Prompt GRB classification through Waterfalls and Deep Learning

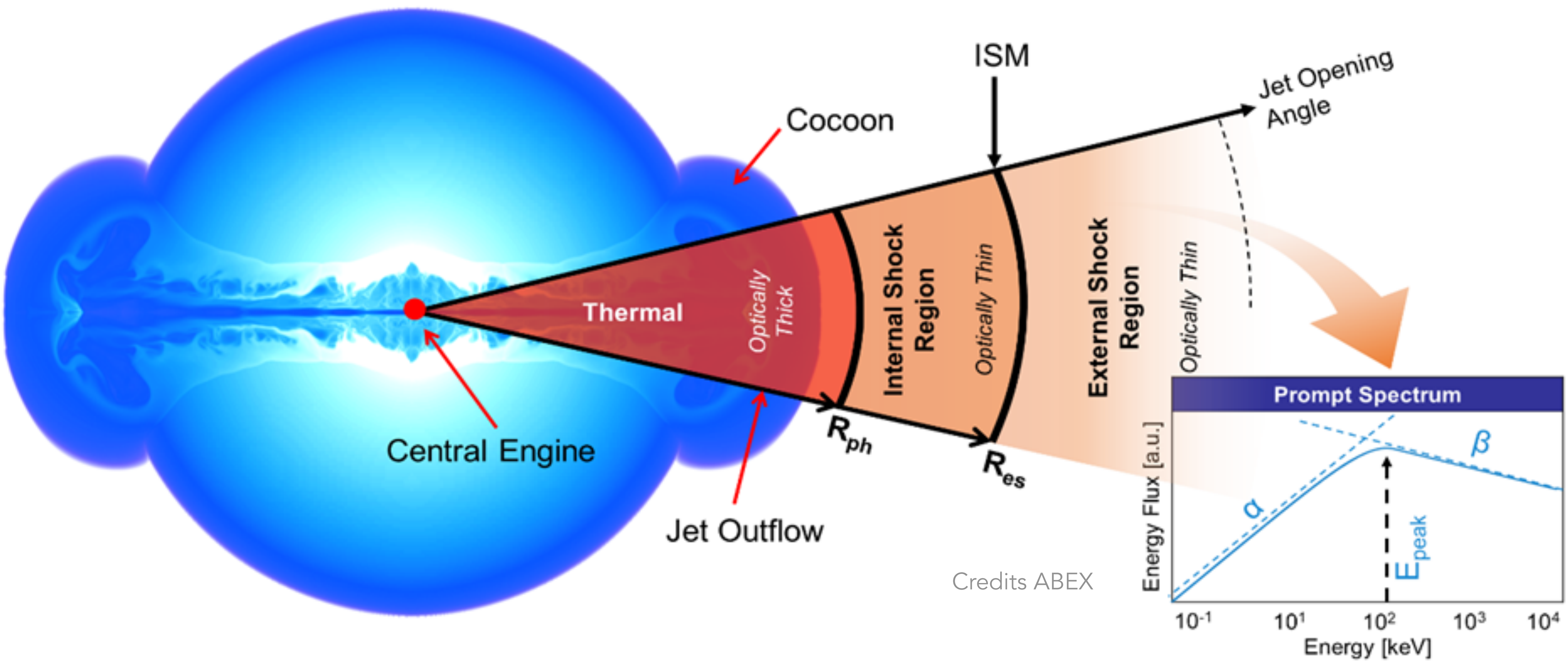
International Conference on Machine Learning for Astrophysics
Valletta, Malta, 01/09/2026

Nicolò Cibrario
cibrario@to.infn.it

On behalf of the team:
M. Negro, E. Burns, S. Strain, J. Wood

Gamma Ray Bursts

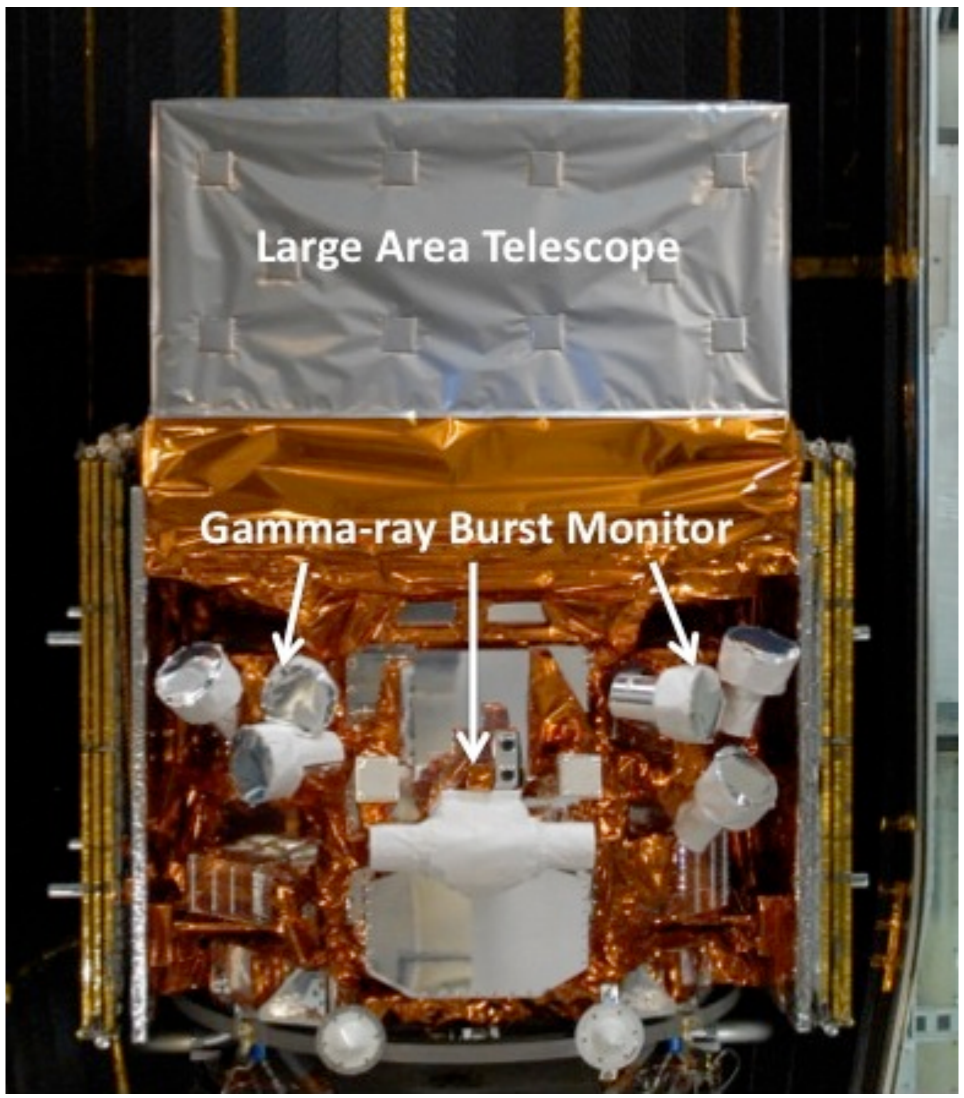
Gamma-Ray Bursts (GRBs) are among the most energetic events known in the universe



Credits ABEX

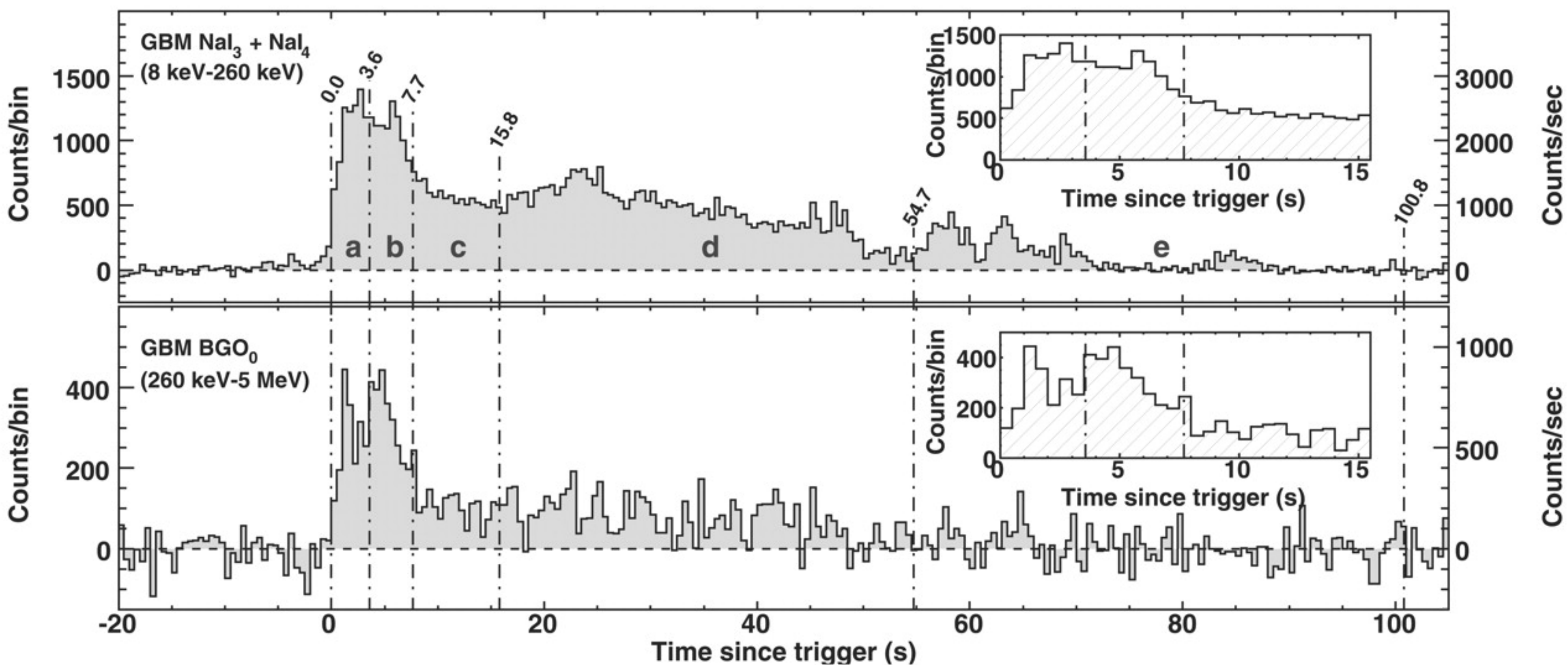
Prompt, multiwavelength, and multimessenger observations allow us to probe physics in extremes which cannot be achieved in terrestrial laboratories

Fermi Gamma-ray Burst Monitor (GBM)

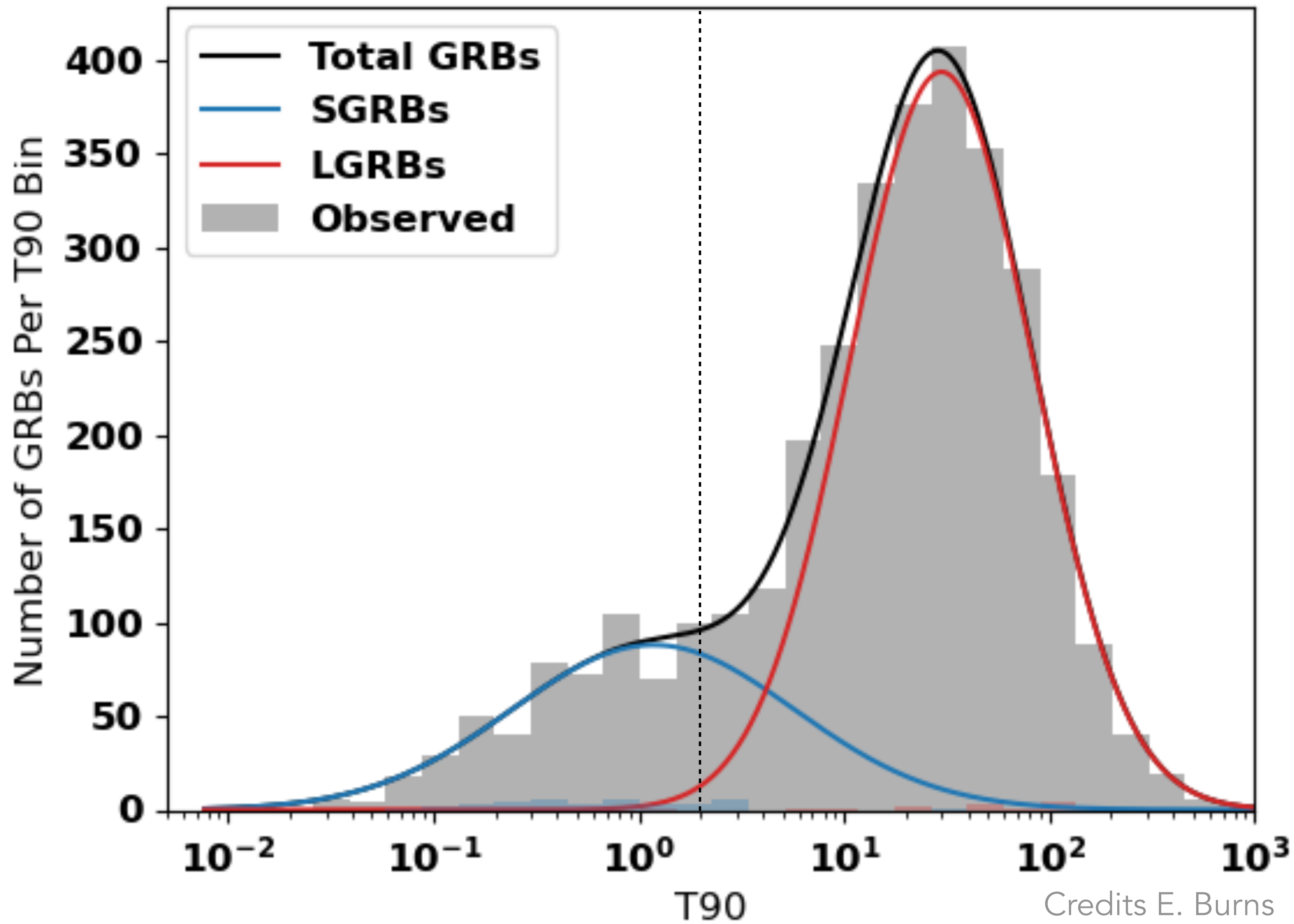


The instrument response is optimal in the energy range between **8 keV and 40 MeV**

GRB 080916C



GRBs progenitors



Binary NS merger

Collapsars

Binary BH-NS merger

Tidal disruptive events

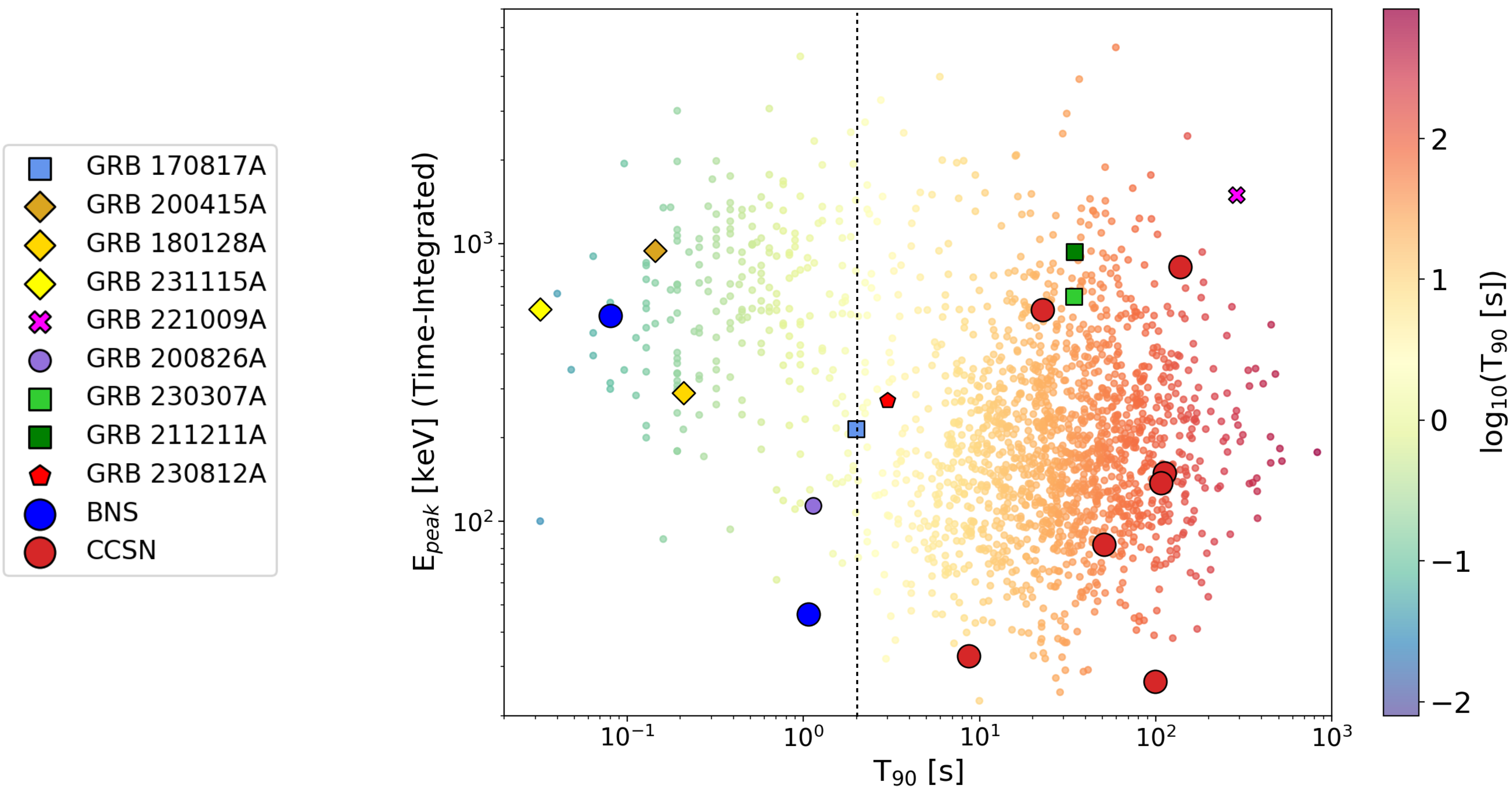
X-ray flashes

Magnetar Giant Flares

Long mergers

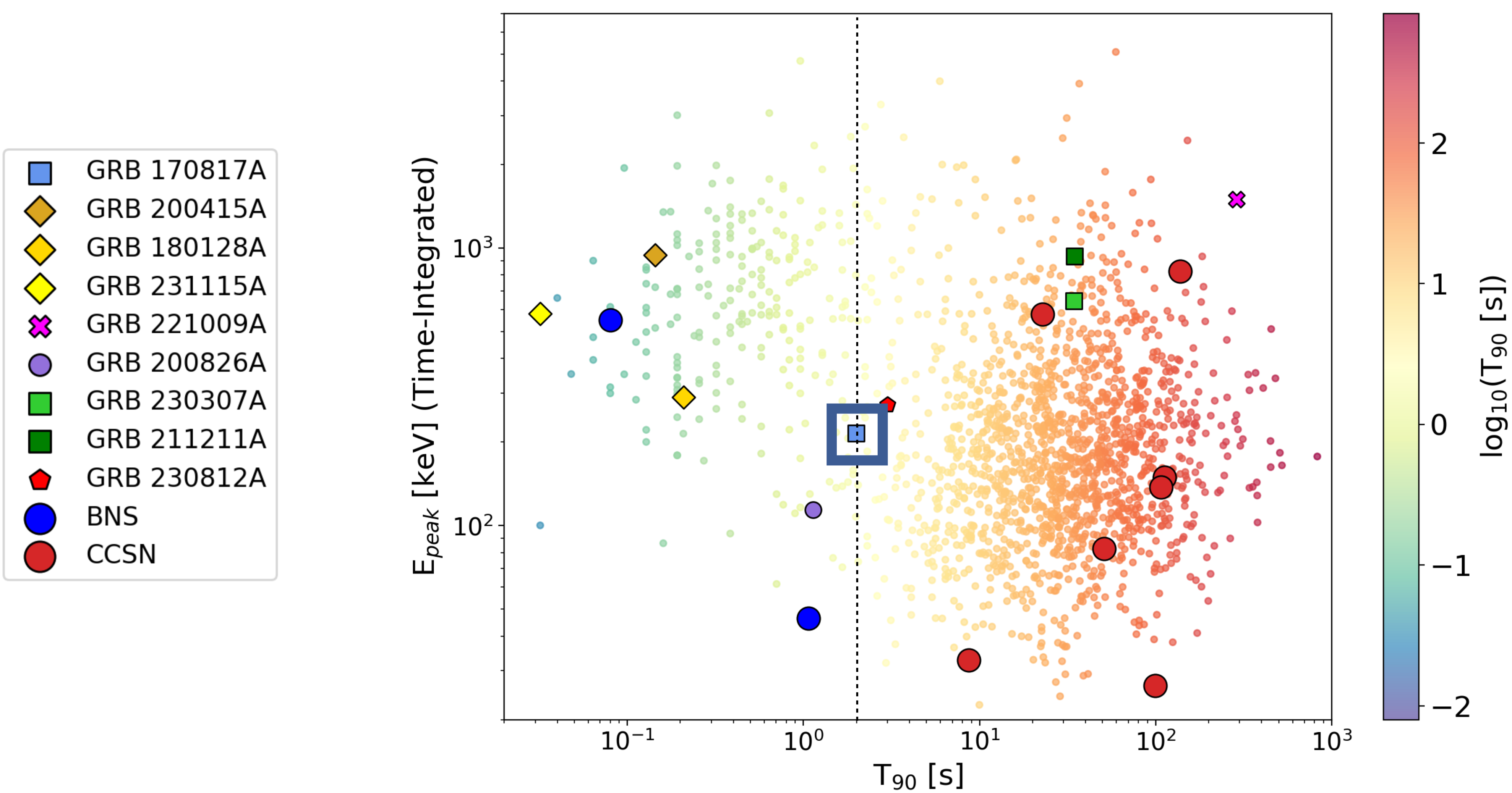
GRBs classification

Categorization of GRBs is not trivial
Recent discoveries challenge the traditional classification



GRBs classification

Categorization of GRBs is not trivial
Recent discoveries challenge the traditional classification

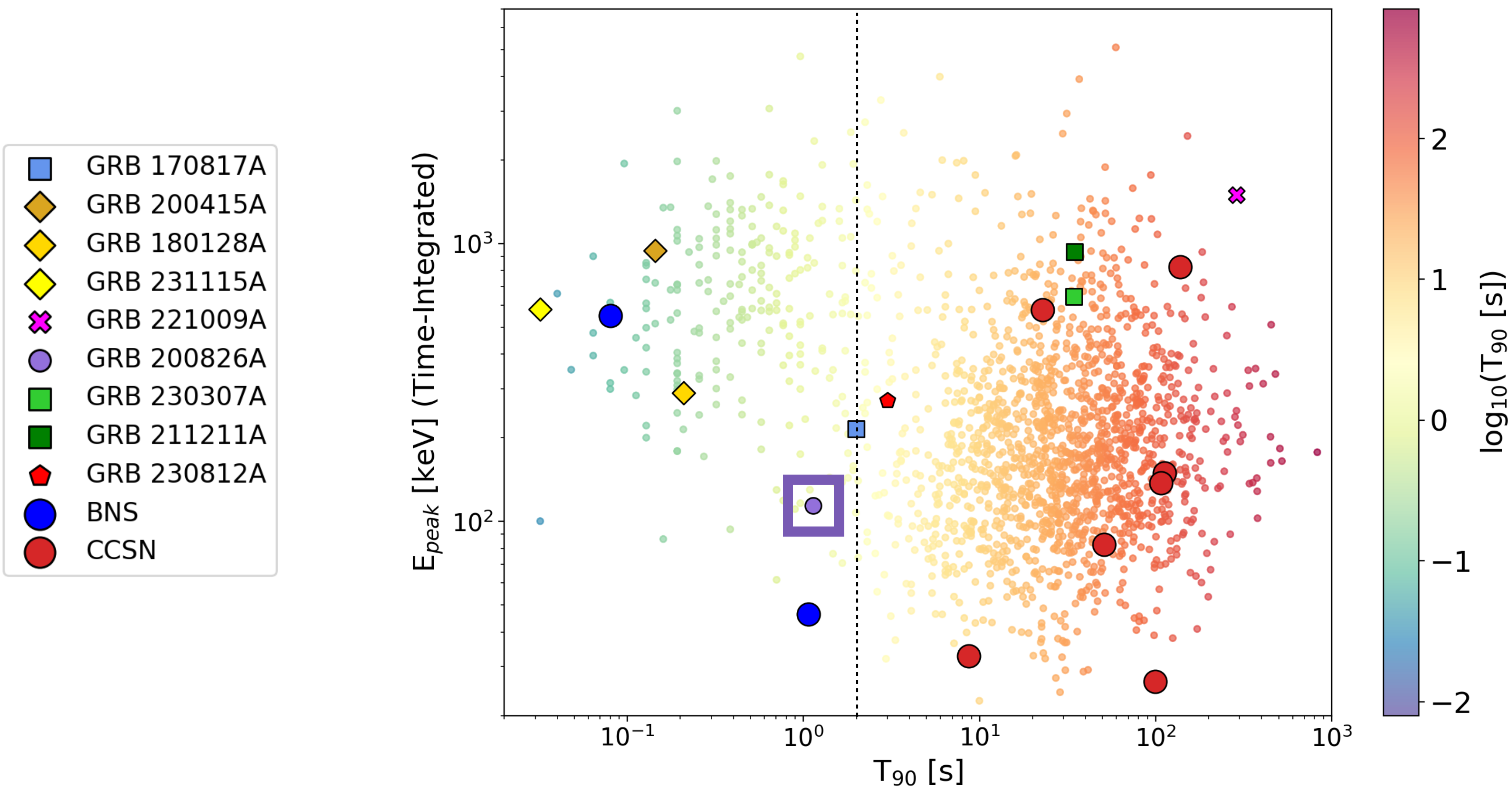


GRB 170817A: A confirmed BNS merger on the 2s threshold

GRBs classification

Categorization of GRBs is not trivial

Recent discoveries challenge the traditional classification

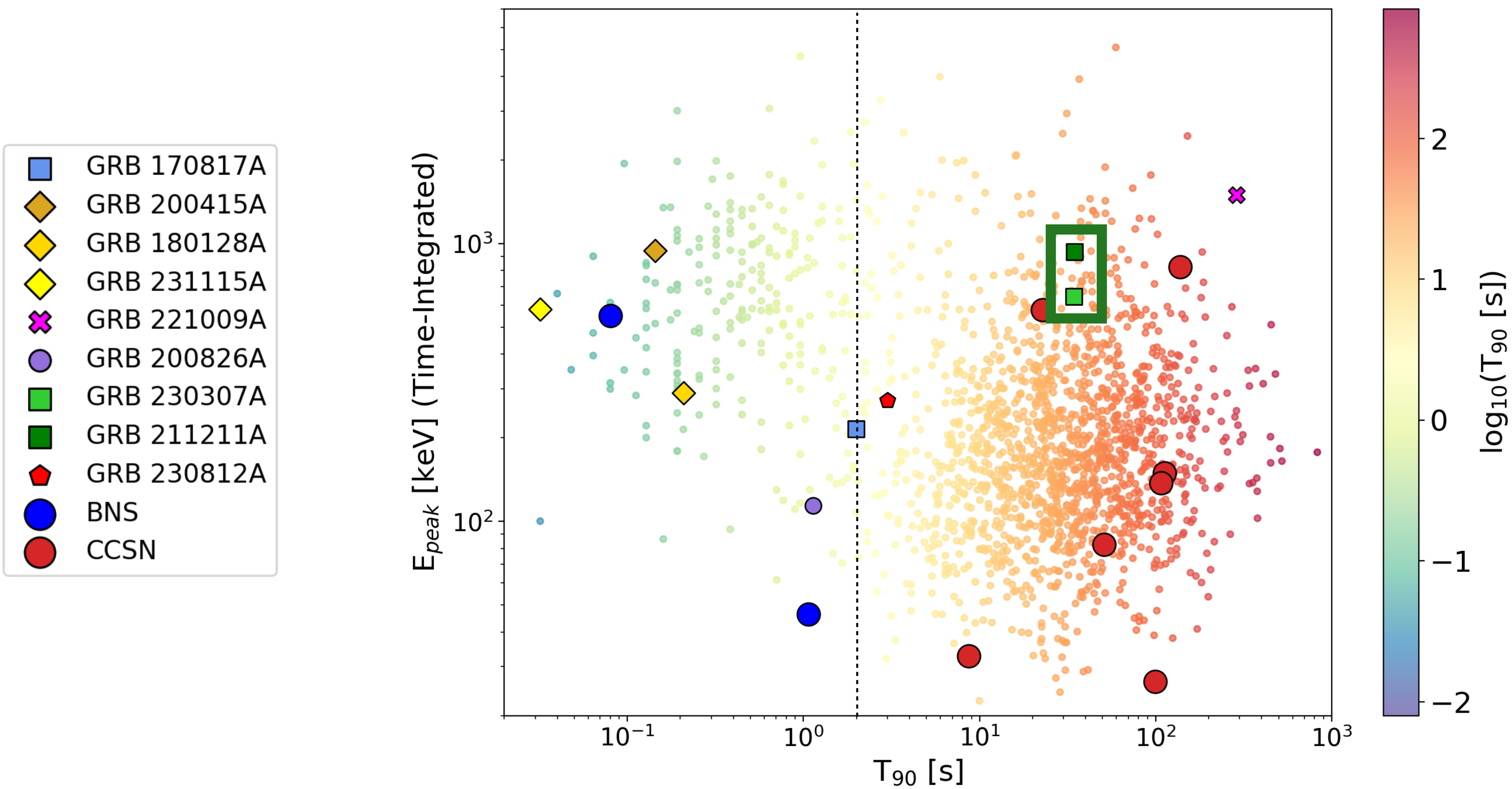


GRB 200826A: a collapsar GRB under the 2s threshold

GRBs classification

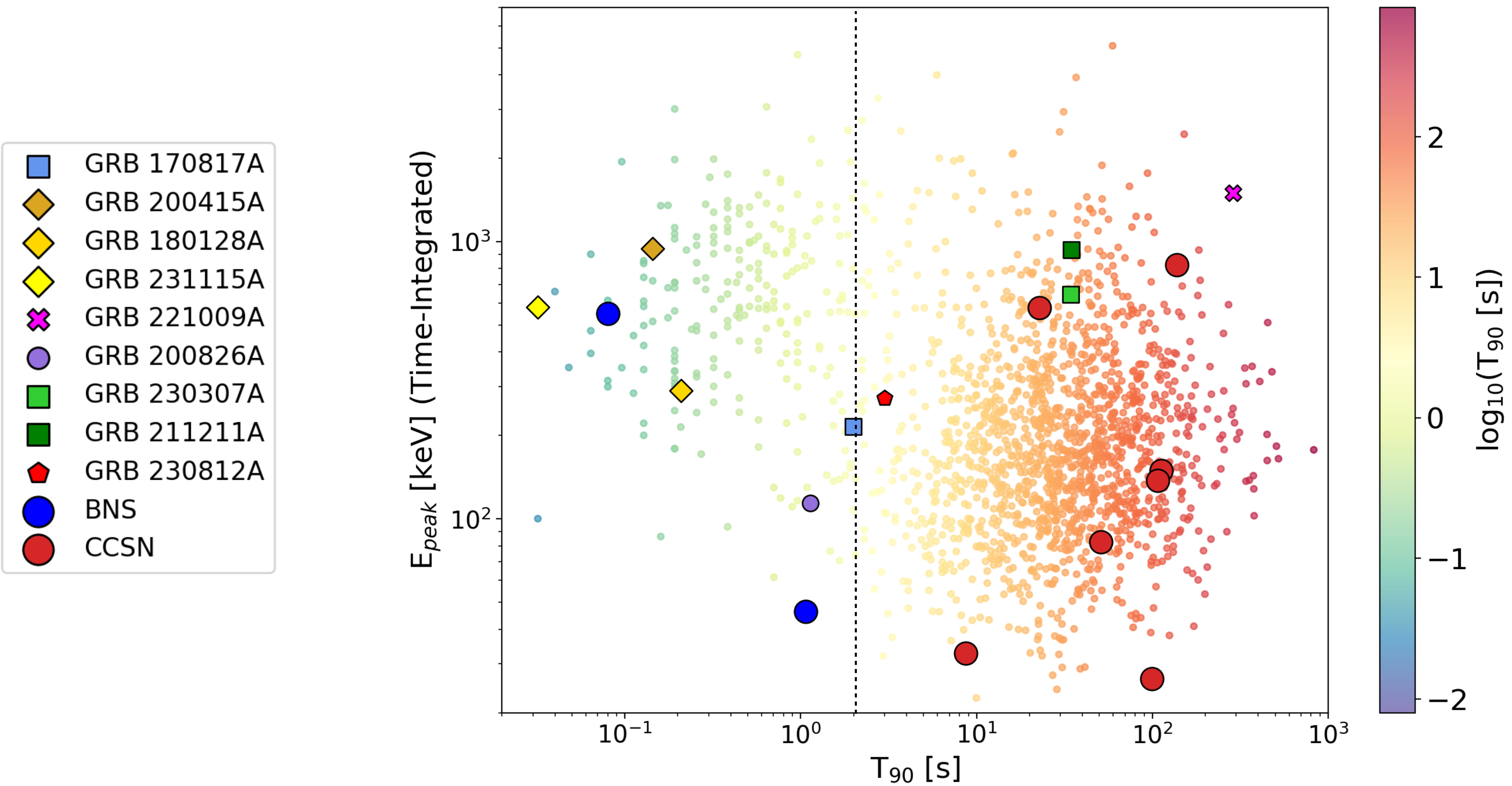
Categorization of GRBs is not trivial

Recent discoveries challenge the traditional classification



GRBs 230207A and 211211A: the long mergers

GRBs classification



Ultimate goal

Rapid identification of the progenitor of a given event, allowing for specific follow-up observations to occur

Unsupervised learning for GRBs classification

The application of unsupervised ML techniques for GRBs classification is not new

Chattopadhyay & Maitra 2017
Acuner & Ryde 2018
Jespersen et al. 2020
Steinhardt et al. 2023
Garcia-Cifuentes et al. 2023
Dimple et al. 2023
Chen et al. 2023
Mehta & Iyyani 2024
Zhu et al. 2024

and more!

GRBs light curve

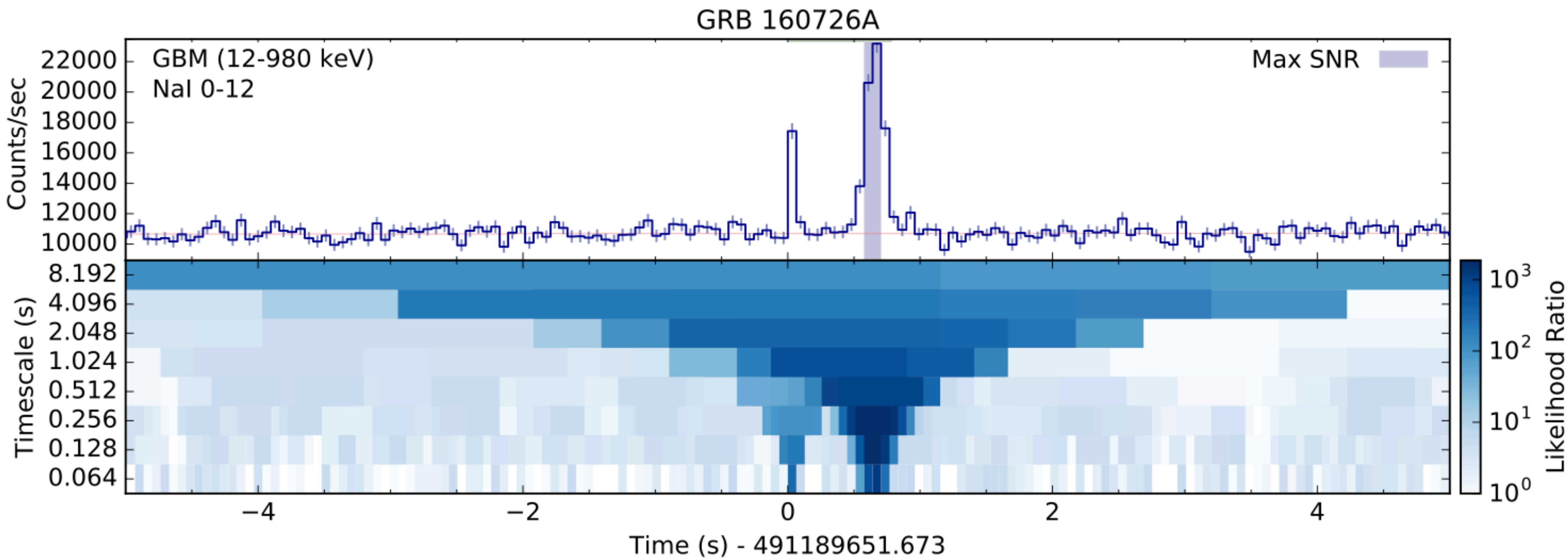
OR

GRBs spectral parameters

Dimensionality reduction algorithm

Final
2D GRBs distribution

A new input: Waterfall plots



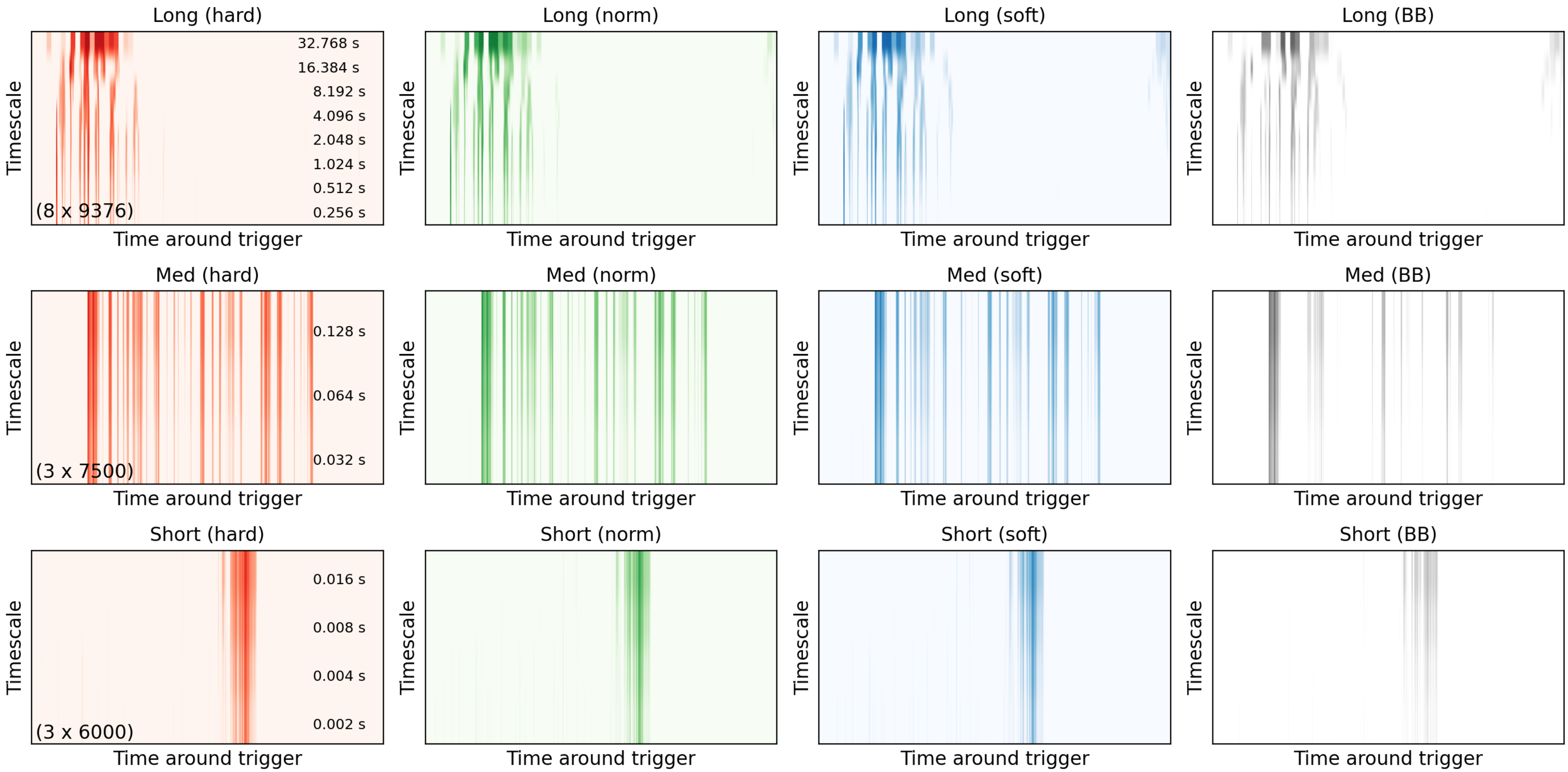
Probability of observed counts being from a source of amplitude S $P(c|\boxed{S})$

Probability of observed counts being from the background (S=0) $P(c|B)$

Likelihood Ratio $ln[\frac{P(c|S)}{P(c|B)}]$

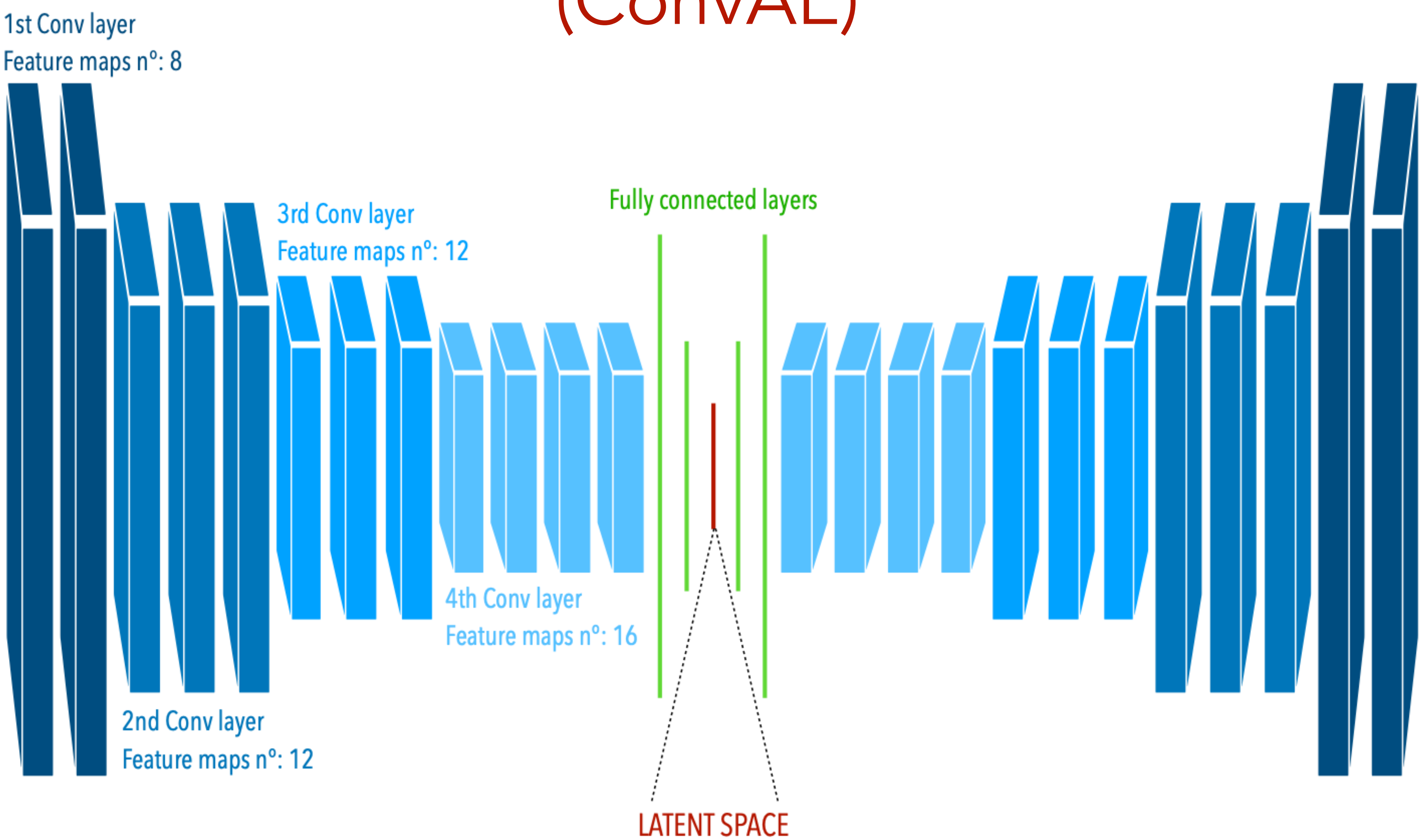
Different spectral shapes representative of typical GRB spectra
Hard - Normal - Soft - Blackbody

Waterfall plots for a single GRB



They contain core prompt information relevant for GRB classification, such as duration, temporal variation, pulse structure, spectral hardness and evolution, and how these parameters relate

Convolutional Autoencoders (ConvAE)

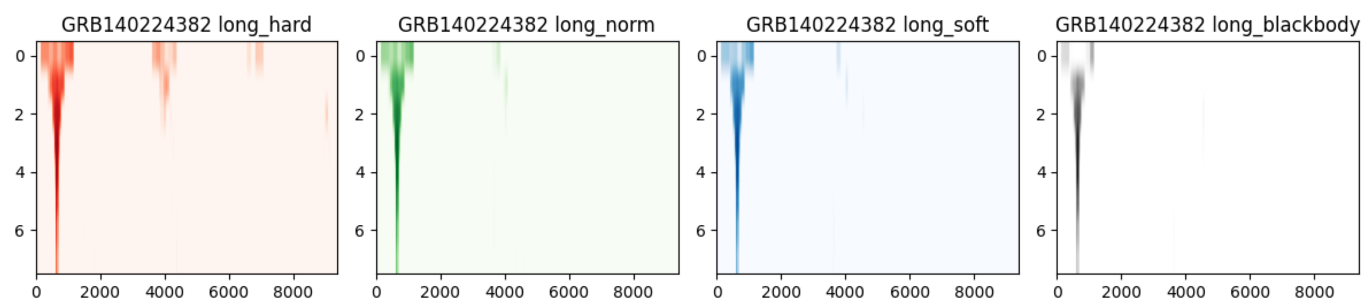


$$Loss = ||\underbrace{x}_{Input} - \underbrace{\hat{x}}_{Output}||^2$$

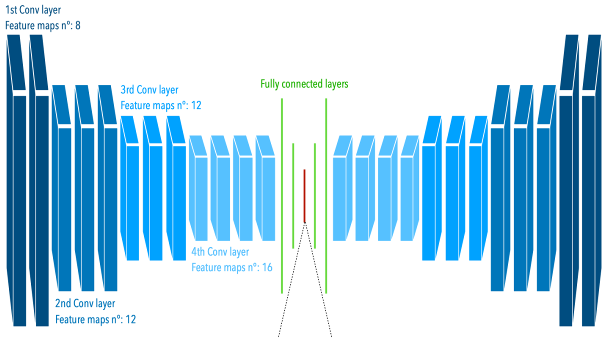
Main goal:
Learn input images representations in a lower dimensionality latent space

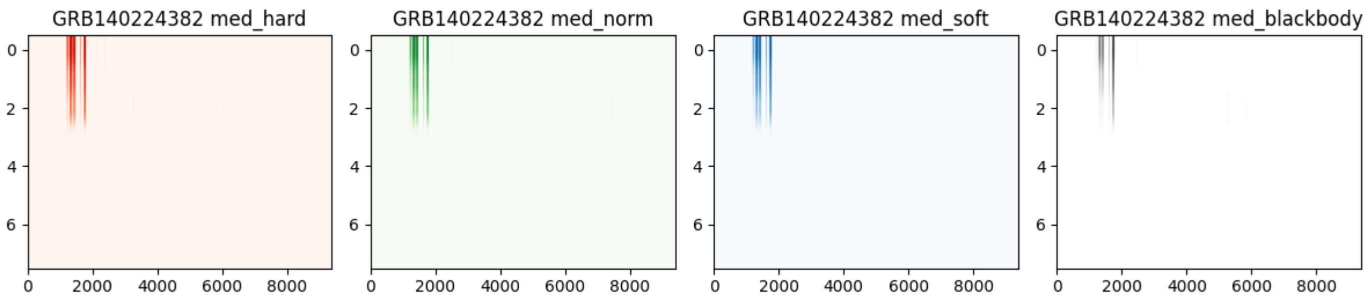
Our model

Dataset: 2371 GRBs (all GBM triggers from Jan 2013 to May 2023)

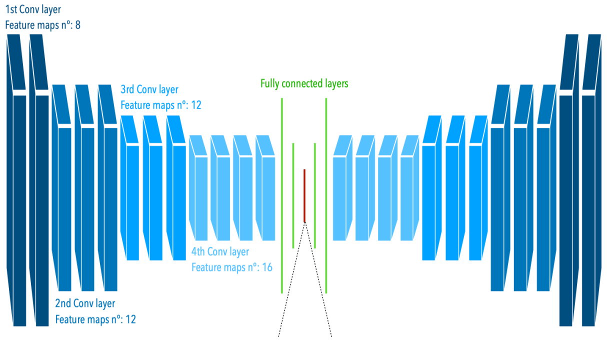


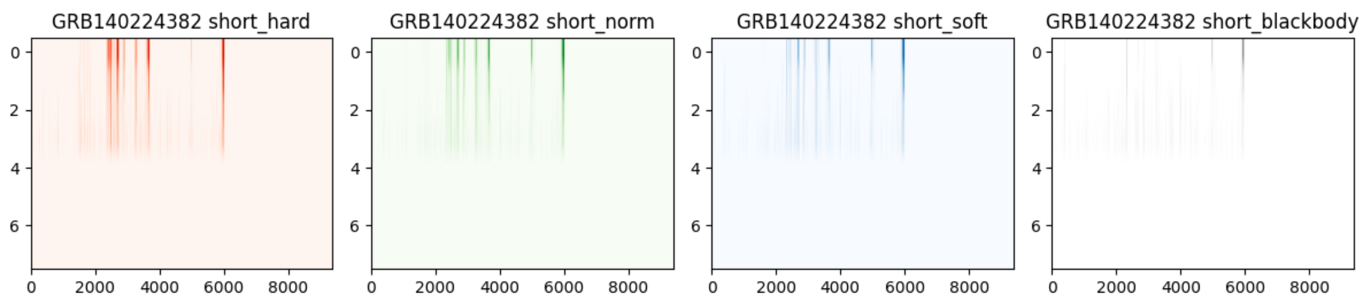
ConvAE 1 — Long timescales
(4 images, 8x9376)



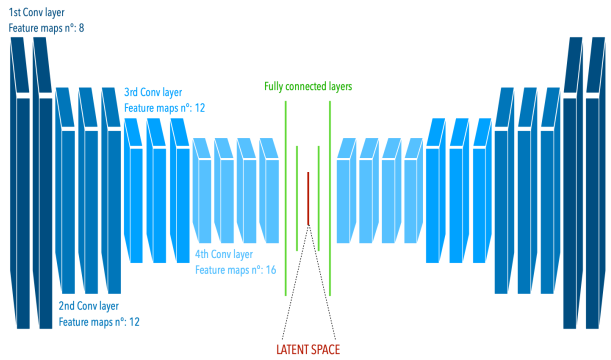


ConvAE 2 — Medium timescales
(4 images, 3x7500)





ConvAE 3 — Short timescales
(4 images, 3x7500)



We train 3 ConvAEs, one for each timescale

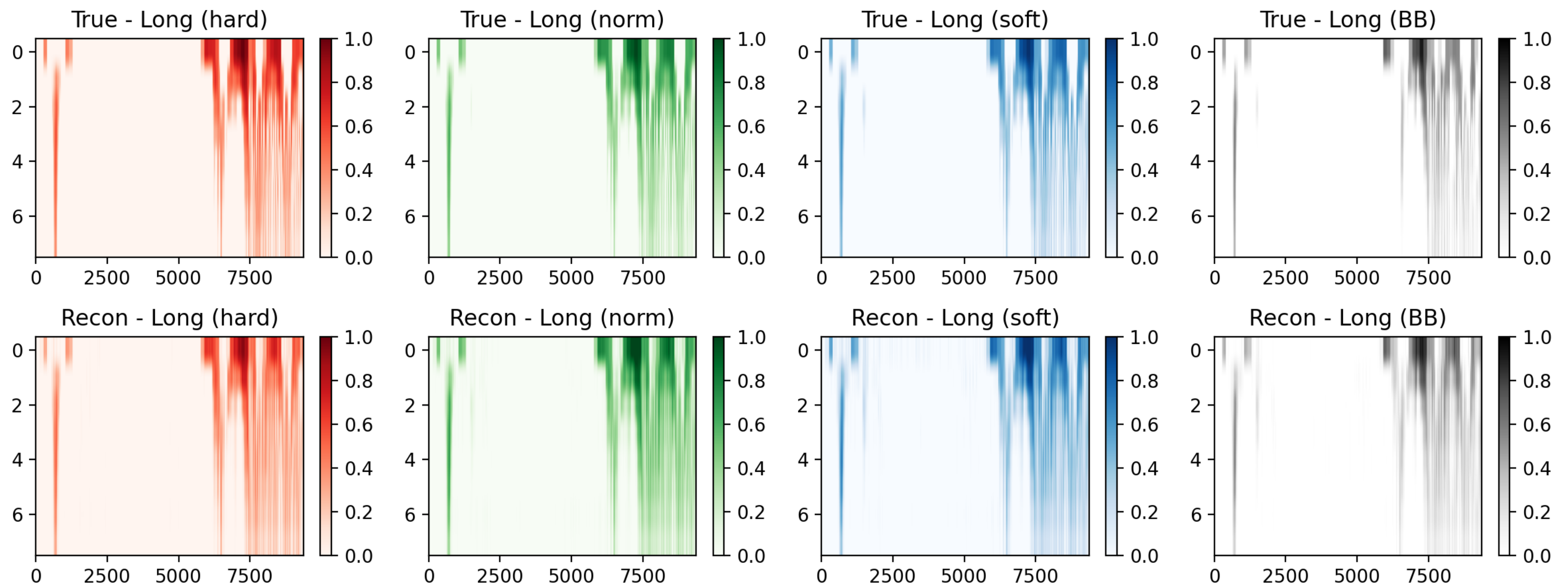
Final latent space 30D	=	Long timescales		Medium timescales		Short timescales
		10D	+	10D	+	10D

Evaluate the performances

QUALITATIVELY

Evaluation of the similarity between the input and output images for random subsamples of the GRBs dataset

GRB 221009A



QUANTITATIVELY

% of reconstructed pixels within 0.1 from the true value

For pixels != 0 in the original image

Average of all images

Reference value

For pixels = 0 in the original image

Average of all images

Check that it's > 95%

Dimensionality reduction with UMAP

Once the ConvAEs are trained, we have **30 variables for each GRB** (10 for each timescale)

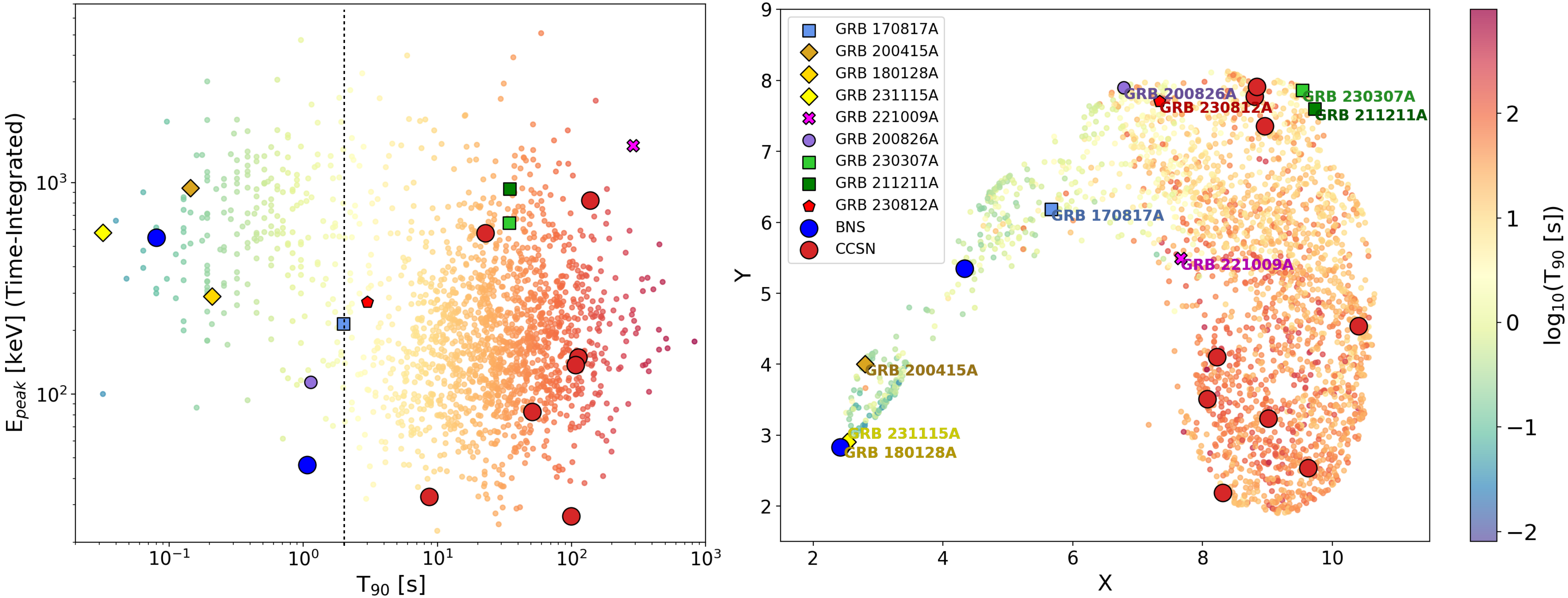
We further reduce the dimensionality of the latent space through UMAP

UNIFORM MANIFOLD
UMAP
APPROXIMATION & PROJECTION

UMAP reduces dimensionality by preserving data structure through manifold learning and neighbor graphs

30D combined latent space $\xrightarrow{\text{UMAP}}$ 2D (or 3D) GRBs distribution

Preliminary results



Results were published on ApJ

Prompt GRB recognition through Waterfalls and Deep Learning

[Negro, Cibrario et al. 2025]

How to improve the algorithm

WORK IN PROGRESS!

Optimization of the input and the ML model

- From 12WP -> 8WP (images dimensions are the same!) -> 1 single AE can be used
- Background model improvement

Reliability of the 2D distribution after UMAP

We can assess the reliability of the AE both qualitatively and quantitatively. How can we do the same for UMAP?

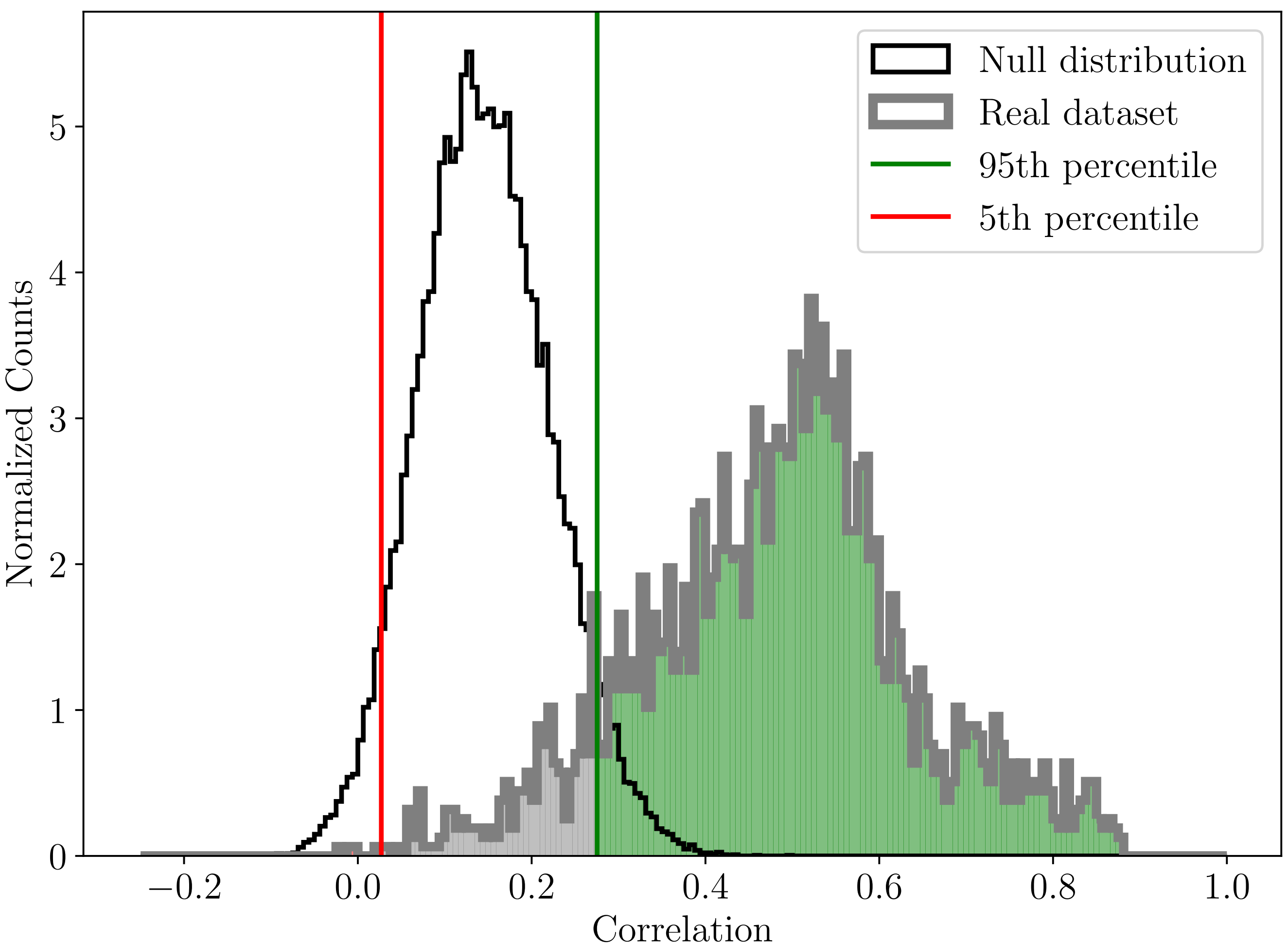
Classification

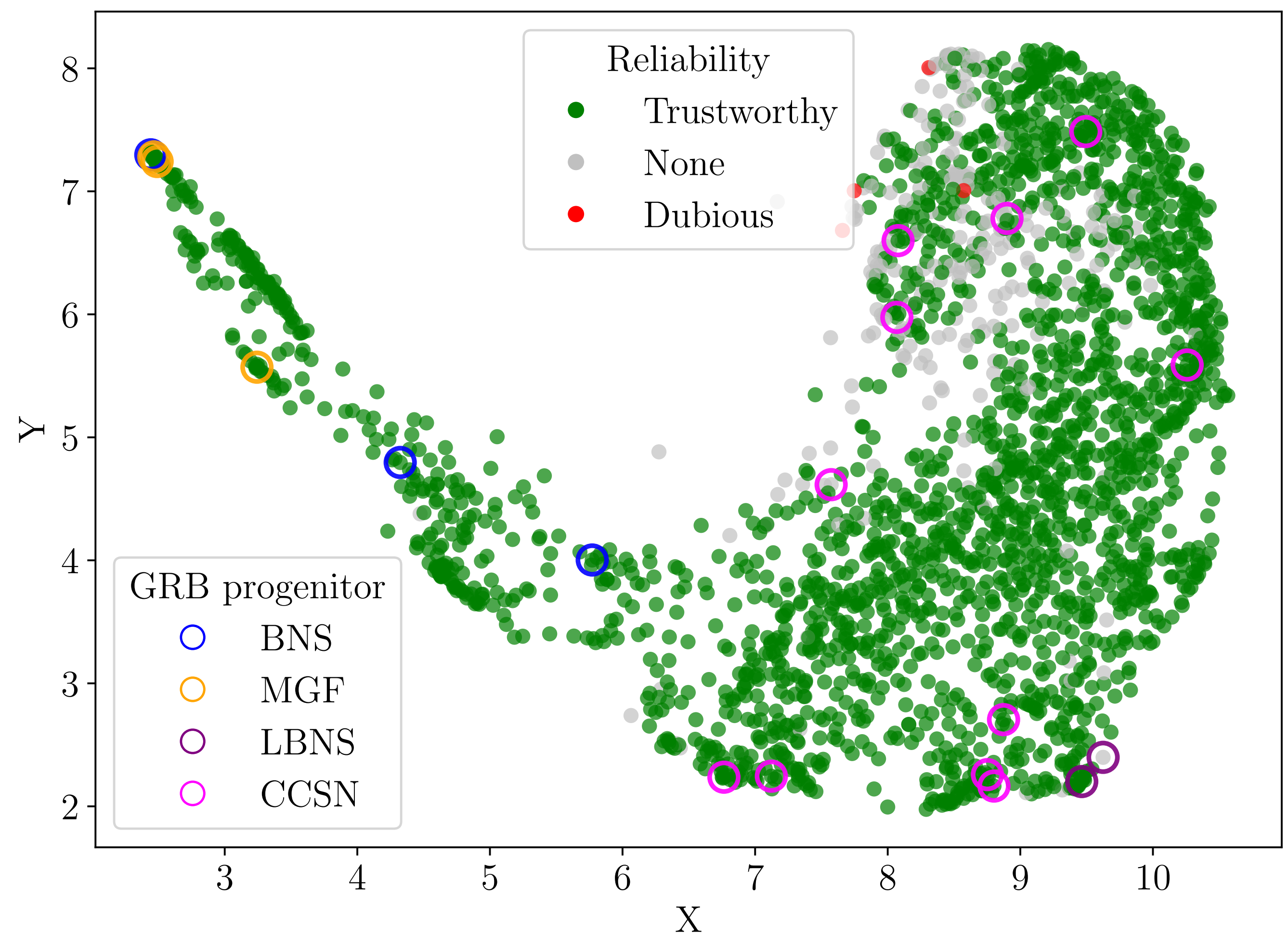
How do we identify a progenitor for each GRB?

2D distribution reliability

When we reduce from 30D to 2D, UMAP has no trivial way to assess the reliability of the position of the single events in the 2D space: is the algorithm distorting what we would see in 30D?

Solution: application of scDEED algorithm (Xia et al. 2024), which can give a reliability score to each 2D point.





Cibrario & Negro (2025)

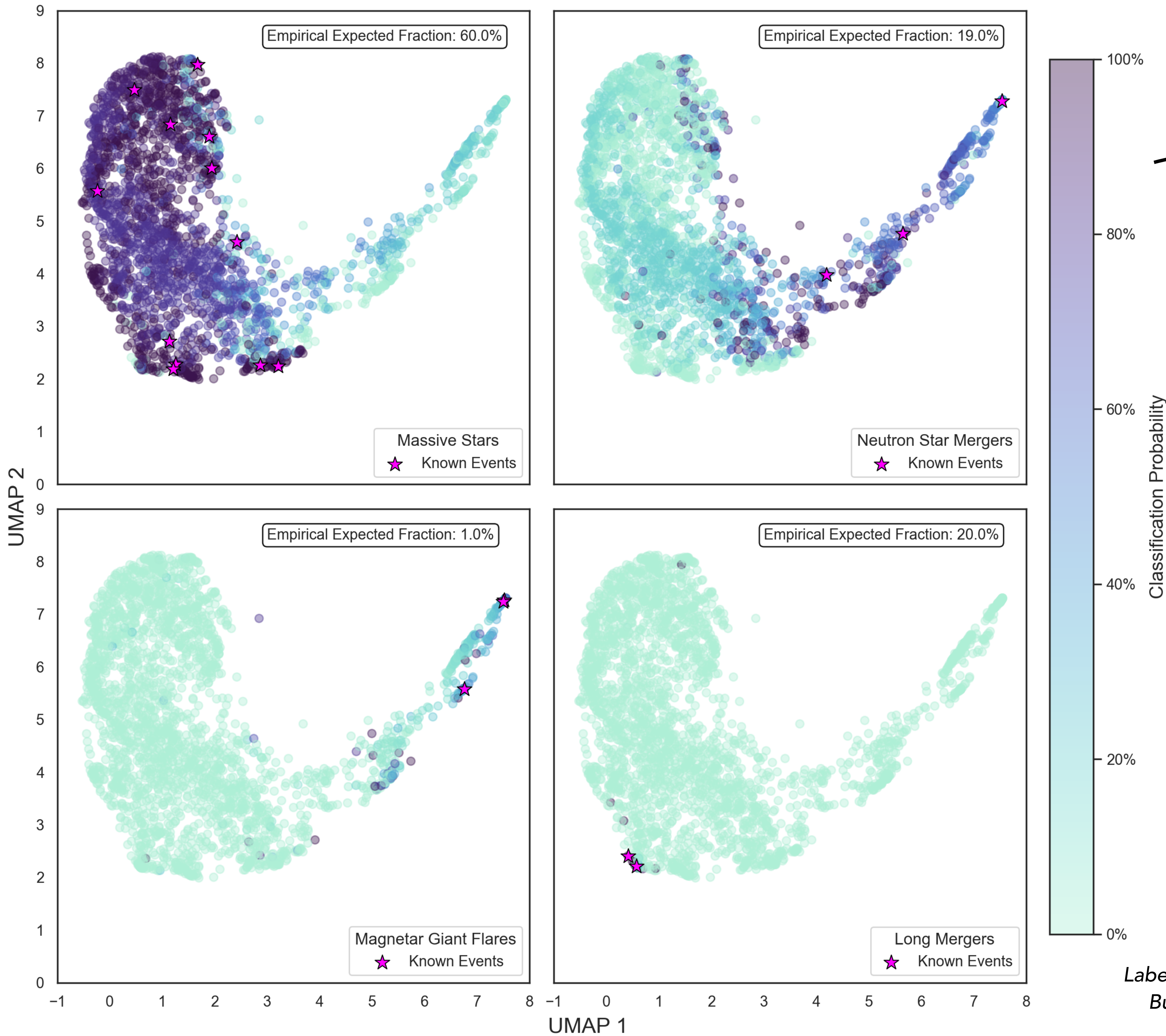
Embedding Reliability for Unsupervised Classification of Gamma Ray Burst progenitors from Prompt Gamma-ray Emission

Pearson correlation

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

Classification step

Semi-supervised approach: we apply the Label Propagation algorithm to the 30D GRB distribution.



For each GRB, a probability of belonging to each progenitor class will be available

Strain, Cibrario et al. (2026)
Label Propagation for Identifying Gamma-Ray Burst Progenitors from Prompt Emission

Conclusions

New comprehensive input for ML algorithm
We built waterfall plots, high-dimensional images containing core prompt emission information

Dimensionality reduction with AE + UMAP
We built an algorithm which allows us to represent the 2D/3D GRB distribution starting from a high-dimensional space

Classification of GRBs
We developed a semi-supervised algorithm for identifying unknown GRB progenitors

|

NEXT STEPS

Optimization of input images and ML model

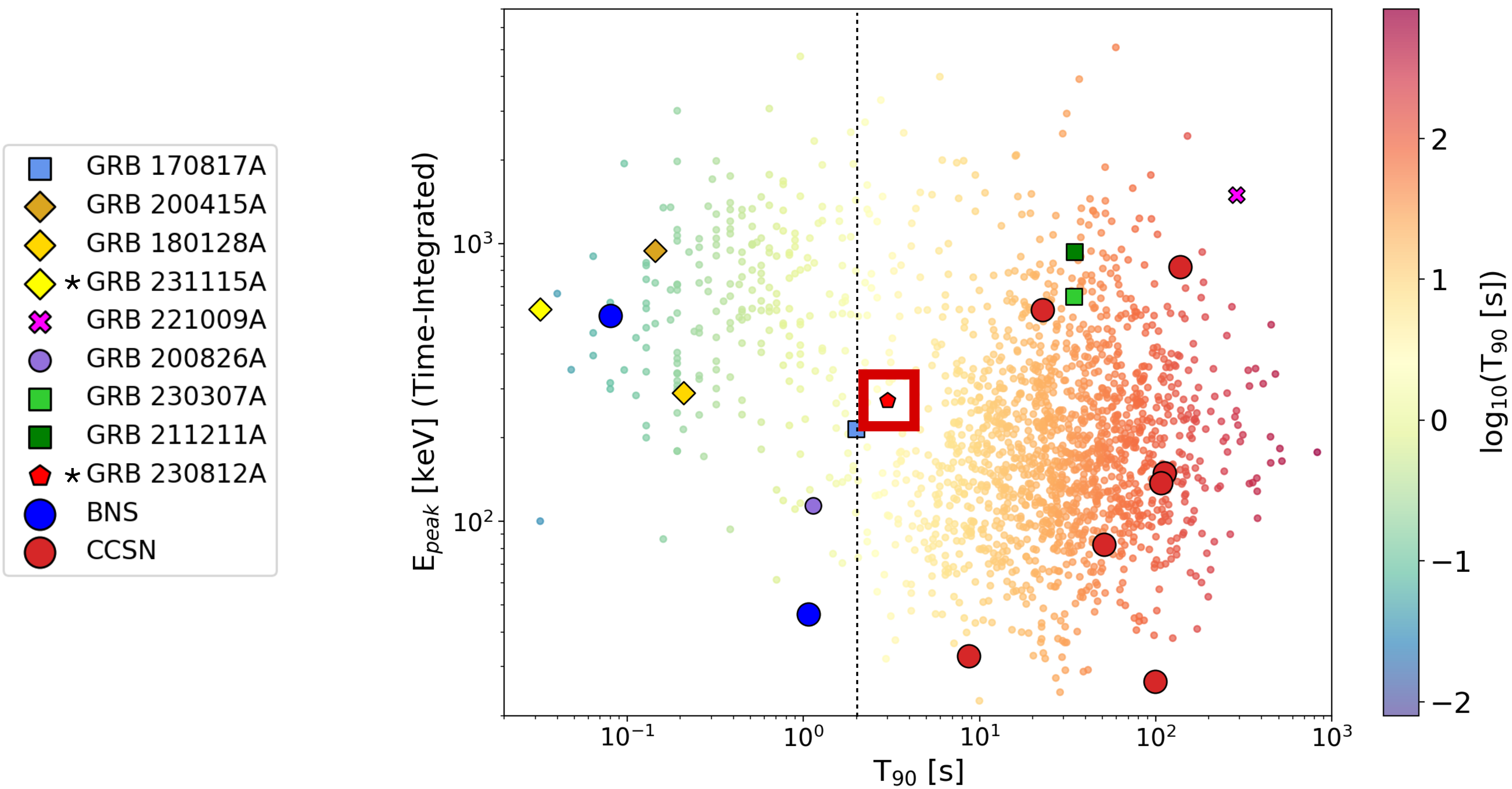
Merging all the steps to achieve a fast algorithm for progenitor association to new detected GRBs

[Interactive portal]

Backup

GRBs classification

Categorization of GRBs is not trivial
Recent discoveries challenge the traditional classification

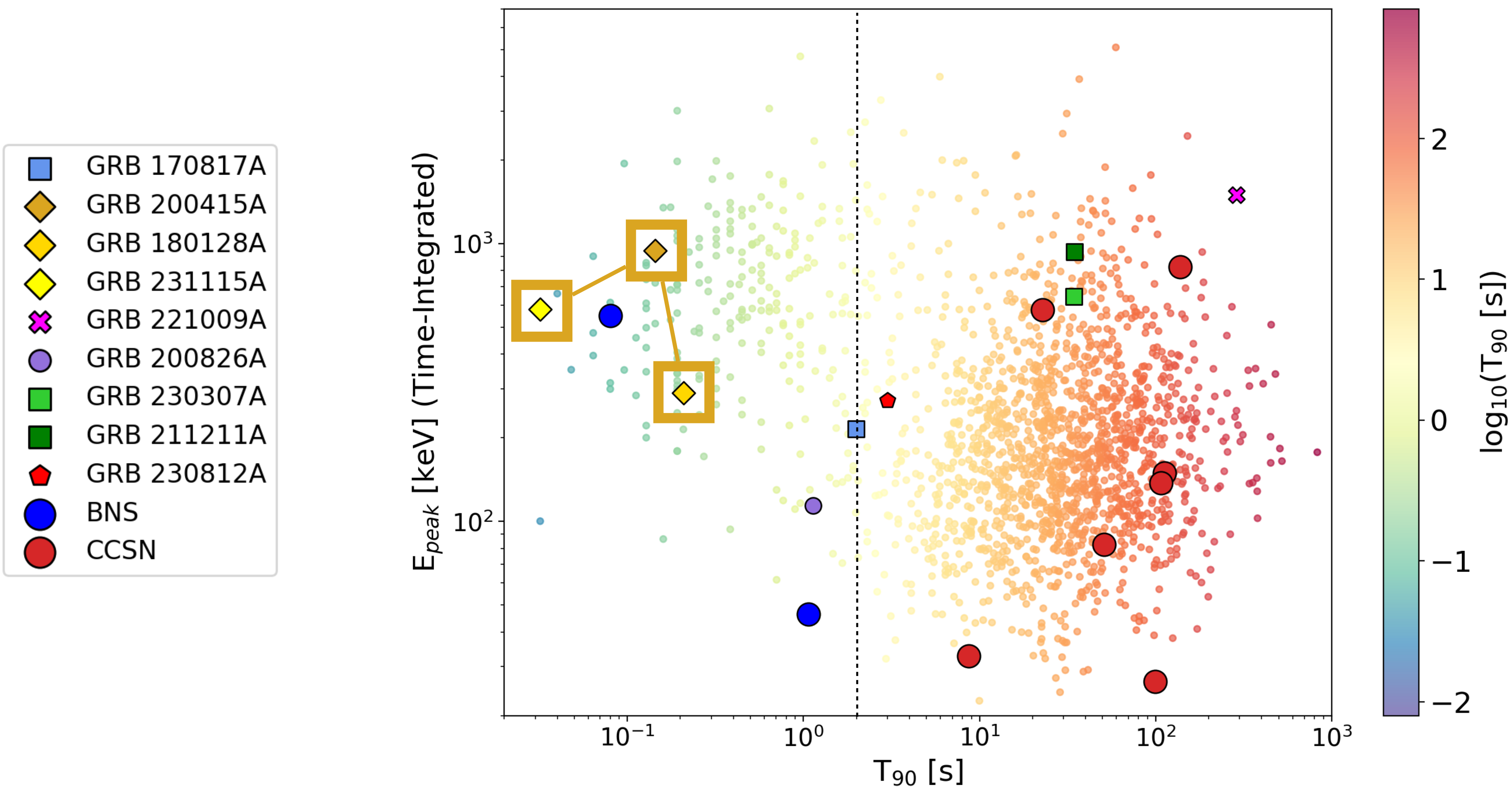


GRB 230812: a close, quite short collapsar

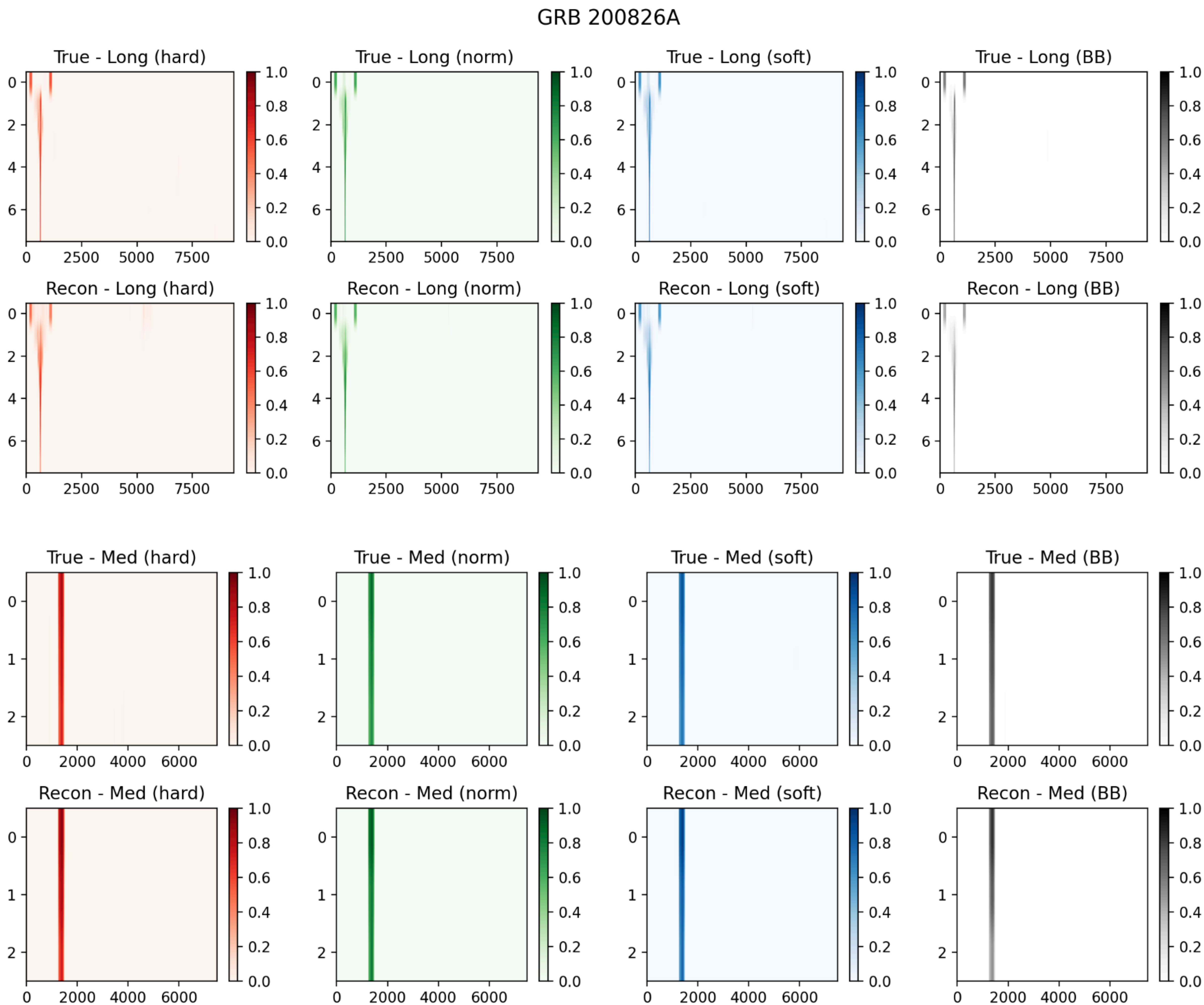
GRBs classification

Categorization of GRBs is not trivial

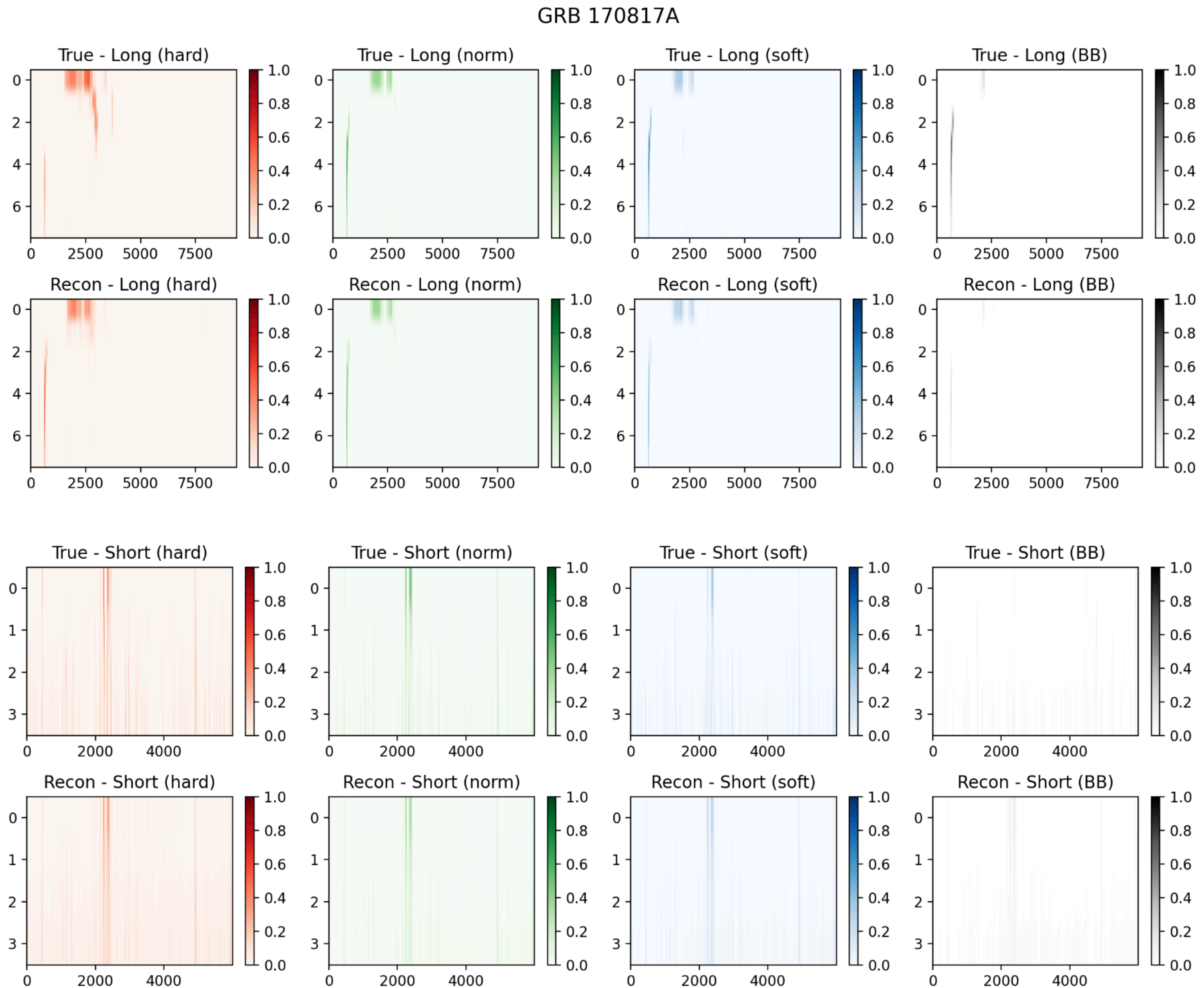
Recent discoveries challenge the traditional classification

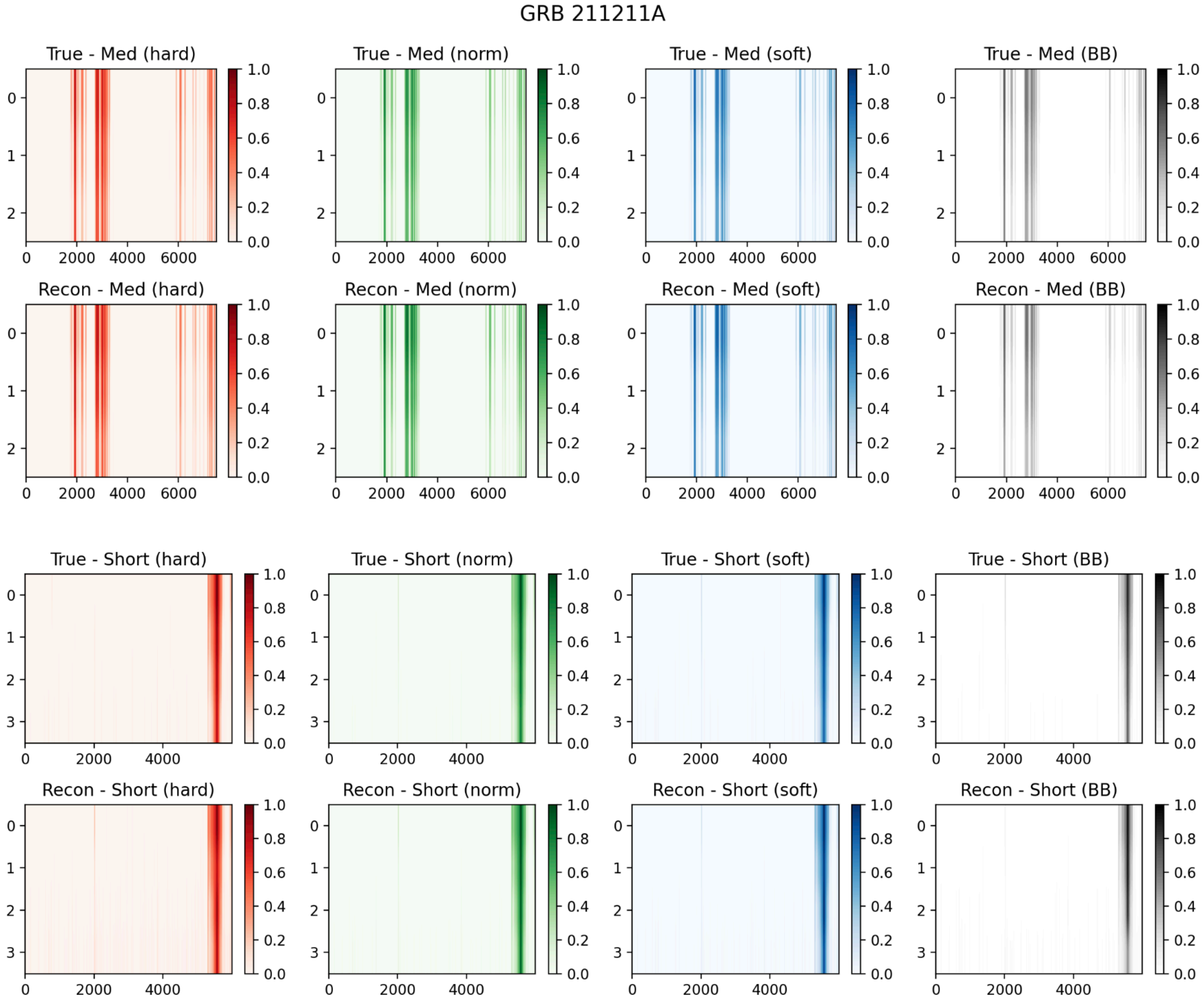


Magnetar Giant Flares candidates: indistinguishable from short GRBs



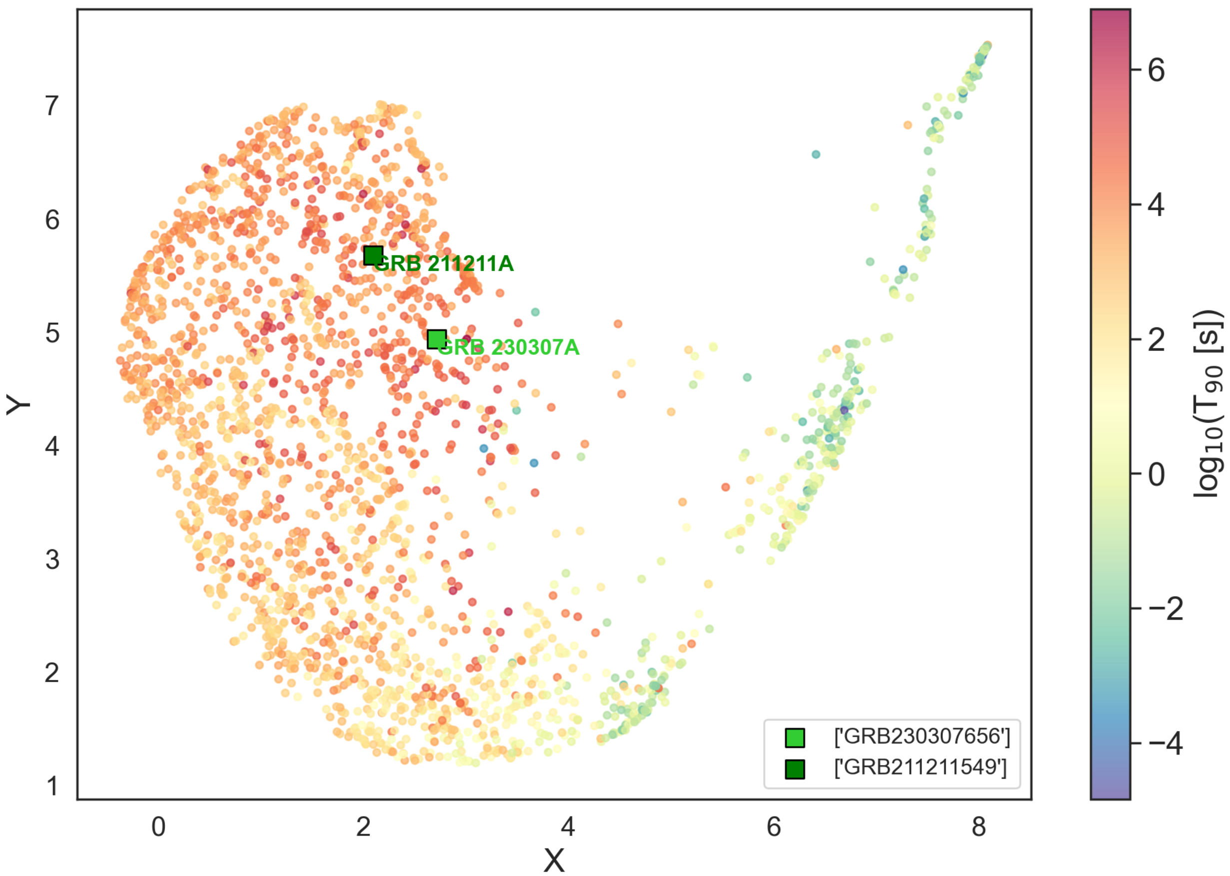
GRB 200826A
True - Long (hard)
True - Long (norm)
True - Long (soft)
True - Long (BB)
Recon - Long (hard)
Recon - Long (norm)
Recon - Long (soft)
Recon - Long (BB)
True - Med (hard)
True - Med (norm)
True - Med (soft)
True - Med (BB)
Recon - Med (hard)
Recon - Med (norm)
Recon - Med (soft)
Recon - Med (BB)


GRB 170817A
True - Long (hard)
True - Long (norm)
True - Long (soft)
True - Long (BB)
Recon - Long (hard)
Recon - Long (norm)
Recon - Long (soft)
Recon - Long (BB)
True - Short (hard)
True - Short (norm)
True - Short (soft)
True - Short (BB)
Recon - Short (hard)
Recon - Short (norm)
Recon - Short (soft)
Recon - Short (BB)

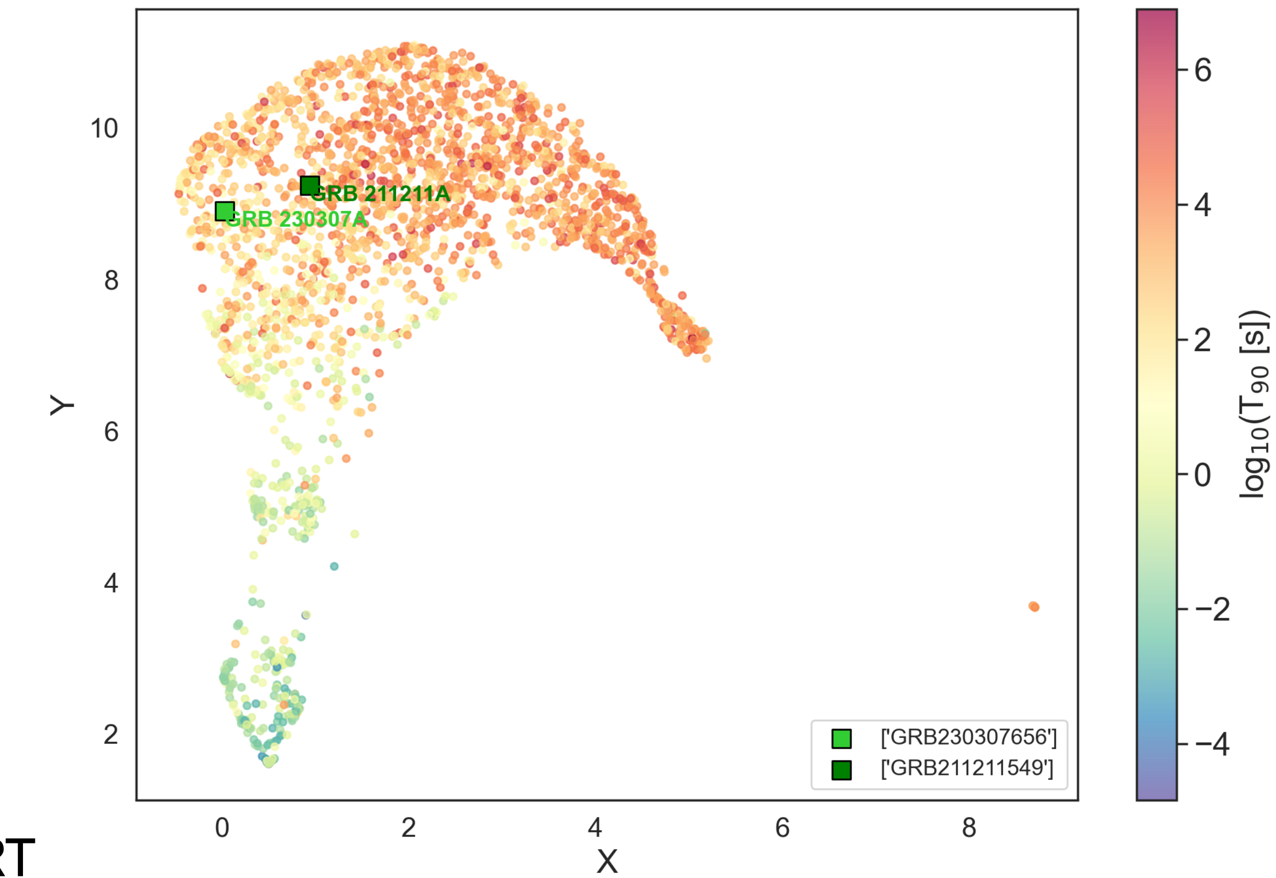

GRB 211211A
True - Med (hard)
True - Med (norm)
True - Med (soft)
True - Med (BB)
Recon - Med (hard)
Recon - Med (norm)
Recon - Med (soft)
Recon - Med (BB)
True - Short (hard)
True - Short (norm)
True - Short (soft)
True - Short (BB)
Recon - Short (hard)
Recon - Short (norm)
Recon - Short (soft)
Recon - Short (BB)

Single timescale distributions

LONG



MEDIUM



SHORT