

The use of simulated spectra for machine learning methods with the Fermi GBM catalogue

S. Sajjad^{1,2}, T. Aldowma^{2,3}, A. Khan⁴, S. Razzaque²

1. Theoretical Research Institute, Pakistan Academy of Sciences, Pakistan

2. Centre for Astro-Particle Physics, University of Johannesburg, South Africa

3. Omdurman Islamic University, Khartoum, Sudan

4. Center for Computational & Data Sciences, Independent University, Bangladesh

HEASA 2026 – Cape Town – 29th September to 3rd October

The Fermi Gamma-Ray Burst Monitor (GBM)

The Fermi Gamma-ray Space Telescope has been observing GRBs for the past ~18 years.

- Scintillation
 - **Observes: 8 keV – 40 MeV**
 - Full sky field of View > 8 sr



Well suited to study the prompt emission of GRBs

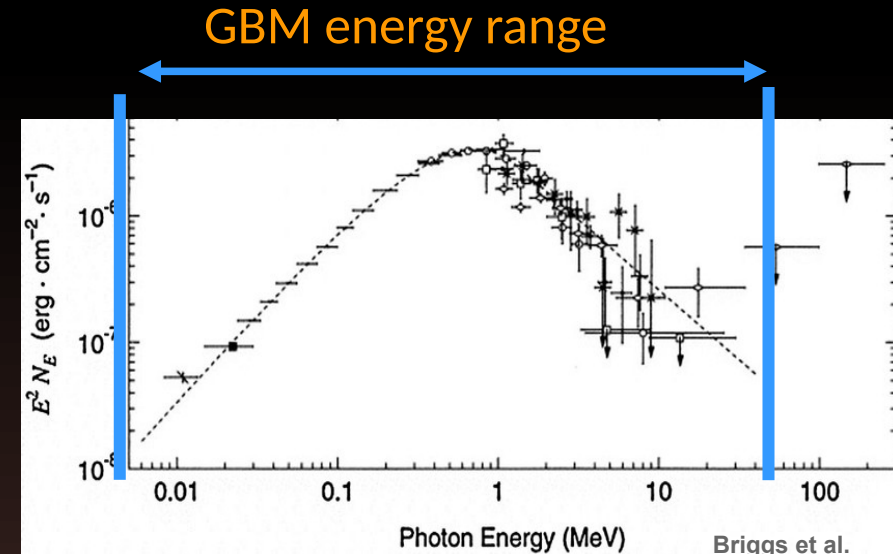
The Fermi GBM Burst Catalogue

The Fermi GBM has been observing GRBs for the past ~18 years.

- Scintillation
- Observes: 8 keV – 40 MeV
- Full sky field of View > 8 sr

Well suited to study the prompt emission of GRBs

Example of GRB prompt emission spectrum



The Fermi GBM Burst Catalogue

Over the course of the past 18 years

• 4236 GRBs → till the 15th of March 2026

• 3590 have spectral data
→ spectra updated till 30th of October 2024

Extensive Fermi GBM Burst Catalogue

→ Contains 306 parameters/columns characterising each of these GRBs

Search in
Xamin or Browse--

FERMIGBRST - Fermi GBM Burst Catalog

HEASARC
Archive

Overview

When referencing results from this online catalog, please cite [von Kienlin, A. et al. 2020](#), [Gruber, D. et al. 2014](#), [von Kienlin, A. et al. 2014](#), and [Bhat, P. et al. 2016](#).

This table lists all of the triggers observed by a subset of the 14 GBM detectors (12 NaI and 2 BGO) which have been classified as gamma-ray bursts (GRBs). Note that there are two Browse catalogs resulting from GBM triggers. All GBM triggers are entered in the [Fermi GBM Trigger Catalog](#), while only those triggers classified as bursts are entered in the Burst Catalog. Thus, a burst will be found in both the Trigger and Burst Catalogs. The Burst Catalog analysis requires human intervention; therefore, GRBs will be entered in the Trigger Catalog before the Burst Catalog. The latency requirements are 1 day for triggers and 3 days for bursts. There are four fewer bursts in the online catalog than in the Gruber et al. 2014 paper. The four missing events (081007224, 091013989, 091022752, and 091208623) have not been classified with certainty as GRBs and are not included in the general GRB catalog. This classification may be revised at a later stage.

The GBM consists of an array of 12 sodium iodide (NaI) detectors which cover the lower end of the energy range up to 1 MeV. The GBM triggers off of the rates in the NaI detectors, with some Terrestrial Gamma-ray Flash (TGF)-specific algorithms using the bismuth germanate (BGO) detectors, sensitive to higher energies, up to 40 MeV. The NaI detectors are placed around the Fermi spacecraft with different orientations to provide the required sensitivity and FOV. The cosine-like angular response of the thin NaI detectors is used to localize burst sources by comparing rates from detectors with different viewing angles. The two BGO detectors are placed on opposite sides of the spacecraft so that all sky positions are visible to at least one BGO detector.

The signals from all 14 GBM detectors are collected by a central Data Processing Unit (DPU). This unit digitizes and time-tags the detectors' pulse height signals, packages the resulting data into several different types for transmission to the ground (via the Fermi spacecraft), and performs various data processing tasks such as autonomous burst triggering.

The GRB science products are transmitted to the FSSC in two types of files. The first file, called the "bcat" file, provides basic burst parameters such as duration, peak flux and fluence, calculated from 8-channel data using a spectral model which has a power-law in energy that falls exponentially above an energy E_{peak}, known as the Comptonized model. The crude 8-channel binning and the simple spectral model allow data fits in batch mode over numerous time bins in an efficient and robust fashion, including intervals with little or no flux, yielding both values for the burst duration, and deconvolved lightcurves for the detectors included in the fit. The bcat file includes two extensions. The first, containing detailed information about energy channels and detectors used in the calculations, is detector-specific, and includes the time history of the deconvolved flux over the time intervals of the burst. The second shows the evolution of the spectral parameters obtained in a joint fit of the included detectors for the model used, usually the Comptonized model described above. The bcat files and their time-varying quantities contained in these two extensions are available at the HEASARC FTP site. Quantities derived from these batch fits are given in the bcat primary header and presented in the Browse table, as described below. The main purpose of the analysis contained in the bcat file is to produce a measure of the duration of the burst after deconvolving the instrument response. The duration quantities are:

- * 't50' - the time taken to accumulate 50% of the burst fluence starting at the 25% fluence level.
- * 't90' - the time taken to accumulate 90% of the burst fluence starting at the 5% fluence level.

Main Search Form

Search of [FERMI](#) Catalog(s)

Tip
Access

Description

Catalog

Data Default Radius (arcmin)

Mission Table Type

Fermi GBM Burst Catalog

fermigbrst.f

180

FERMI Object

1. Enter any constraints on the query below. [\[Help on constraint syntax \]](#)

Examples of query constraints:

▼

(What about [wildcards](#), [spaces](#), and [case sensitivity](#)?)

2. To change the fields that are returned, select the box in the "View" column beside each field desired.

3. To sort the results by any field, select one box in the "Sort" column beside the field to sort on.

View	Sort	Parameter (Units)	Query Terms	Min Value	Max Value	Value Type
☒	☐	name		GRB080714086	GRB260921395	string
☒	☐	ra		00 01 04.8	23 58 57.6	position
☒	☐	dec		-89 00 33	+89 11 01	position
☒	☐	trigger_time		2008-07-14 02:04:12.053 UTC	2026-09-21 09:29:08.423 UTC	date
☒	☐	t50 (s)		0.008	984.084	float
☒	☐	t50_error (s)		0.018	77.118	float
☒	☐	t90_start (s)		-807.424	222.212	float
☒	☐	fluence (erg/cm ²)		5.0105e-09	3.8836e-02	float
☒	☐	fluence_error (erg/cm ²)		3.6450e-09	3.4802e-04	float
☒	☐	flux_1024 (photon/cm ² /s)		-0.5465	8285.3000	float
☒	☐	flux_1024_error (photon/cm ² /s)		0.0869	4460.4100	float
☒	☐	flux_1024_time (s)		-137.664	438.597	float
☒	☐	flux_64 (photon/cm ² /s)		1.2511	33889.7000	float
☒	☐	flux_64_error (photon/cm ² /s)		0.3503	86691.7000	float
☒	☐	flnc_band_area (photon/cm ² /s/keV)		1.141184e-03	1.845692e+06	float
☒	☐	flnc_band_area_pos_err (photon/cm ² /s/keV)		0.000000e+00	7.319149e+05	float
☒	☐	flnc_band_area_neg_err (photon/cm ² /s/keV)		0.000000e+00	7.319149e+05	float
☒	☐	flnc_band_area_pos_err (keV)		1.000000e-02	1.175259e+17	float
☒	☐	flnc_band_area_neg_err (keV)		0.000000e+00	1.175259e+17	float
☒	☐	flnc_band_area_pos_err (keV)		0.000000e+00	1.175259e+17	float
☒	☐	flnc_band_area_neg_err (keV)		-1.977372e+00	2.663325e+01	float
☒	☐	flnc_band_area_pos_err		0.000000e+00	1.788019e+03	float
☒	☐	flnc_band_area_neg_err		0.000000e+00	1.788019e+03	float
☒	☐	flnc_band_area		-3.910338e+01	-3.102754e+00	float

Machine learning approaches with the catalogue

Rich dataset: 3590 GRBs with 306 parameters for each

→ Potentially very interesting for machine learning approaches

→ focusing on the prompt emission

Potential uses include :

- Classifying GRBs
- Clustering
- Looking for indications of emission type
- Regression approaches such as for estimating redshift
- ...

Machine learning approaches with the catalogue

Rich dataset: 3590 GRBs with 306 parameters for each

→ Potentially very interesting for machine learning approaches

→ focusing on the prompt emission

Idea proposed in this work:

The use of simulated/sampled spectra in conjunction with
the GBM catalogue for such approaches

The Fermi GBM Burst Catalogue

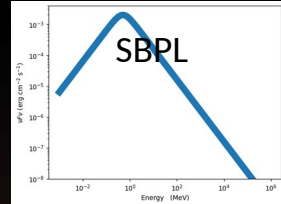
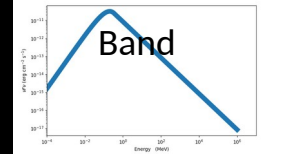
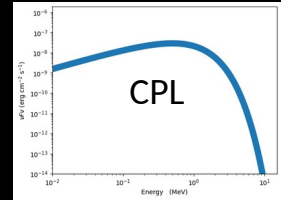
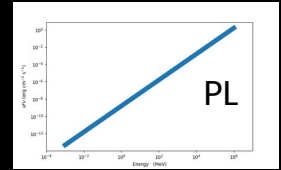
General parameters: 52 → Duration, fluence in various ranges, etc.

The Fermi GBM Burst Catalogue

General parameters: 52 → Duration, fluence in various ranges, etc.

Spectral parameters: 254 → Four spectral models

3590 GRBs, 306 parameters



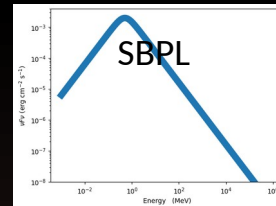
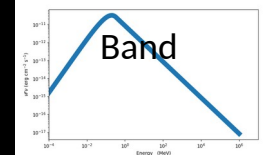
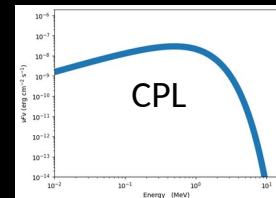
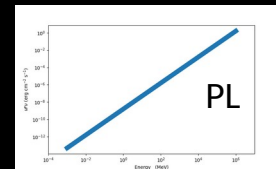
The Fermi GBM Burst Catalogue

General parameters: 52 → Duration, fluence in various ranges, etc.

Spectral parameters: 254 → Four spectral models

→ For each of the four spectral models:

- Fit parameters: 17
- Positive and negative uncertainties : 34



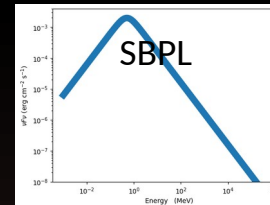
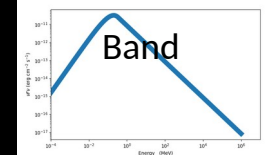
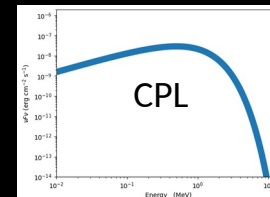
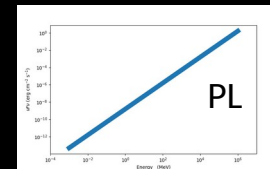
The Fermi GBM Burst Catalogue

General parameters: 52 → Duration, fluence in various ranges, etc.

Spectral parameters: 254 → Four spectral models

→ For each of the four spectral models:

- Fit parameters 17
- Positive and negative uncertainties : 34
- Flux and fluence from the fits : 56
- Other fit parameters: 20
 - Chi square, degrees of freedom, best fit model, time interval, ...



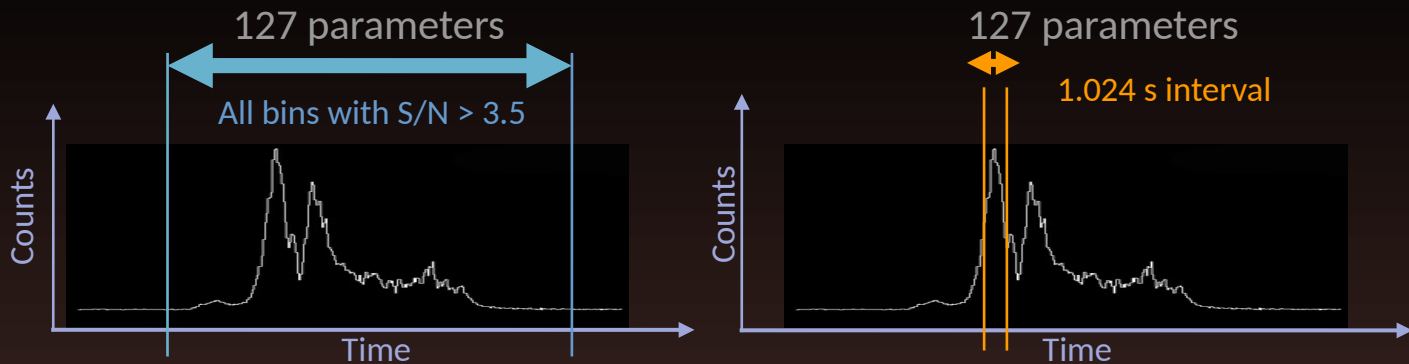
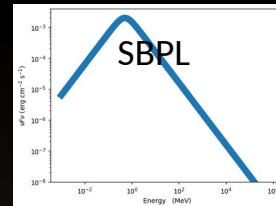
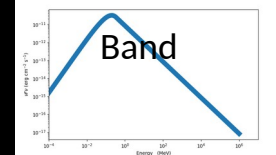
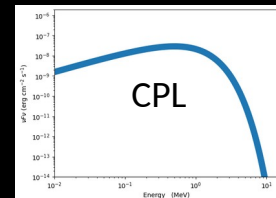
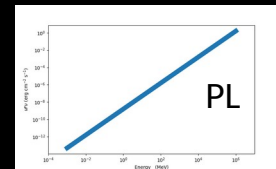
The Fermi GBM Burst Catalogue

General parameters: 52 → Duration, fluence in various ranges, etc.

Spectral parameters: 254 → Four spectral models fitted

→ For each of the four spectral models:

- Fit parameters 17
- Positive and negative uncertainties : 34
- Flux and fluence from the fits : 56
- Other fit parameters: 20
 - Chi square, degrees of freedom, best fit model, time interval, ...



The Fermi GBM Burst Catalogue

General parameters: 52 → Duration, fluence in various ranges, etc.

Spectral parameters: 254 → Four spectral models fitted

→ For each of the four spectral models:

- Fit parameters 17
- Positive and negative uncertainties : 34
- Flux and fluence from the fits : 56
- Other fit parameters: 20
 - Chi square, degrees of freedom, best fit model, time interval, ...

Parameters usually used

The Fermi GBM Burst Catalogue

General parameters: 52 → Duration, fluence in various ranges, etc.

Spectral parameters: 254 → Four spectral models fitted

→ For each of the four spectral models:

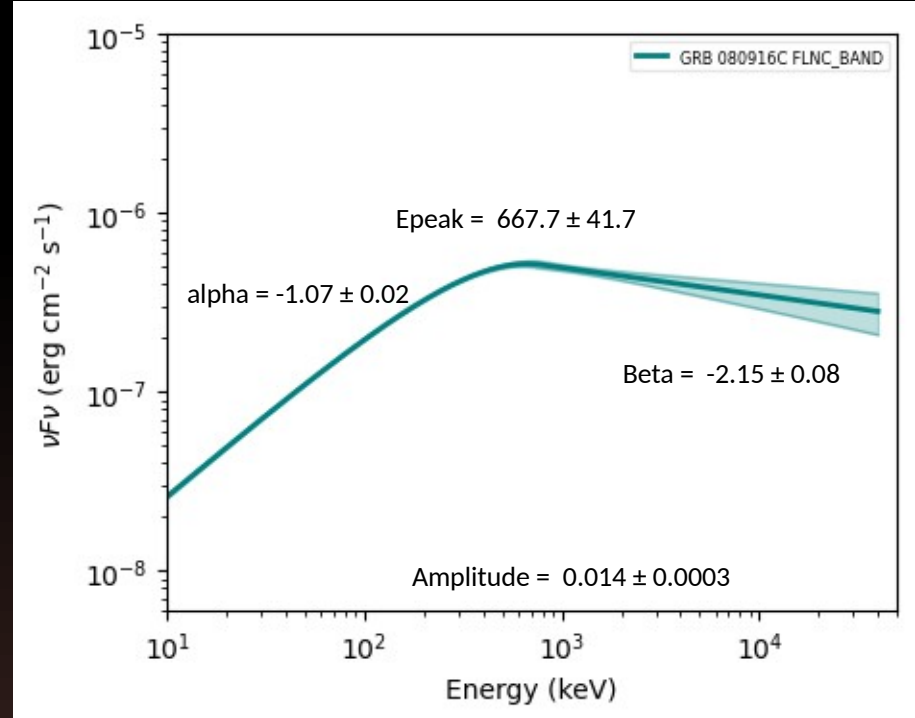
- Fit parameters 17
- Positive and negative uncertainties : 34
- Flux and fluence from the fits : 56
- Other fit parameters: 20
 - Chi square, degrees of freedom, best fit model, time interval, ...

The errors on the fit parameters are not used

The use of error bars in machine learning

Several comments can be made about this

- Not using the uncertainties → All fits/GRBs are treated as equal
- Yet, not all fits are equal
- → Example:
- We can have well-constrained spectral fits



The use of error bars in machine learning

Several comments can be made about this

- Yet, not all fits are equal

But the GBM catalog has poorly constrained fits too.

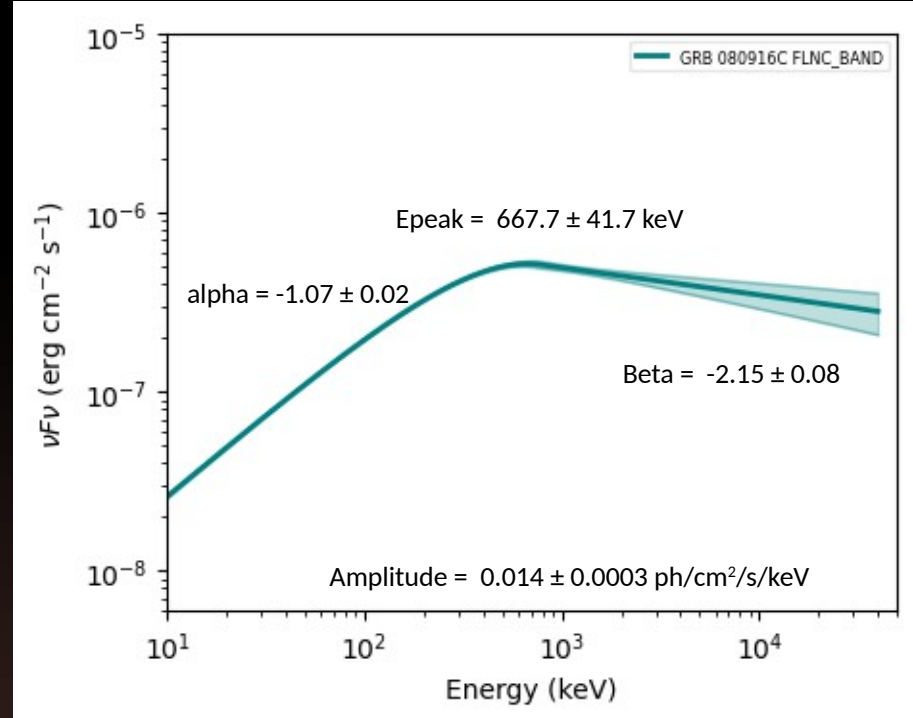
GRB 100112A

→ $E_{\text{peak}} = 1122.8 \pm 2258.4$ keV

→ negative E_{peak} → unphysical

GRB 120403B

→ $\text{Beta} = -3.72 \pm 20.6$



The use of error bars in machine learning

Several comments can be made about this

- Not all fits are equal
- → Some fits also have unphysical values → mentioned in the catalogue notes

GRB 191011A

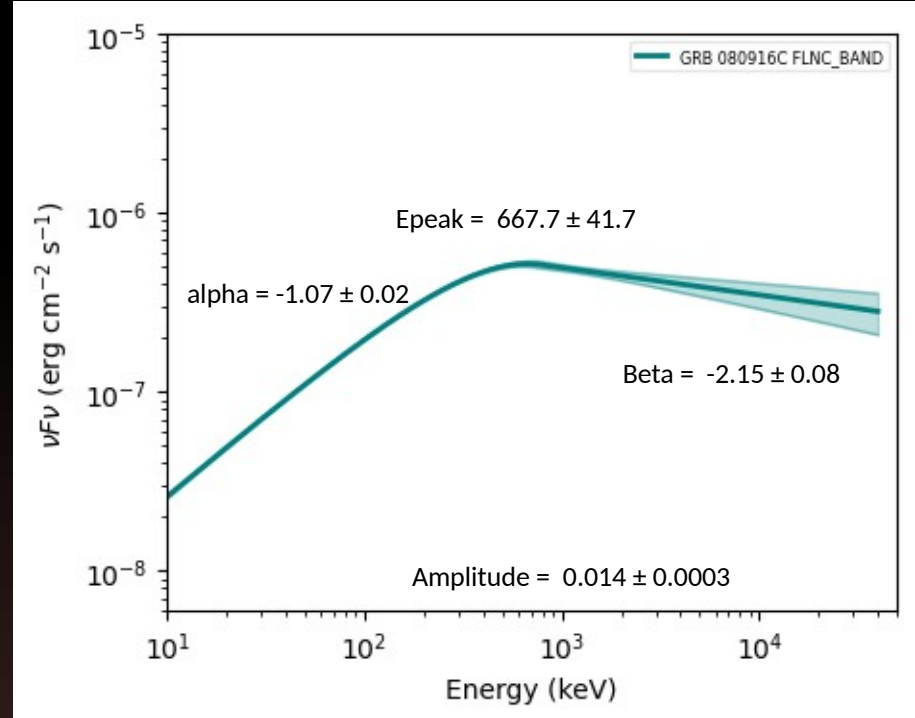
→ Amplitude = $5.72 \times 10^{23} \pm 1.15 \times 10^{23}$ photons/cm²/s/keV

The use of error bars in machine learning

Several comments can be made about this

- Finally, an error bar simply tells us about the compatible range of spectral parameters

→ and therefore also the range of possible underlying physics



The use of error bars in machine learning

Note

- Error bars are often not used in machine learning, in general
 - The emphasis is usually on massive datasets
 - → capture the intrinsic variations as an aggregate

However, in our case 3000-4000 GRBs do not seem to be enough to capture the intrinsic variations

→ This is a relatively small dataset

All this to say, we probably want to use the uncertainties in some way

Machine learning approaches

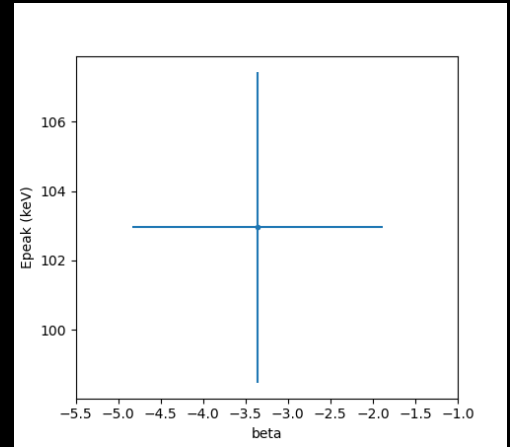
There do exist methods

- Use errors as features (additional parameters)
- Use weighted methods → assign a weight to each point proportional to $1/\sigma^2$
- Adapted machine learning techniques
 - → Bayesian neural networks → Treat weight and inputs as probability distributions
 - → Gaussian processes → Use error bar into prediction uncertainty
 - → Total least square
- Simulate data → in our case, simulate spectra

Simulations

Approach used

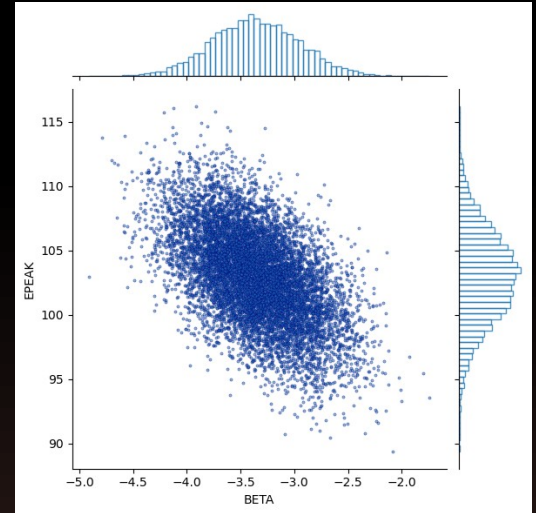
- Download covariance matrix of the fits → Also available through the Fermi website
- Simulate cloud of points 10, 100, 1000, 10000 spectra compatible with it



→ Shown for only 2 parameters here
But done for all parameters of the fit

Additionally

- Reject any spectrum that does not pass cuts
 - Negative amplitude
 - Negative Epeak
 - $\text{Beta} > -2$ (peak calculations not possible)
 - Negative, or infinite flux/fluence
- Keep simulating until the required number: 10, 100, 1000 10000 is reached



Numbers of GRBs left after applying cuts

Example for the Band function

- Apply cuts without simulations:
 - We keep 1566 GRBs out of 3590
- Apply cuts with simulations:
 - 10 Simulations → We keep 3434 GRBs out of 3590
 - 100 Simulations → 3434 GRBs
 - 1000 Simulations → 3434 GRBs

Brief example of application

Principal Component Analysis (PCA)

- PCA can be foundational to many machine learning approaches

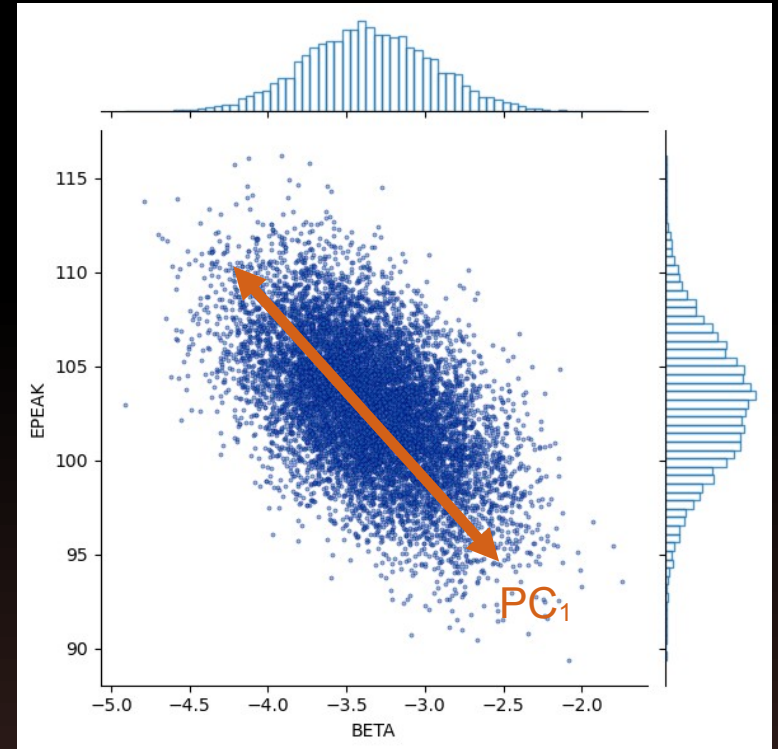
Brief example of application

Principal Component Analysis (PCA)

- PCA can be foundational to many machine learning approaches

Simple 2 parameter problem:

$PC_1 = a_1 * \text{Beta} + a_2 * \text{Epeak} \rightarrow$ Explains most of the variance



Brief example of application

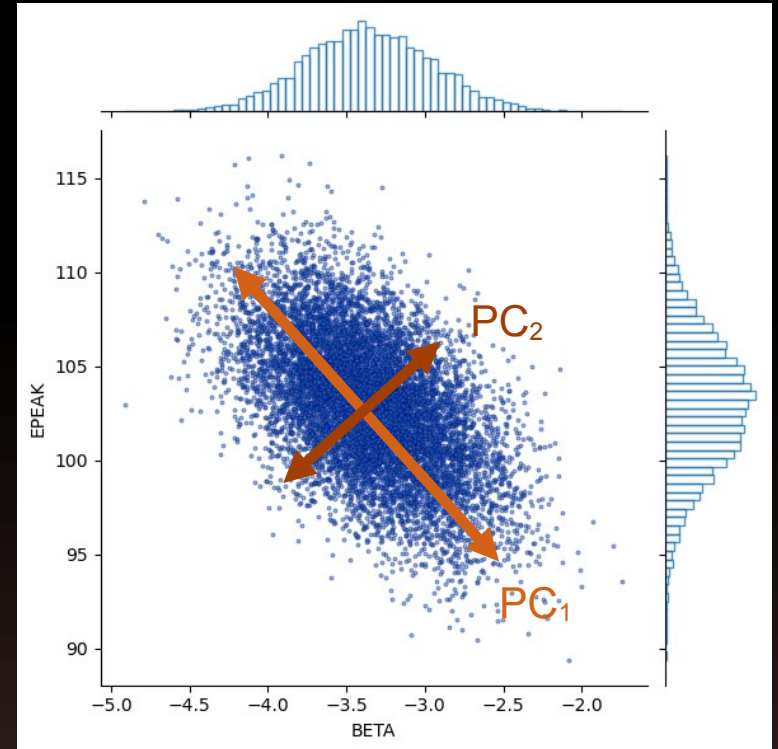
Principal Component Analysis (PCA)

- PCA can be foundational to many machine learning approaches

Simple 2 parameter problem:

$$PC_1 = a_1 * \text{Beta} + a_2 * \text{Epeak}$$

$$PC_2 = a_3 * \text{Beta} + a_4 * \text{Epeak} \rightarrow \text{Explains the rest of the variance}$$



Brief example of application

Principal Component Analysis (PCA)

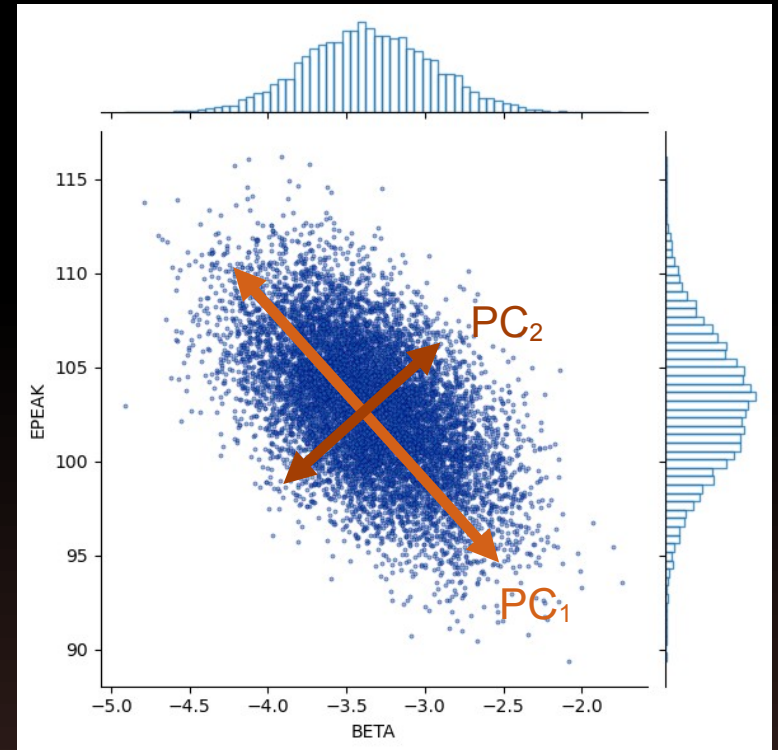
- PCA can be foundational to many machine learning approaches

Simple 2 parameter problem:

$$PC_1 = a_1 * \text{Beta} + a_2 * \text{Epeak}$$

$$PC_2 = a_3 * \text{Beta} + a_4 * \text{Epeak} \rightarrow \text{Explains the rest of the variance}$$

Can be used to identify new variables/components



Brief example of application

Principal Component Analysis (PCA)

- PCA can be foundational to many machine learning approaches

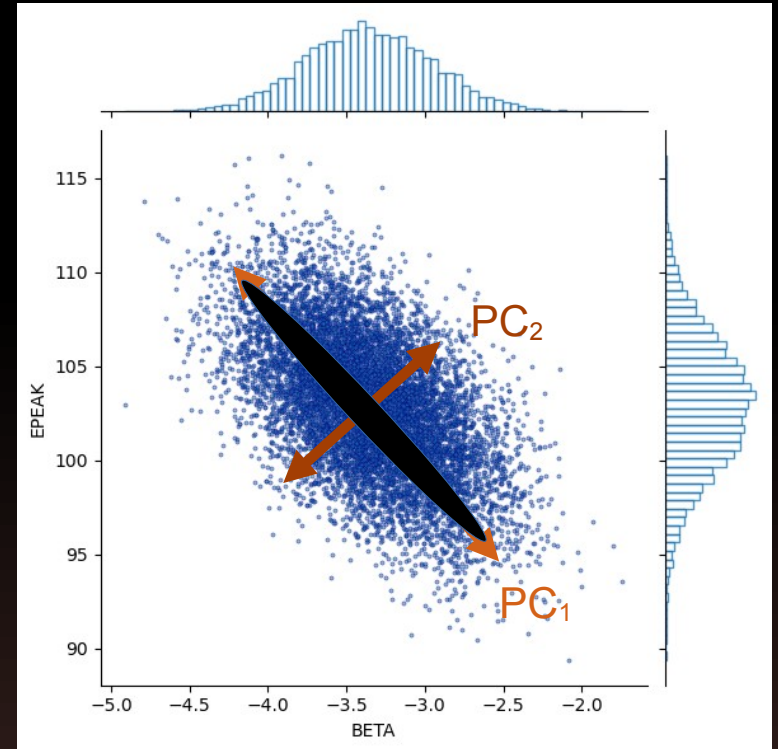
Simple 2 parameter problem:

$$PC_1 = a_1 * \text{Beta} + a_2 * \text{Epeak}$$

$$PC_2 = a_3 * \text{Beta} + a_4 * \text{Epeak} \rightarrow \text{Explains the rest of the variance}$$

Can also be used for reducing dimensions

If the ellipse was flat $\rightarrow$ PC_2 would be negligible



Brief example of application

5 parameters

- 5 Parameters
 - Tint → duration
 - Alpha, Beta, Epeak, Fluence → From Band function spectral fits

We should get 5 principal components:

PC1, PC2, PC3, PC4, PC5

Focus on the PC1

Without simulations: it explains 0.41 of the total variance

With 10, 100, 1000 simulations → ~0.31 of the variance

Brief example of application

PC1

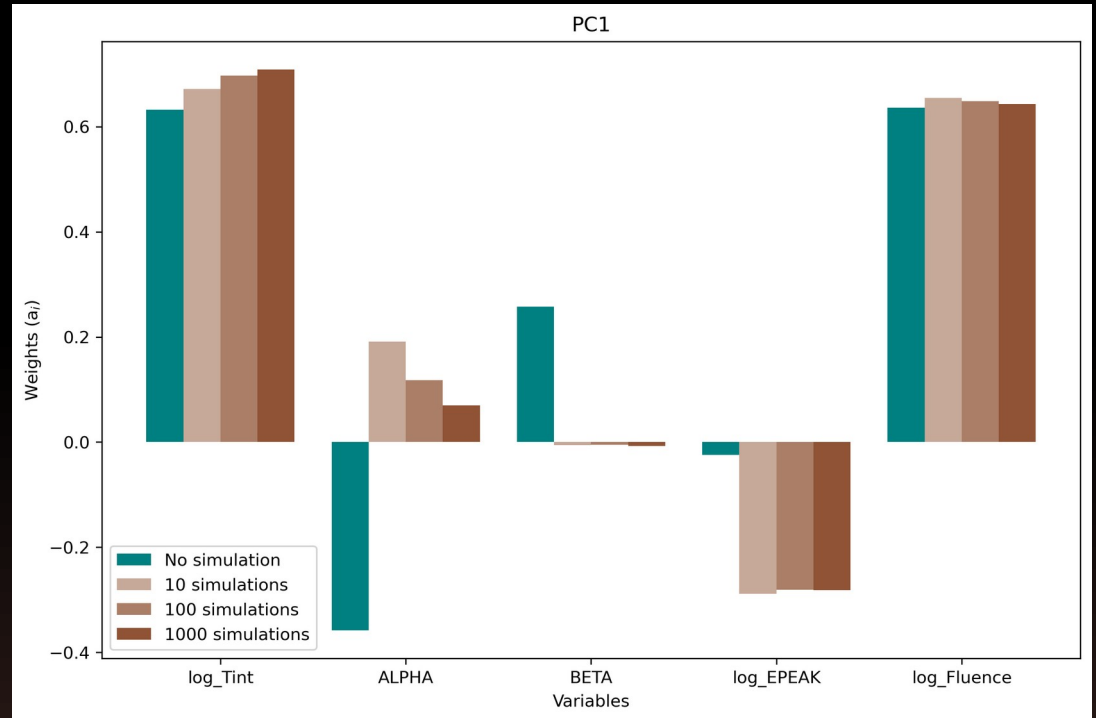
Without simulations

- Tint, alpha, beta, and fluence all contribute
 - alpha contributes negatively
 - Epeak → negligible

With simulations

- Dominated by Tint and fluence
- Beta → negligible
- Alpha → diminishing contribution → change of sign
- Epeak → moderate negative contribution

$$PC_1 = a_1 * \log(Tint) + a_2 * \alpha + a_3 * \beta + a_4 * \log(Epeak) + a_5 * \log(Fluence)$$

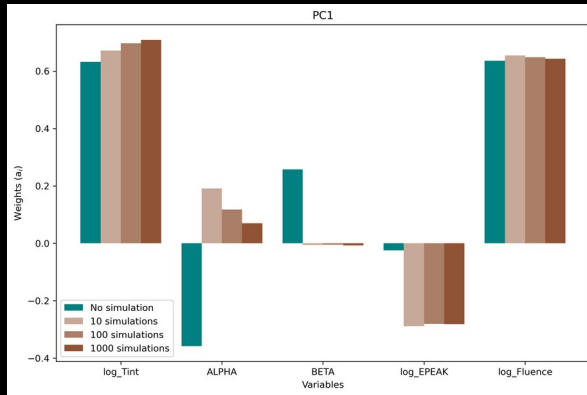


Suggests a duration-fluence component... more apparent with simulations

Brief example of application

Comparison

$$PC_1 = a_1 * \log(\text{Tint}) + a_2 * \alpha + a_3 * \beta + a_4 * \log(\text{Epeak}) + a_5 * \log(\text{Fluence})$$



PC₁ → Fluence and Tint dominated

PC₂ → alpha, Epeak dominated, but also fluence

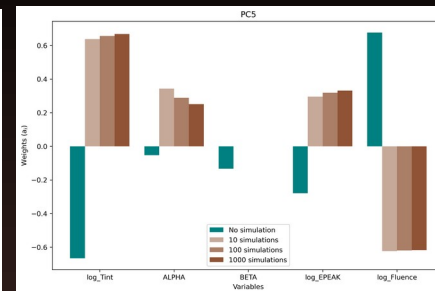
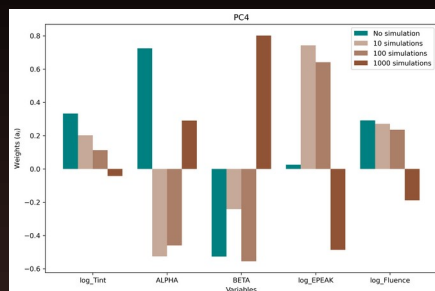
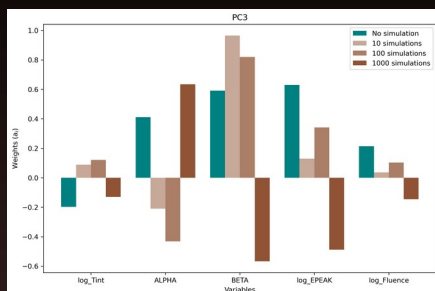
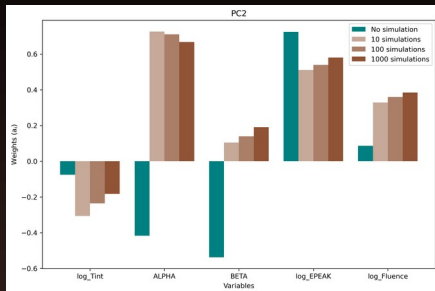
PC₃ → dominated by alpha and beta → flips sign with 1000 sims

PC₄ → Shifts from mixed to beta dominated → flips sign with 1000 sims

PC₅ → Tint and fluence but with opposite signs

Needs to be better understood, especially what happens with 1000 sims

→ Also study effects of further cuts



Thank you

Extra slides

Brief example of application

5 parameters

- 5 Parameters
 - Tint → duration
 - Alpha, Beta, Epeak, Fluence → From Band function spectral fits

We should get 5 principal components:

PC1, PC2, PC3, PC4, PC5

Without simulations

→ they explain 0.41, 0.24, 0.18, 0.14 0.026 of the variable

With 100 simulations

→ 0.31, 0.22, 0.2, 0.2 0.08 → More evenly spread

Numbers of GRBs left after applying cuts

Example: Band FLNC (time-integrated)

- Without simulations:
 - 1566 GRBs out of 3590
- With simulations:
 - 10 Simulations → 3434 GRBs
 - 100 Simulations → 3434 GRBs
 - 1000 Simulations → 3434 GRBs

