run times.

Figure 4.5 depicts the fitted energy spectra obtained with this process for two different bursts that favor either a NDP or a SCS model.
4.3.3 Information Criteria

Besides the evidence, three information criteria are presented throughout the papers in this thesis. As with the Bayesian evidence, the purpose of these criteria are to make comparisons of different models as objectively as possible. Predominantly in Paper 3, we make use of Bayesian information criterion (BIC) (Schwarz, 1978), Akaike information criterion (AIC) (Akaike, 1974) and the deviance information criterion (DIC) (Spiegelhalter et al., 2002). We explain them briefly below and discuss how they compare to the Bayesian evidence. These criteria are driven from the Information Theory, hence the name information criteria and are based on the maximum likelihood for a given model. The Bayesian evidence on the other hand, takes into account the likelihood throughout the parameter space.

AIC is a frequentist criterion defined as,

$$AIC = -2 \ln L_{\text{max}} + 2k$$ (4.18)

where $L_{\text{max}} = P(d \mid \theta, M)$ is the maximum likelihood and $k$ is the number of parameters of the model. BIC is a Gaussian approximation of the Bayesian evidence,

$$BIC = -2 \ln L_{\text{max}} + k \ln N$$ (4.19)

where $N$ is the number of data points used in the fit. BIC further assumes that the data points are independent and identically distributed (i.i.d.). DIC is similar to AIC but now $L_{\text{max}}$ is replaced by the Kullbeck-Leibler divergence $D_{KL}$,

$$DIC = -2D_{KL} + 2C_b$$ (4.20)

and $C_b$ is the effective number of parameters which is the number of parameters that are constrained. Kullbeck-Leibler divergence is a measure of relative entropy between two probability distributions.

The Bayesian evidence does not penalize unconstrained parameters whereas AIC, BIC and DIC do. AIC and BIC further penalize extra parameters regardless
of their being constrained or not. However, not penalizing the unconstrained parameters can be useful during model selection, since if a model fits the data good enough with fewer parameters, we can always shrink the model not to include those, instead of discarding the entire model.

Although very simple to calculate and readily available in many fitting procedures, unlike the Bayesian evidences, information criteria require the Gaussianity or near-Gaussianity of the posterior, which is more often than not, is not the case for real-life scientific problems. AIC and BIC further assume that all parameters are well constrained (Trotta, 2008; Liddle, 2007; Mukherjee, Parkinson, and Liddle, 2006).

We now proceed to give some examples regarding the relationship between information criteria and the Bayesian evidence from the results of Paper 3. Figure 4.6 shows the correlations between evidence differences (\(\ln K\)) and information criteria differences between NDP and SCS models. All three plots have a linear relation which means for our sample, when compared to the Bayesian evidence, AIC, BIC and DIC give reasonable estimates for comparing these two models. This result means that the Bayesian evidences are also dominated by the peaked maximum likelihood region and for this sample of GRBs, models and prior intervals, a comparison of information criteria can be used as a diagnostic tool for model comparison when calculation of Bayesian evidences are unfeasible.
Chapter 5

Discussion

5.1 Clusters in GRB catalogs

To give the reader another perspective, we plot the Paper 1 results according to the three main groups that have been described in Section 4.2. This means, instead of the photospheric (clusters 1, 3, 5) and synchrotron (clusters 2, 4) groups that have been shown in the paper, Figure 5.1 show groups based on $T_{90}$ and fluence.

Figure 5.1. Same as Figure 4.1 but for fluence, $E_{pk}$ and $\alpha$ values from the Band function fits respectively.
Since the main variance in the data is due to $T_{90}$ and fluence, the three groups, namely bright-long, intermediate-intermediate and short-faint bursts, can be observed well in each plot. Furthermore, the density plots of $\alpha$ are clearly peaked in different regions of $\alpha$ values. Thermal and non-thermal bursts are still clearly classified in $T_{90}$-fluence space. To further test this proposition, we carry out the analysis for model comparison described in Section 4.3.2 for 5 bursts from the clustering analysis sample. These bursts are selected such that they each have very high probability of belonging to their cluster (one selected for each cluster) and are thought to be representing their clusters’ spectral properties well. Figure 5.2 shows these 5 bursts each labelled according to their cluster number. As expected, all bursts from the clusters that are predicted to be of non-thermal origin falls to the left side of the so called Line of Death (LOD) of synchrotron emission ($\alpha_{LOD} \sim -0.8$ (Preece et al., 1998; Burgess and Ryde, 2015)) with evidences suggesting a SCS model over NDP with certainty. The LOD of -0.8 is from Burgess and Ryde (2015) who fitted the simulated synchrotron spectra with a Band function, finding that the LOD shifts to a more constraining value than the original reported value of -2/3 given in Preece et al. (1998). This happens due to the inability of the Band model to capture the true synchrotron spectrum, biasing the obtained $\alpha$ values. Cluster 5 representative clearly prefers a thermal model, although with an $\alpha$ value softer than the theoretically predicted of value 0.4. Bursts from clusters 1 and 3 are
in the region of uncertainty pertaining to which of the models they prefer better. However, because the SCS cannot explain harder values of $\alpha$ than $\alpha_{LOD}$, between the two models these are assigned to a photospheric origin as well. Results from Paper 2 support this proposition. In Paper 2, we have shown that a NDP type spectrum, when fitted with empirical models such as CPL and Band functions, can have $\alpha$ values as soft as -0.4, peaking around -0.1. Furthermore, photospheric emission can become non-thermalized via subphotospheric dissipation, in which case, the spectrum will be narrower than SCS but broader than a NDP spectrum.

It is interesting to see the relationship between $T_{90}$ and fluence values of bursts and their preferences of thermal or non-thermal shaped spectra. From Figures 5.1 and 5.2, it can be speculated that longer bursts are brighter and more non-thermal whereas shorter bursts are dimmer and more thermal in spectra. If the burst length is associated with the emission radius, it can be argued that the long and non-thermal bursts are of an external shock origin whereas shorter and thermal bursts require a different, likely photospheric origin with smaller emission radii.

Through the findings of Papers 1 and 3, besides classifying the observed bursts, a probability statement can be made in terms of the yet unobserved data. When an unknown spectrum arrives there is a 1/3 ($\sim$ 33 percent) chance that it will be well described by a non-thermal physical process such as SCS. On the other hand, there is a 2/3 (67 percent) probability that the spectrum will be explained by the following three physical scenarios. With softer $\alpha$ values ($\alpha < 0$ down to $\alpha < -0.8$) the burst is in the domain of subphotospheric dissipation. With $\alpha > -0.5$ the data starts to prefer a NDP scenario that occurs in the coasting phase. With very hard $\alpha > 0.6$ values, spectrum will most likely prefer a NDP in the accelerating phase, which is theoretically the narrowest spectrum that can be observed from GRBs.

Combining the findings of this thesis as well as of the $\alpha_{LOD}$ for the synchrotron emission (Preece et al., 1998; Burgess and Ryde, 2015), we can speculate the observation probabilities for non-thermal, SPD, coasting and accelerating NDP scenarios in a way that depends on the catalogue values of $\alpha$ derived from the Band and CPL functions.

Up until the date 29 October 2019, 2284 bursts were observed with full data on $\alpha_{CPL}$ and $\alpha_{BAND}$. For $\alpha_{BAND}$ values, 193 of these have $\alpha > 0.65$ which is consistent with an accelerating NDP, 634 of them have $-0.4 < \alpha < 0.6$ that is consistent of a coasting NDP while at least 724 of them are in the range $-0.8 < \alpha < -0.4$ which can be considered as SPD bursts. This gives the following percentages for the overall GBM catalogue:

- Accelerating NDP : 8 %
- Coasting NDP : 28 %
- SPD : at least 32 %
- Synchrotron/non-thermal : 32 %

SPD percentage gives only a lower limit since in principle this process can have a large range of $\alpha$ values that could occupy some portions of the coasting NDP interval.
as well. Furthermore, SPD can produce $\alpha$ values that are generally attributed to synchrotron emission, i.e. $-0.8 > \alpha$, which can complicate the identification of a particular emission mechanism for bursts falling into this region.

### 5.2 Spectral shape due to subphotospheric dissipation

The inverse Compton scattering of photons coming from the cooling electrons are mainly responsible for the SPD spectral shape, which is broadened due to the high-energy photons created. This takes place when the kinetic energy dissipation takes place below or near the photosphere, where the optical depth is still quite high in the flow. While the high-energy powelaw index is modified by these photons, newly produced low energy photons from the synchrotron emission of heated electrons soften the low-energy powerlaw spectral index. These two effects modify the steepness of the indices both below and above the thermal peak (Vurm, Beloborodov, and Poutanen, 2011; Pe’er, Mézáros, and Rees, 2006; Ahlgren et al., 2015). With SPD, the radiative efficiency of the flow is higher than that of a synchrotron based emission and it naturally explains the steep $\alpha$ values without the efficiency problems that synchrotron emission encounters (Pe’er, Mézáros, and Rees, 2007; Rees and Mézáros, 2005). The predicted energy peaks ($E_{pk}$) in the $\nu F_\nu$ spectra is at the order of $1/(1 + z)$ MeV (Vurm, Beloborodov, and Poutanen, 2011) with a non-thermal looking spectral shape, where $z$ is the measured redshift. Commonly observed low and higher energy indices are around $\alpha \sim -1$ and $\beta \sim -2.5$. These values are for a jet with magnetization, for which $\alpha$ values could be as high as 0.4 if the magnetization is low. These values are compatible with our synchrotron and SPD groups, which can be an argument for simply explaining all non-thermal looking spectra via an SPD model.

### 5.3 Model comparison

The most reliable model comparison is given by the comparison of Bayesian evidences, since it takes into account the whole likelihood region without the need of assumptions, point estimates or approximations. However, practically, we need to make use of the catalogue parameter values for different models like Band and CPL functions for the projects included in this thesis. Below we discuss the goodness of fit measures and parameter values, specifically the low-energy powerlaw index, given by these two models in more detail. It should be noted that it is not a good idea to directly use any parameter values from empirical models to explain physics, as also have been shown in Paper 2.

Yu, Dereli-Bégüé, and Ryde (2019) describe that both in the GBM GRB time-resolved spectra (Yu et al., 2016) and their Bayesian equivalent, they find CPL to be the best fit model according to several different criteria such as Castor C-Statistic (CSTAT) and the DIC, as explained in Section 4.1. In the model comparison that
we have carried out in Paper 3, we utilised the same sample of 37 bursts as that of Yu, Dereli-Bégué, and Ryde (2019) and found that the Band function gives larger evidences in most cases. This might be due to the fact that Bayesian evidences do not punish the models for unconstrained parameters whereas DIC does. Since the quality of high-energy data is low, it is a common occurrence in fits that the high-energy index $\beta$ of the Band function remains badly constrained, taking on spurious values. Besides from the array of different statistical measures used for the model comparison across different catalogues, a main problem for the obtained fit parameters for the GBM catalogue is the automated fits. This pipeline doesn’t always arrive at the maximum likelihood region since the fitting algorithms and the statistics used do not guarantee this in the first fitting trial.

### 5.4 Catalogue parameter values

Figure 5.3 shows the distributions of the $\alpha$ values obtained for Band and CPL functions through the GBM automated fits on the peak flux interval. The $\alpha$ values obtained from CPL fits are, in general, softer than that of the Band $\alpha$ values.

Table 5.1 summarizes the distribution properties of Band and CPL low energy indices. It can be seen that the Band $\alpha$ distribution is wider with heavier tails. CPL $\alpha$ distribution peaks at a softer $\alpha$ value compared to the CPL $\alpha$ distribution. This difference becomes crucial when we attempt to do a similar analysis as done...
Table 5.1. Some statistical properties of the two $\alpha$ distributions. SD stands for the standard deviation and IQR is the inter-quantile range.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Median</th>
<th>Mean</th>
<th>SD</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{BAND}$</td>
<td>-0.6</td>
<td>-0.4</td>
<td>1.08</td>
<td>0.7</td>
</tr>
<tr>
<td>$\alpha_{CPL}$</td>
<td>-0.75</td>
<td>-0.7</td>
<td>0.64</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Figure 5.4. KDE plots of $\alpha$ values from the NDP and SCS bursts that are fitted with the CPL function (from Paper 3). Only the bursts preferring their respective models with a very high probability are plotted ($\sim 100$ per cent, 14 NDP and 10 SCS bursts). Densities are normalized to match the sample sizes.

for the Band $\alpha$ values in the beginning of this section. We see that many more bursts, when judged by the CPL $\alpha$ values, are classified as non-thermal. Hence, the physical interpretations of the $\alpha$ values from the generic catalogues should be done with great caution. Following the discussion on the model comparison, even when sophisticated statistical inference methods and physical models are being used, there can be degeneracies in model space for certain spectra with the physical models available. This may either be remedied with better quality data or the expansion of current physical models to account for the majority of the spectra observed.

Although the mass GRB catalogues might be misleading in terms of physical interpretation of the parameters, the major information about the physical processes responsible for the production of GRB prompt spectra can be obtained from the $\alpha$ values, if the analysis is carried out carefully and with the correct methodology. Figure 5.4 shows the bursts strongly preferring either NDP or SCS model from Paper 3, with $\sim 100$ per cent probabilities. As can be seen, the distributions are clearly separated and are peaking in the physically expected regions, when fitted
with the CPL function. Therefore, empirical functions are giving some indications of what the physical emission processes might be, but the model comparisons should always be carried out with physical models for hypothesis building since the prior for an empirical function is in principle zero. For the two samples plotted, $E_{pk}$ distributions remain identical, which can be explained under a SDP scenario as well, with the model comparison picking out the most extreme cases in the wide $\alpha$ distribution that SDP can provide.

5.5 Synchrotron model fits

Beniamini and Piran (2013) investigated the parameter space of possible fast cooling synchrotron bursts characterised by the ratio of electron energy to that of the energy in magnetic fields as well as the bulk Lorentz factor ($\Gamma$). They find that $\Gamma$ lies between 300 to 3000 with an emission radius of $10^{15}$ to $10^{17}$ cm.

Ravasio et al. (2018) fit GRB160625, one of the brightest bursts observed by Fermi, with a double smoothly broken powerlaw (2SBPL) and a 2SBPL plus a blackbody. They find that a single 2SBPL is a better fit to this burst and that it indicates photon indices which might be generated from a synchrotron emission coming from a non-thermal electron distribution. They further argue that these electrons do not cool fully, leading to a low-energy cut-off well described by the 2BPL model. However, they find that the cooling frequency associated with the $E_{pk}$ found for the fits imply low magnetic fields, which remains a contradiction. They find similar low magnetic fields in a later work (Ravasio et al., 2019). In Oganesyan et al. (2019), they find spectra again well described by synchrotron emission, with the claimed prompt phase as late as 200 seconds.

Burgess et al. (2019) fit a sample of bursts with a physical synchrotron model which allows the for time dependent cooling of the electrons. Among their sample of 19 bursts, $\sim 44\%$ fall into our non-thermal group of bursts from Paper 1. The remaining falls into clusters 1 and 3, which would be falling near the log-evidence $\sim 0$ line in Figure 5.2. From our Paper 3 sample, one common burst with Burgess et al. (2019) is GRB141028 which they describe as well fitted with the synchrotron model whereas the model comparison between the physical photospheric and synchrotron models in Paper 3 shows that the data from this burst strongly prefers a NDP spectral model. In their analysis Burgess et al. (2019) conclude that the empirical model fits such as done with a Band function that give $\alpha$ values from -0.8 up to 1 can actually be explained by a physical synchrotron model and the interpretation of harder $\alpha$ values as an indication of photospheric emission might not be needed for these bursts. Their fits require large emission radii around $10^{20}$ cm and a moderate degree of magnetization which comes with high Lorentz factors. Compared to the $r_{dec}$ calculated in Section 3.4.1 this value for the emission radius is quite large and is not in line with the values calculated from the afterglow measurements.

In Papers 1 to 3, we find that at least 1/3 of the bursts could be of an emission mechanism apart from the photospheric emission, yet we also find evidence for the fact that many bursts between $\alpha$ values of -0.8 to 0 need other explanations besides
synchrotron emission as revealed by Paper 3 results which was the only work so far that compared a thermal and a non-thermal physical model by Bayesian evidences.

One caveat Paper 3 has is that it accounted only for a single instance of synchrotron emission, namely the slow cooled synchrotron which has an electron spectral index $p$ of 3.5 at the time of electron injection. The $p$ value is theoretically established to be 2.24 by relativistic shock predictions but this value has not been observationally recovered. The value of 3.5 was referenced from Burgess et al. (2019) as their sample of 18 bursts strongly preferred this higher $p$ value for their best fit parameters. This dichotomy could be overcome by interpreting a larger fraction of bursts with soft $\alpha$ values, both from our non-thermal and NDP groups as resulting from SPD.

The dimensionless entropy ($\eta$) as mentioned in Section 3.1 can be directly related to the emission mechanism. Bégué and Burgess (2016) discusses that internal and external shocks as well as the thermal emission from the photosphere can be responsible for the observed emission from GRBs. For large values of $\eta$ the photospheric emission is expected to dominate. For moderate and low values, external and internal shocks are expected to take over, respectively. According to the evolution of the parameter $\eta$, different emission mechanisms can take place in a single GRB at different times.

5.6 Polarization

Apart from the broadness of the spectral shapes, linear photon polarization (consisting of the measurements of the degree and the angle of polarization) predictions are different for thermal and non-thermal GRBs. A very narrow and thermal spectrum is expected to have unpolarized emission while broader, non-thermal looking spectra are expected to be highly polarized. Spectra with Comptonized photons, like those that can be produced by SPD, are predicted to have low to moderate polarizations. Therefore, polarization measurements can help to distinguish between SPD and synchrotron models, especially for the bursts in which the comparison between the two models remain inconclusive (Lundman, 2013). However, there are still large measurement errors on the few polarization measurements performed on GRBs which hinders a precise statement about the state of polarization from observed events.

5.7 Time variability

There are many methods to decompose GRB lightcurves for the purpose of extracting the length of the minimum time variability scales. Golkhou and Butler (2014) and Golkhou, Butler, and Littlejohns (2015) have used a structural function analyses, autocorrelation function analyses and Fourier power spectral density analyses. See Golkhou and Butler (2014) and the references therein for more details on the use of these methods in GRBs.
5.8. The general view

analysis method based on the Haar wavelets\footnote{Wavelets are mathematical functions that are used for compact descriptions of narrow time features and the variations in local averages. In GRB lightcurves, they help to keep track of the signal power over noise versus timescale.} to measure the uncorrelated minimum variability time scale in *Swift* bursts. Their method first identifies the smooth temporal features and then registers the longer, temporally unsmooth features as the true variability timescales, which avoids a direct comparison of any variability against the noise floor measurements. They find a median minimum time-scale for long bursts around 0.5 s and a minimum time variability scale of the order of 10 ms in the source frame. In a later work, they analyse *Fermi* bursts this time, finding similar values of 45 and 10 ms respectively, for long and short burst median variability timescales. In the observer frame, these values are 134 ms for long and 18 ms for short bursts.

Considering an emission radius of \( R \sim 2\Gamma^2 c \Delta t_{\text{min}} \), long bursts have a typical emission radii of \( R \sim 1.3 \times 10^{14} \) cm while short bursts have \( R \sim 3.3 \times 10^{13} \) cm for these shortest variability time scales. These approximations are all lower limits derived from the most rapid time variabilities observed, corresponding to about 10 per cent of the bursts. Therefore, they show how time variability measurements can limit the size of the progenitor.

5.8 The general view

The main results arrived at in this thesis are summarized in Figure 5.5. The grey and turquoise curves represent the photospheric and synchrotron groups as specified in Paper 1. Magenta and purple curves are the simulated and fitted \( \alpha \) values from the coasting and accelerating NDP models respectively. Some known theoretical limits are also shown. The blue and light blue dashed lines respresent the slow and fast cooling synchrotron \( \alpha \) values whereas the purple dashed lines are the limits for a coasting phase non-dissipative spectrum. The \( \alpha \sim 0.4 \) represent the theoretically expected slope for a coasting NDP while the \( \alpha \sim -1 \) is the limit up to where a dissipative coasting phase photospheric spectrum can soften.

As can be seen, the synchrotron distribution has two peaks near the expected values of SCS and FCS emissions (denoted by the turquoise lines). The photospheric group extends from \( \alpha \approx 1 \) to \( \alpha \) as small as -2. The interval \(-1 < \alpha < 1\) can be explained by non-dissipative photospheres for larger values and dissipative photospheres for the smaller values which can all be produced by SPD. The soft \( \alpha \) values can be attributed to the Comptonization of soft photons from synchrotron emission between the optical depths of 100 up to 1, the photosphere. If there is continous dissipation throughout the jet, it is expected to have \( \alpha \) values around -1 which would become even softer if larger magnetizations are involved (Vurm, Beloborodov, and Poutanen, 2011). Although the SPD model can account for a large range of \( \alpha \) values without the need of an additional non-thermal component, it would still need to account for the naturally occuring peaks at the SCS and FCS values which is currently a matter of research. Furthermore, it remains also
Figure 5.5. KDE plots for the $\alpha$ distributions from the simulated coasting and accelerating NDP bursts that are fitted with the CPL function (magenta and purple curves) from Paper 2, as well as the $\alpha$ distributions for the synchrotron and photosphere groups from Paper 1 (turquoise and silver respectively). Superimposed are the minimum and maximum $\alpha$ values theoretically predicted for the coasting NDP spectrum (purple dashed line), dissipative coasting phase spectrum (black dashed line) as well as the SCS (turquoise dashed lines) and FCS (light turquoise dashed lines) limits.
5.8. The general view

A possibility that more than one emission mechanism is taking place in GRBs at different times, which would ultimately lead to softer $\alpha$ values.

As have been discussed extensively throughout this thesis, the possible groupings presented can be due to different dissipation patterns in the jet such as internal, external shock or magnetic reconnection which depends on the amount of magnetic energy compared to the thermal energy in the flow.

Besides from different emission and dissipation mechanisms, different viewing angles to each GRB observed might create the cumulative differences in the observed distributions, as discussed previously in Section 3.2.

Gathered from the three papers presented in this thesis, 2/3 of the bursts could be of photospheric origin. These bursts include dissipative and non-dissipative photospheres including the possibility of the sub-photospheric dissipation. Almost 1/4 of the bursts have spectra that are narrow enough to be described as non-dissipative photospheres. The remaining 1/3 of the bursts could be explained by a non-thermal process such as the synchrotron emission.
Chapter 6

Outlook

A full understanding of the GRB physics has evaded researchers so far. This is related to a few factors. The high energy photons, especially gamma-rays, are notoriously hard to work with due to the sparsity of the incoming photons as well as the statistical complications, such as the need for the forward folding method to generate spectra to be interpreted. This has created many physical models and ideas with a lack of tools to certainly verify them. As this is the case, researchers oftentimes had to pick their favorite models before having a decent rapport with the data. However, taking sides against Nature has somewhat undermined the quest to find the ultimate explanation as the investments for a certain hypothesis or the other increased in time.

I believe two important tasks are required to be undertaken by any researcher in the field of GRBs. The first would be to use the appropriate statistical methods and be creative about it. The second would be to focus on the big picture rather than trying to prove a limited view of a certain hypothesis. These attempts will be even more fruitful with better data in the future that can differentiate between the already available hypotheses with a good precision.

Better data combined with a Bayesian way of thinking in terms of expanding the models would hopefully create the best environment in which the most useful ideas would flourish to explain the nature of the enigmatic GRBs.

In regard to this thesis, supervised machine learning is the next step to uncover the mysteries of GRBs.

Also, future polarization measurements will be of fundamental importance in breaking some degeneracies between these different models.
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Diversity of the light curves and variability time scales for different BATSE bursts. Figure taken from Fishman and Meegan (1995).</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>The Band function. Superimposed are its index parameters and the energy ranges from GBM and LAT instruments.</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Distributions of Band function for $\alpha$ and $E_{\text{pk}}$ parameters from the Fermi GBM catalog.</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>The Fermi Gamma-ray Telescope before launch. LAT sits on the top inside a reflective covering. Six GBM NaI and one BGO detector can be seen. The right image shows the orientations of the detectors in detail. Figures taken from Meegan et al. (2009).</td>
<td>8</td>
</tr>
<tr>
<td>2.2</td>
<td>GRB100707 lightcurves for one NaI and one BGO detector, labeled with the background selections, the background fit with the polynomial order of 1 and the source selection. The source selection is the orange region and it is the time bin for which the photon spectrum is analyzed.</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>GRB100707032 photon count spectrum with residuals, fitted with a thermal and a non-thermal model, respectively. See Chapter 3 for the details of these models.</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Schematic of the fireball model expressing the relationship of Lorentz factor $\Gamma$ to the radius. $r_s$ is the saturation radius as mentioned in the text. $r_{ph}$, $r_{is}$ and $r_{es}$ are the radii that photospheric, internal and external shock emissions occur respectively. Figure taken from Meszaros (2006).</td>
<td>14</td>
</tr>
<tr>
<td>3.2</td>
<td>The analytical (dashed line) and the expected blackbody emissions from GRBs. Figure taken from Goodman (1986).</td>
<td>16</td>
</tr>
<tr>
<td>3.3</td>
<td>Energy spectra ($F_{\nu}$; arbitrary units) from non-dissipative photospheres (NDP). The blue lines are for a photosphere occurring the coasting phase and the red line is for an acceleration phase photosphere. The black line is for a Planck function. The dashed lines show the approximations given by Eqns. (3.6) and (3.7).</td>
<td>17</td>
</tr>
</tbody>
</table>
List of Figures

3.4 Schematic spectra for the synchrotron cooling, with slow, marginally fast and fast cooling scenarios. Note that $\gamma_{\text{min}} \approx \nu_{\text{min}}$ and $\gamma_{\text{cool}} \approx \nu_{\text{cool}}$. The characteristic photon distribution slopes are given for each case below $\nu_{\text{min}}$. Figure taken from Yu et al. (2015).

3.5 Fermi GBM $\alpha$ distribution superimposed by predicted fast (FCS) and slow (SCS) cooling synchrotron values of -3/2 and -2/3, respectively.

3.6 Schematic of different processes producing a spectrum in the co-moving frame. Thermal component and its comptonized component occur closer to the start of the outflow whereas the synchrotron component comes from farther out. There could be another component at higher energies due to the shock with pair formation as described in the text. Note that an ideal black-body would be narrower than the thermal spectrum shown here. Figure taken from Rees and Mészáros (2005).

4.1 Kernel density estimation (KDE) plots for short, intermediate and long bursts as classified in Paper 1. The curves are normalized to account for the different number of bursts in each group. Note that the long bursts consists of clusters 2 and 4 and the intermediate length bursts consists of clusters 1 and 3. Cluster 5 is represented as short bursts.

4.2 Prior ranges for scale and normalization parameters for NDP and SCS models.

4.3 PPC plots of a good fit (modelled with NDP) and a bad fit (modelled with SCS) on the one NaI spectrum of GRB100707.

4.4 Prior ranges for scale and normalization parameters for NDP and SCS models.

4.5 The energy spectra of GRB 110721 and GRB 110920 fitted with the NDP and SCS models respectively. Shown are the best fitted models for each burst, among the NDP, SCS and CPL models.

4.6 $\ln K$ versus difference in the three different information criteria (AIC, BIC and DIC) for NDP and SCS models, as discussed in Paper 3.

5.1 Same as Figure 4.1 but for fluence, $E_{\text{pk}}$ and $\alpha$ values from the Band function fits respectively.

5.2 Difference in log-evidence, between the fits to the NDP model and the fits to the SCS model, versus the low-energy power-law index, $\alpha$, of the cutoff powerlaw function. Positive values of the log-evidence difference (corresponding to $\ln K$ in Eq. 4.8) indicate a preference for the NDP model. The vertical, dashed, line indicates $\alpha = -0.8$. The burst names and the clusters they belong to are indicated at the individual data points.
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3</td>
<td>KDE plots of CPL and Band function low-energy indices from the time resolved GBM burst spectra observed to date. Note that no cuts were made to $\alpha$ errors in this figure. When $\Delta(\alpha) &lt; 0.5$ are selected, the two distributions become similar (see also Yu, Dereli-Bégué, and Ryde (2019)).</td>
<td>45</td>
</tr>
<tr>
<td>5.4</td>
<td>KDE plots of $\alpha$ values from the NDP and SCS bursts that are fitted with the CPL function (from Paper 3). Only the bursts preferring their respective models with a very high probability are plotted (~100 per cent, 14 NDP and 10 SCS bursts). Densities are normalized to match the sample sizes.</td>
<td>46</td>
</tr>
<tr>
<td>5.5</td>
<td>KDE plots for the $\alpha$ distributions from the simulated coasting and accelerating NDP bursts that are fitted with the CPL function (magenta and purple curves) from Paper 2, as well as the $\alpha$ distributions for the synchrotron and photosphere groups from Paper 1 (turquoise and silver respectively). Superimposed are the minimum and maximum $\alpha$ values theoretically predicted for the coasting NDP spectrum (purple dashed line), dissipative coasting phase spectrum (black dashed line) as well as the SCS (turquoise dashed lines) and FCS (light turquoise dashed lines) limits.</td>
<td>50</td>
</tr>
</tbody>
</table>
Bibliography


Burgess, J. Michael et al. (2019). “Gamma-ray bursts as cool synchrotron sources”. In: Nature Astronomy. DOI: 10.1038/s41550-019-0911-z


Gelman, Andrew et al. (2014). *Bayesian data analysis*. CRC Press.


Lundman, Christoffer (2013). “Photospheric emission from structured, relativistic jets: applications to gamma-ray burst spectra and polarization”. PhD thesis. KTH.


Pe’er, A., F. Ryde, and et al. (July 2007). “A New Method of Determining the Initial Size and Lorentz Factor of Gamma-Ray Burst Fireballs Using a Thermal


