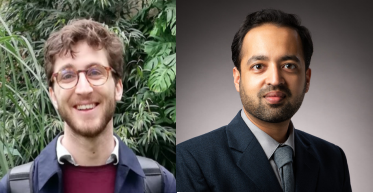


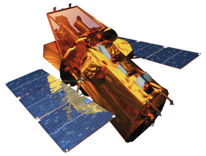
Machine Learning Applications in High-Energy Time-Domain Astrophysics

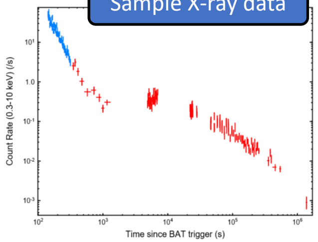
Eliot Ayache*, Tanmoy Laskar*, Tara Howard, Claire Anderson, Hendrik van Eerten

e.h.r.ayache@bath.ac.uk, t.laskar@bath.ac.uk

The Physics

Sample X-ray data

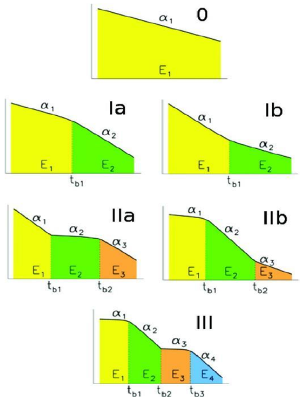




Gamma-ray bursts (GRBs) are bright flashes of light signalling the death of stars. **Above:** Example X-ray light curve (brightness over time) of a GRB from the *Swift* satellite. *Swift* is discovering roughly 100 GRBs every year, autonomously slewing to capture their fading X-ray counterparts within minutes. *Swift's* quick sleuthing has revealed an inexplicably large diversity of X-ray light curves (below), not all of which is fully understood.

Right: A traditional phenomenological classification scheme for GRB X-ray light curves. The brightness declines as a power law with time ($F \propto t^{\alpha}$). Different power law slopes (α) and transitions between them may encode information about the physical processes producing the X-rays. From Margutti et al. (2013)

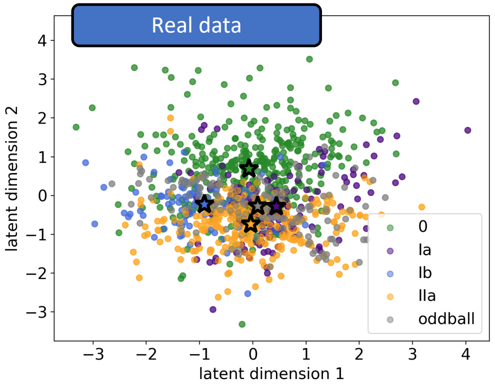
Traditional classification



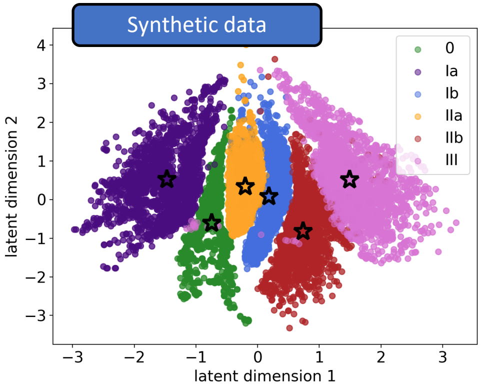
Are there hidden patterns in the properties of **Gamma-ray bursts**, the most energetic explosions in the Universe? We trained autoencoders (right column) on X-ray observations of 1300 Gamma-ray bursts (left column) to search for physical patterns in the latent space (results below).

Analysis and Results

Real data

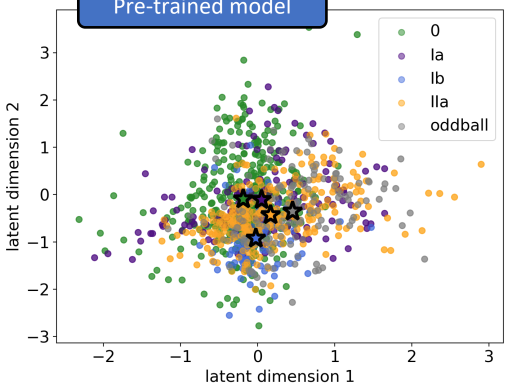


Synthetic data

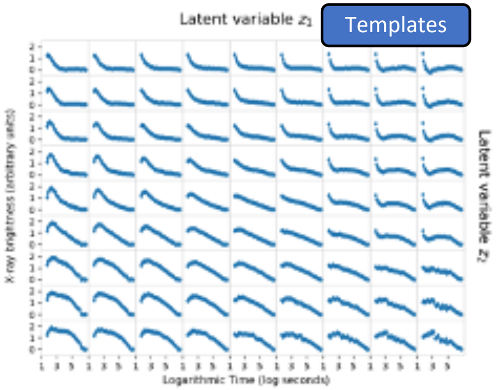


Above: A CNN-based autoencoder provides no evidence for latent-space-based clusters in the X-ray light curves (left panel), but correctly clusters a stratified, synthetic training set (right panel). **Below:** Even when trained on synthetic data with well-defined clusters, the CNN model does not yield clusters in the real data, suggesting that the traditional, phenomenological GRB X-ray light curve classes may be somewhat arbitrary (left panel). Instead, templates from the latent-space representation of the data reveal that light curve morphology spans a continuum (right panel).

Pre-trained model

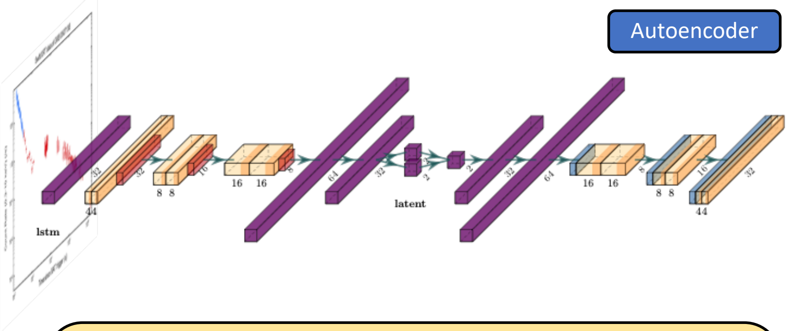


Templates



Architecture & Implementation

Autoencoder



Above: CNN-based autoencoder for modelling Gamma-ray burst X-ray light curves using time-series (with LSTM). Can the data be whittled down to a few parameters, and can we interpret the resulting latent space using physics? **Below:** Sample input/output pairs for an image-based autoencoder employed to tackle the same problem. The decoder reconstructs the images with high fidelity.

Sample fits

Conclusions: There is a large diversity in X-ray afterglow light curves, and no clear evidence for traditional categories when using ML-based clustering methods. The inferred continuum of light curves may be representative of a competition between multiple underlying physical processes. ML recovers clusters for large enough data sets (>1k samples / label); such data sets may be feasible with future instruments (e.g. SVOM, Athena).

Future work:

- Improved pre-processing steps
- Data augmentation
- Other clustering algorithms (t-SNE, SOM)