Comparing Classical and Quantum Deep Learning Techniques for Anomaly Detection of Short-Duration Gamma-Ray Signals

USC-C General assembly, Trieste, 10th March 2026

Alessandro Rizzo - alessandro.rizzo@inaf.it
INAF - OACT

The goal of this work

- **Goal** → Compare the detection capabilities of short GRBs in the **COSI BGO light curves** using machine learning and quantum machine learning anomaly detection techniques.

- **Anomaly detection task** → Detect anomalies (short GRBs) from COSI BGO background light curves → An anomaly might indicate a GRB.

- Reproduce neural networks developed by **AGILE** team to compare quantum and classical machine learning algorithms. ***(Parmiggiani et al., 2023)***

- A key focus is comparing perfromance under resource-contrained scenarios such as limited training data and fewer trainable parameters

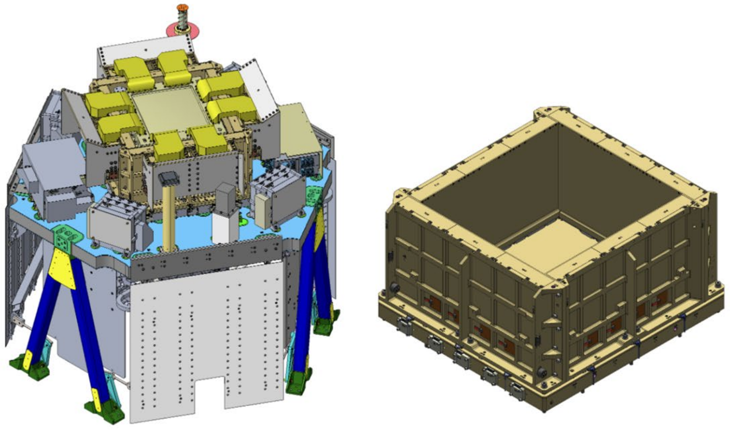
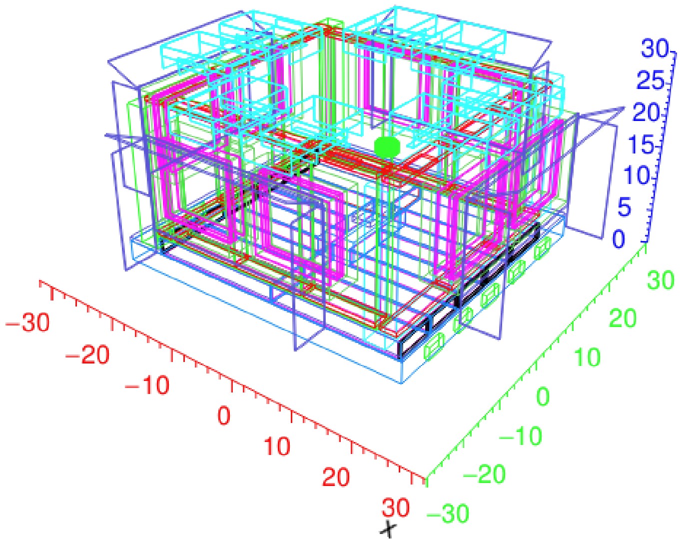


Fig. 1 COSI payload design (left) and ACS design (right)



COSI (Compton Spectrometer and Imager)

- **COSI (Compton Spectrometer and Imager)** is a compact soft gamma-ray survey telescope designed to perform pioneering studies of gamma-ray polarization, or probe the origins of Galactic positrons.

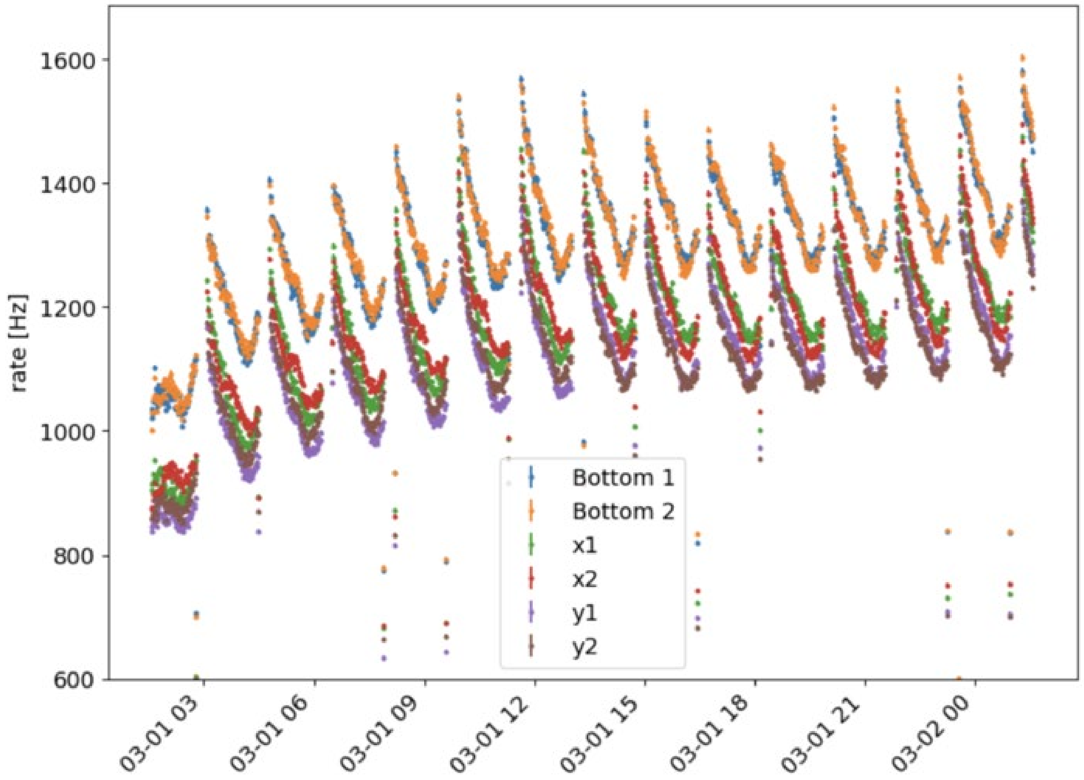
- It is designed to study the **0.2 – 5 MeV energy band**, an important region of the electromagnetic spectrum.

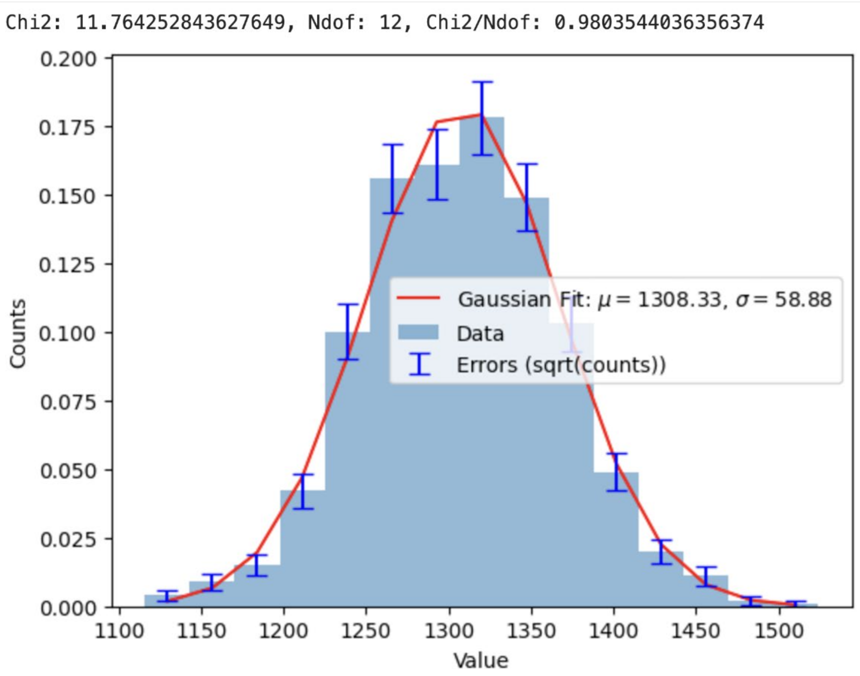
- Its **germanium detectors** have an instantaneous field of view of more than 25% of the sky.
    - They are surrounded on the sides and bottom by active shields (BGO), which provide background rejection while also allowing for detection of **gamma-ray bursts** and other gamma-ray flares across the majority of the sky.

- Scheduled for launch in 2027 as a NASA mission.

Dataset (1)

- To simulate new background light curves, we sampled from a Gaussian distribution with a mean of 1308 and a standard deviation of 58.





Samples	Training Set	Validation Set	Test Set	GRB Set
100	80	20	1×10^4	1×10^4
	80	20	1×10^4	1×10^4
1000	800	200	1×10^4	1×10^4
	800	200	1×10^4	1×10^4
5000	4000	1000	1×10^4	1×10^4
	4000	1000	1×10^4	1×10^4

Dataset (2)

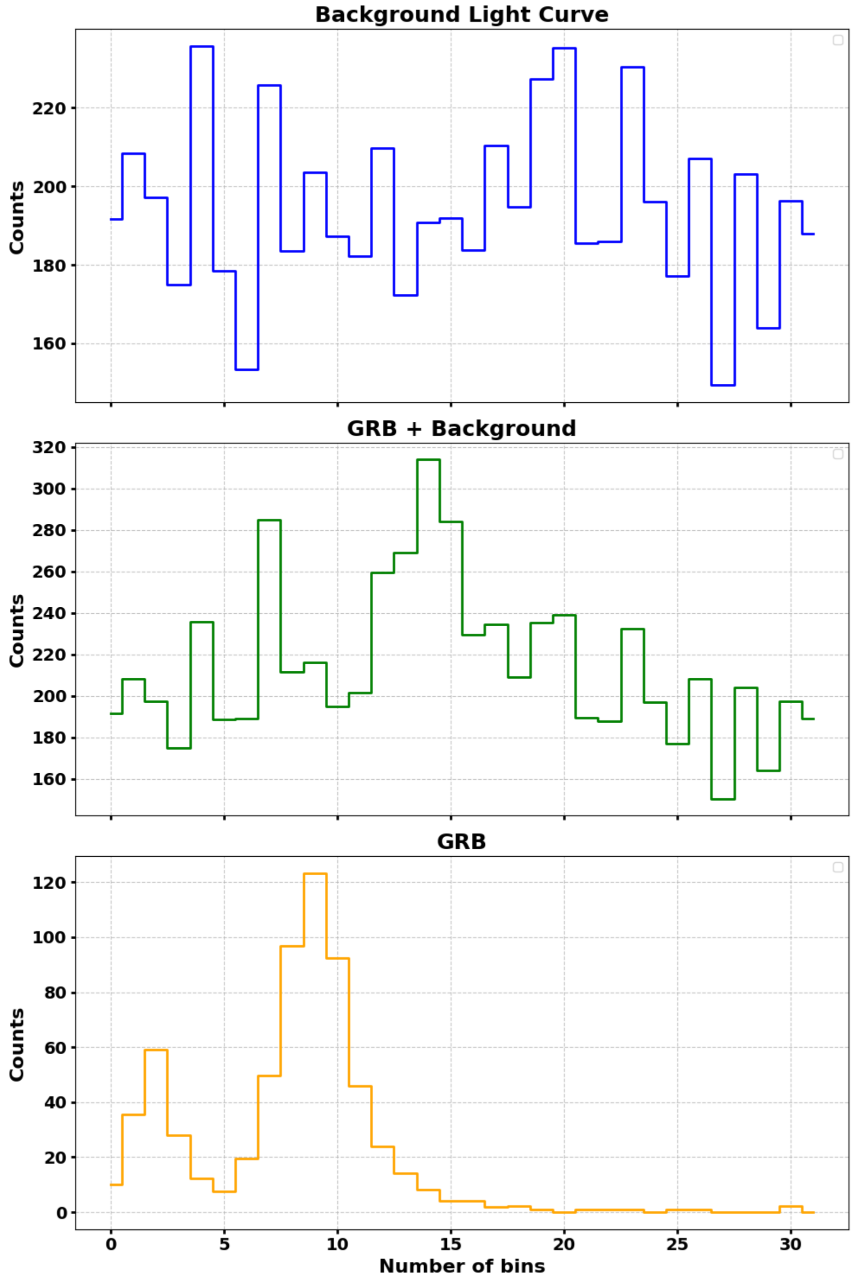
Gamma-Ray Bursts (GRB) Light Curves:

- Generated from the MEGAlib software with the DC3 COSI mass model, extracting the template from the **Fermi/GBM GRB** catalog.
- Number of counts reduced to increase detection challenge.
- Used only for model evaluation.

Dataset Preprocessing:

- **Min-Max normalization** applied to all data.
- Multiple dataset sizes tested.
- **Data Augmentation** used to expand GRB dataset → To create new GRB light curves, we multiplied each bin of the original one by a random value up to 10% of the counts of that bin.

- The models are trained only on background light curves for reconstructing normal patterns.

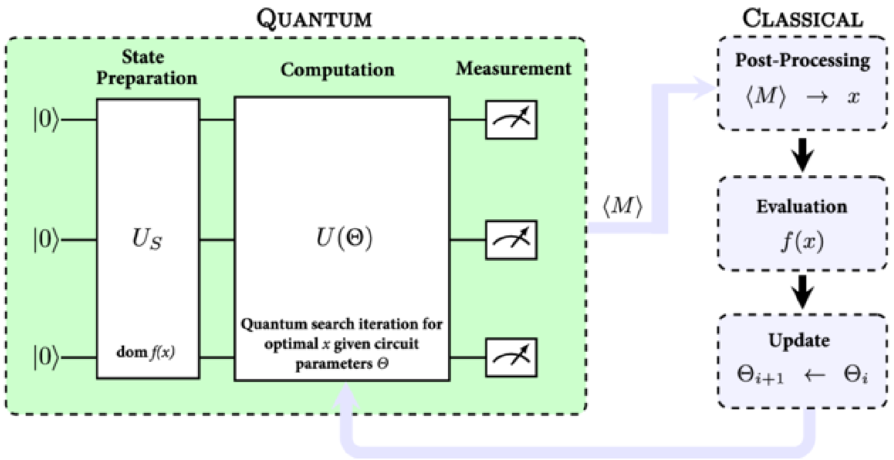


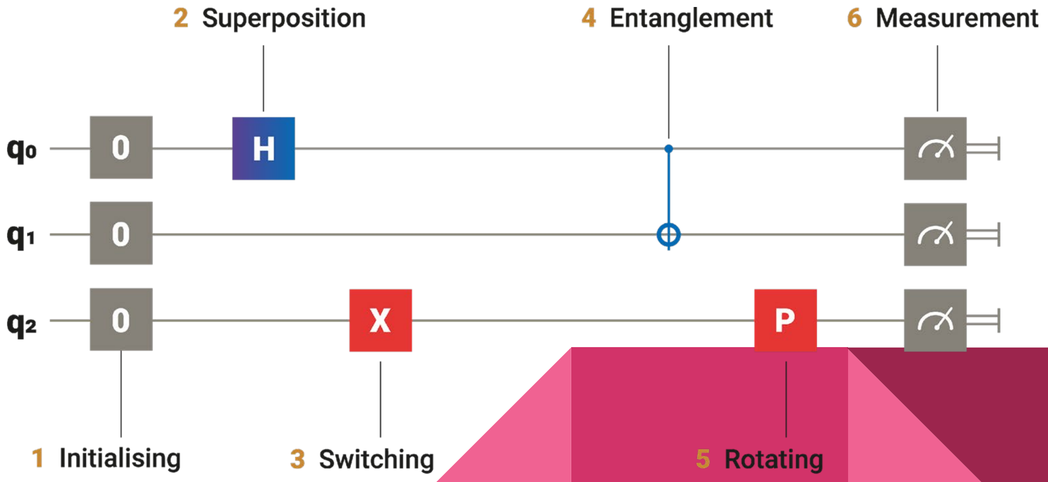
Quantum Computing

- **Quantum computing** is a revolutionary field that combines principles from **quantum mechanics** and computer science to create a new paradigm of computation.
- **Qubit** → Fundamental unit of quantum information.
- Fundamental Principles of Quantum Mechanics:
  - **Superposition** → A qubit can exist in a combination of states at the same time.
  - **Entanglement** → Two or more qubits become correlated in a way that the quantum state of each qubit cannot be described independently of the others..
  - **Quantum Measurement** → Process by which a quantum system in superposition collapses into a definite state when observed.
- **Current challenges** → Limited qubit availability, lack of fault tolerance and decoherence effects restrict the performance and computational time of quantum systems.
- **Quantum Advantages** → Leveraging superposition and entanglement to explore larger computational spaces, offering speedups for hard problems for classical computers.

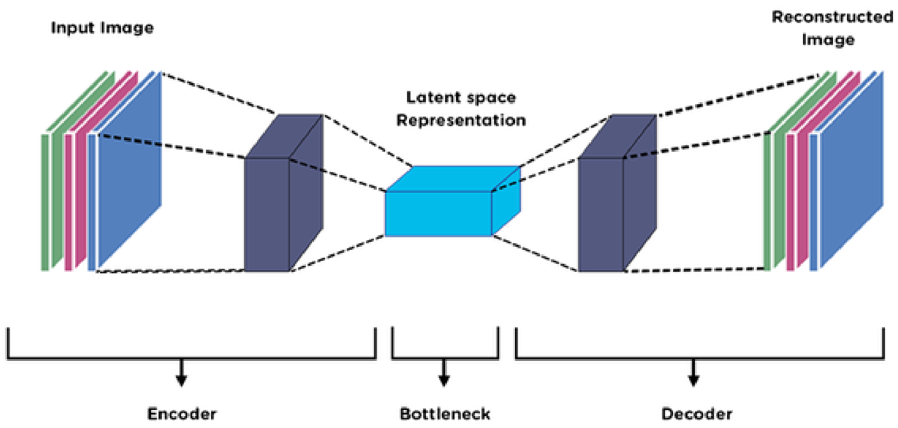
Quantum Machine Learning

- **Quantum Machine Learning** (QML) integrates quantum algorithms into ML models to potentially outperform classical approaches in specific tasks.
- **Quantum gates** serve as a fundamental building blocks for quantum circuits, analogous to logic gates in classical computing.
- **Parameterized Quantum Circuits** (PQCs) form the backbone of QML models, containing trainable parameters that are optimized during the learning process.
- QML leverages quantum properties to potentially represent highly complex functions with exponentially fewer parameters than classical networks.
- We need to represent and input the data into a **quantum system**, so that it can be processed by a quantum machine learning algorithm → A quantum embedding represents classical data as quantum states.





Autoencoder



- In Deep Learning, an **autoencoder** is a type of neural network used to compress and reconstruct data, often for tasks like data denoising or anomaly detection.

- It consists of three main parts:
  - **Encoder** → compresses the input data into a lower-dimensional representation → Two 1D Convolutional Layers.
  - **Bottleneck Layer** → holds the compressed representation of the input data and captures the most essential features.
  - **Decoder** → reconstructs the original data from its compressed form → Three Transposed 1D Convolutional Layers.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- A **1D Convolution** applies a moving filter along a one-dimensional sequence, capturing local patterns in temporal data.

- This network is trained to minimize the difference between the original input and the reconstructed output →
  - It learns to capture the essential features of the data → When an **autoencoder** encounters anomalous data, it is less effective at reconstructing it, leading to **higher reconstruction errors** which can be used to identify anomalies.

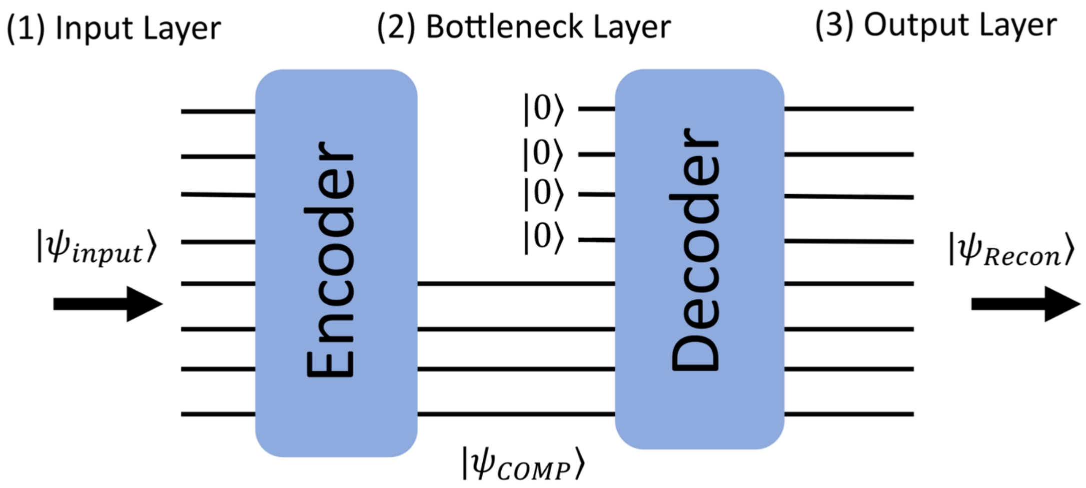
- We used the **Mean Squared Error (MSE)** as the loss function to minimize the difference between the original and reconstructed light curves.

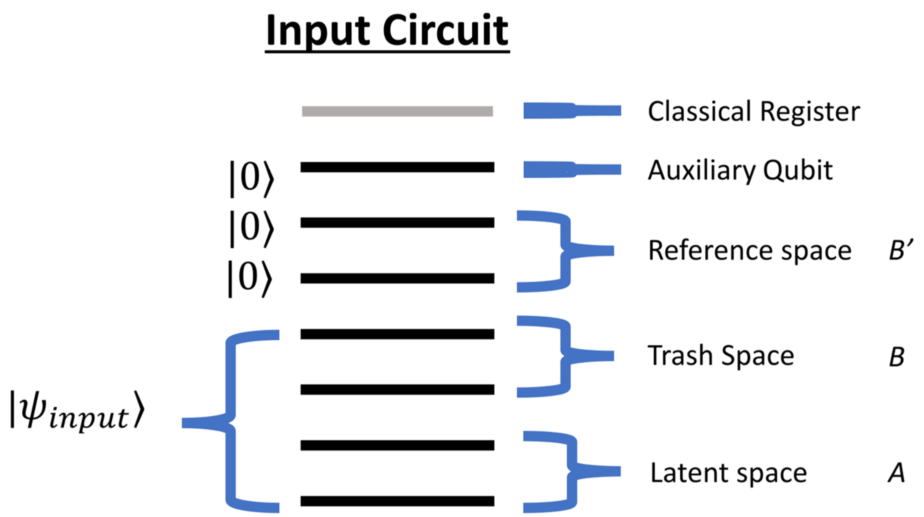
Quantum Autoencoder

- We implemented the quantum counterpart to the classical autoencoder that compresses quantum information into a reduced latent space using PQCs.
  (Romero et al., 2017)

- **Circuit structure:**
  - **Input Layer**: Original quantum state with **reference state** and **ancilla qubit**.
  - **Bottleneck Layer: Latent space** with *n* - *k* qubits → Compressed representation of the input state.
  - **Output Layer**: **Reconstructed quantum state** after adding back *k* qubits in the zero-state.

- **Training Process**:
  - **SWAP Test:** It is used to evaluate the **fidelity** between the input quantum state and the reconstructed state.
  - **Local Function**: Trained by pushing discarded qubits to the zero state, improving compression quality and training efficiency





Results & Comparison between Classical and Quantum

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

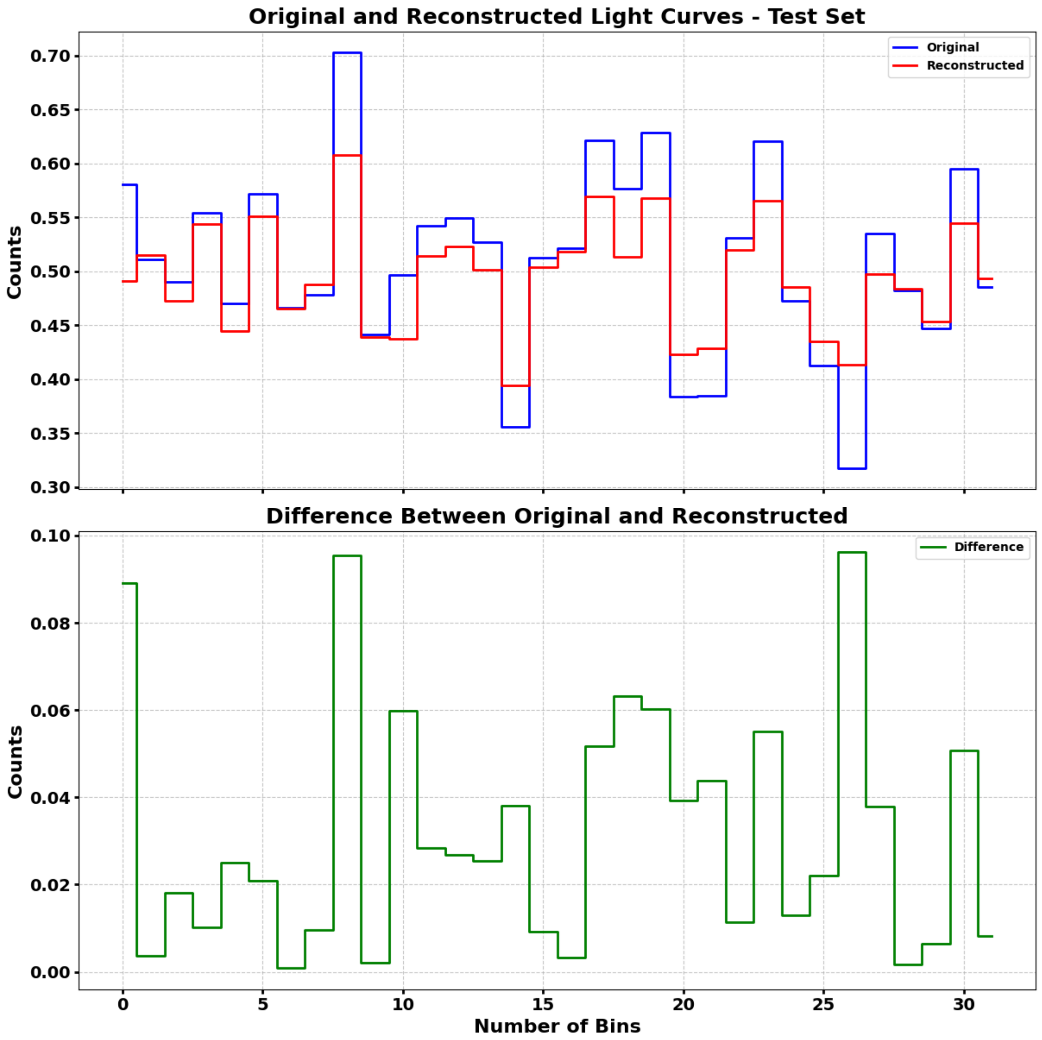
$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

$$\text{Light Curve Relative MAE \%} = \left(\frac{MAE_{lc}}{\mu_{lc}}\right) \times 100$$

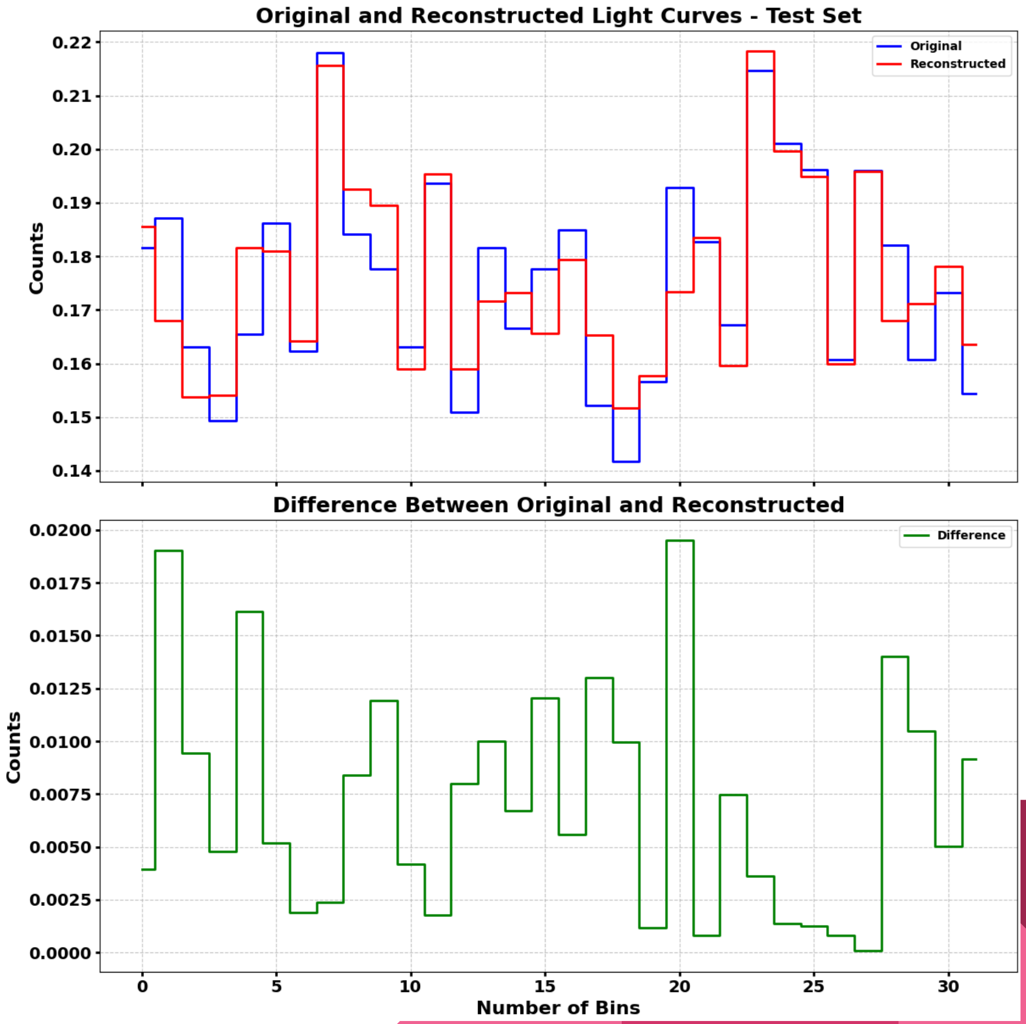
Dataset Configuration	Approach	Model Details	Mean Absolute Percentage Error
100 Samples - 32 Bins	Classical	6.5×10^5 Parameters	6.87 %
	Classical	705 Parameters	34.77 %
	Quantum	Real Amplitudes (10 parameters)	6.75%
1000 Samples - 32 Bins	Classical	6.5×10^5 Parameters	2.29 %
	Classical	705 Parameters	19.72 %
	Quantum	EfficientSU2 (20 parameters)	6.84 %
5000 Samples - 32 Bins	Classical	6.5×10^5 Parameters	0.82 %
	Classical	705 Parameters	9.90 %
	Quantum	EfficientSU2 (20 parameters)	6.73 %

Results & Comparison between Classical and Quantum

Classical - Reconstructed Curve

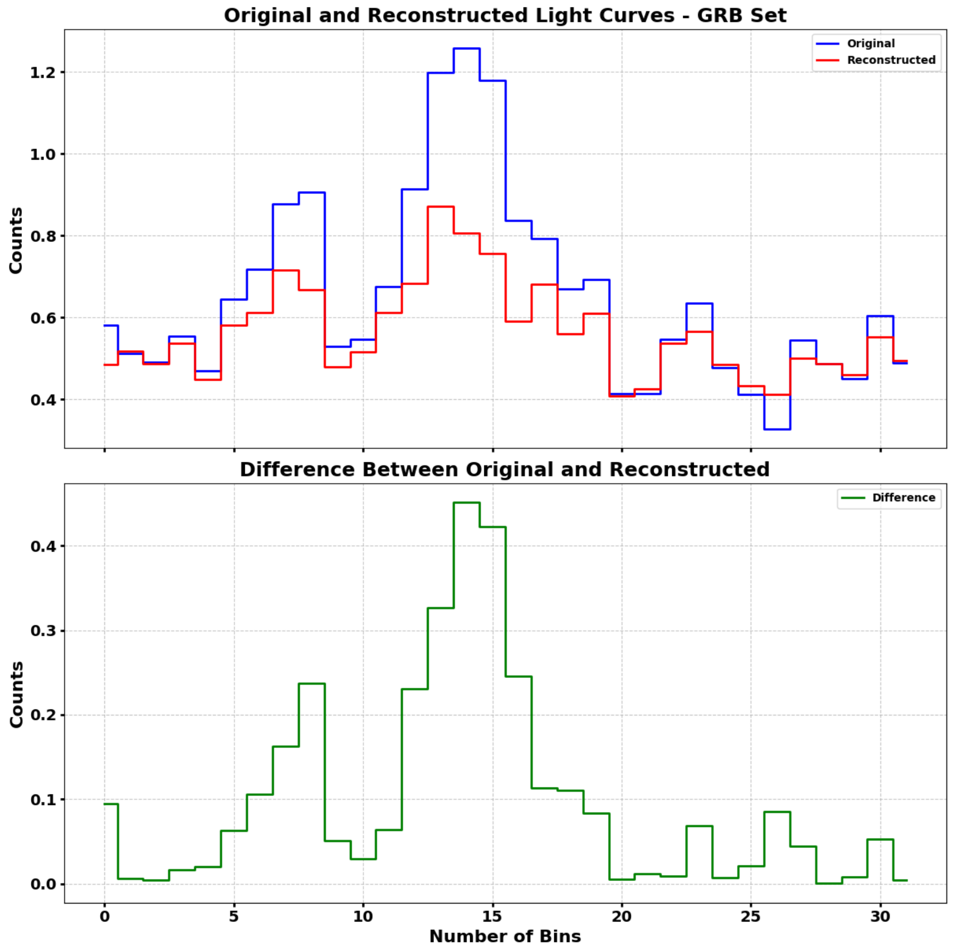


Quantum - Reconstructed Curve

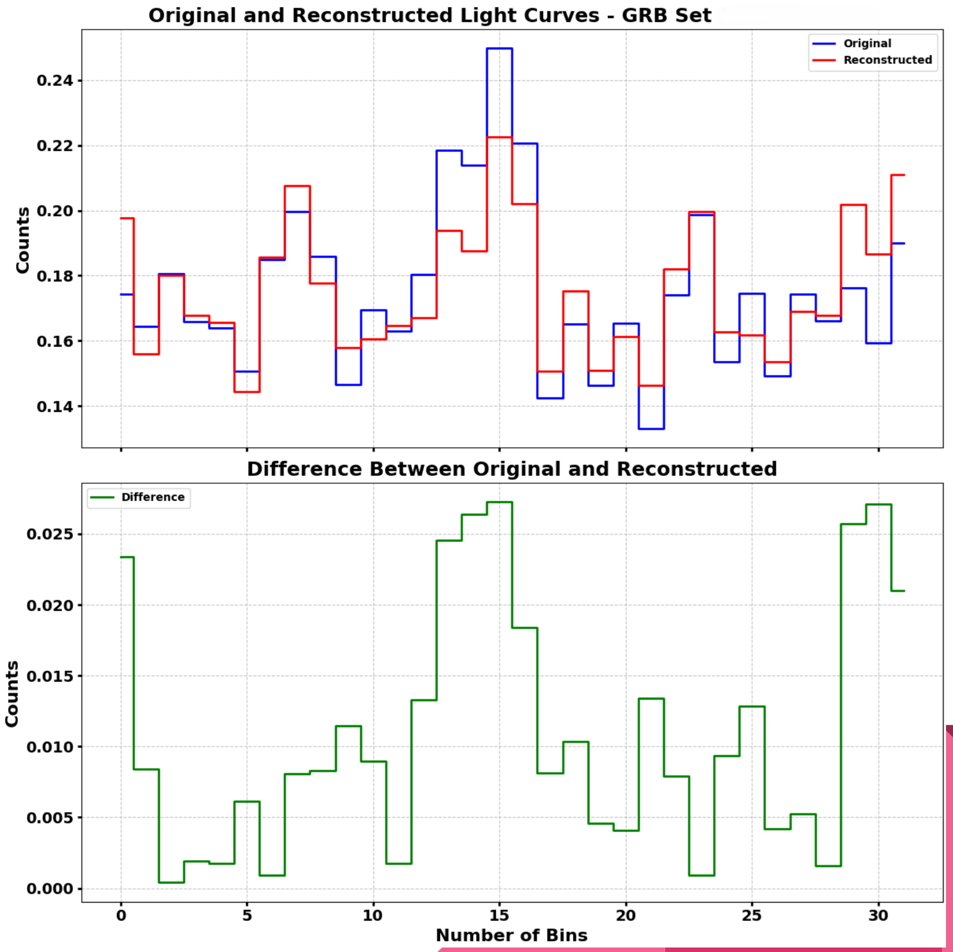


Results & Comparison between Classical and Quantum

Classical - Anomaly detected



Quantum - Anomaly detected



Conclusions

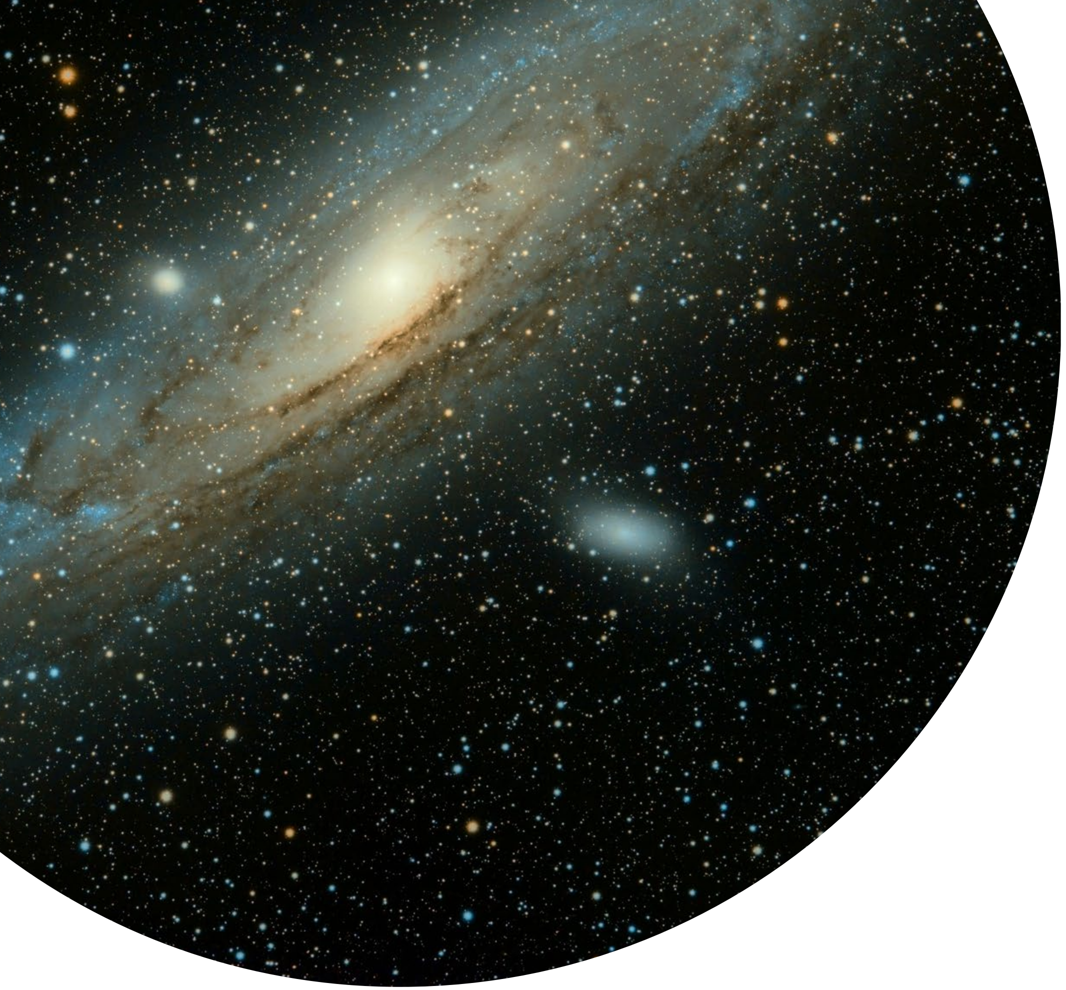
- The classical approach achieved **excellent** GRB anomaly detection with large data and parameters, but its performance significantly **drops under data scarcity or when limited to a few parameters.**

- **Quantum improvements**:

    - **Fewer Parameters**: Quantum models achieved strong results with significantly fewer parameters, where approximately ten parameters can achieve superior performance compared to over a thousand.

    - **Fewer Samples**: Quantum models show robust learning from smaller datasets, outperforming the classical approach.

    - **Good for resource-constrained tasks**: QML is promising for resource-constrained tasks, including scenarios with data scarcity or where less complex models are essential, especially as quantum hardware improves.



THANKS FOR YOUR ATTENTION

alessandro.rizzo@inaf.it
INAF - Osservatorio Astrofisico di Catania

Results & Comparison between Classical and Quantum

- With 100 training samples, the quantum autoencoder (10 parameters) achieved an MSE of 2.24 x 10^{-4}, outperforming the classical model (705 parameters) by two orders of magnitude.
- At full capacity with 5000 samples, the classical model achieves the lowest overall MSE of 2.97 x 10^{-5}, confirming its superiority with larger amount of data and parameters.

Dataset Configuration	Trainable Parameters	MSE - Training Set	MSE - Test Set
1000 Samples - 32 bins	6.5×10^5	1.65×10^{-4}	1.76×10^{-4}
	4×10^4	5.30×10^{-4}	5.65×10^{-4}
	1081	9.70×10^{-3}	9.78×10^{-3}
	705	1.32×10^{-2}	1.33×10^{-2}
5000 Samples - 32 bins	6.5×10^5	2.65×10^{-5}	2.97×10^{-5}
	4×10^4	1.51×10^{-4}	1.58×10^{-4}
	1081	5.42×10^{-3}	5.50×10^{-3}
	705	4.07×10^{-3}	4.11×10^{-3}

Model	5q, 1L	7q, 1L	5q, 3L	7q, 3L
Efficient SU2	20	28	40	56
Real Amplitudes	10	14	20	28

Dataset Configuration	Qubits	Layers	Ansatz	MSE - Training Set	MSE - Test Set
100 Samples - 32 Bins	4+1	1	Real Amplitudes	2.17×10^{-4}	2.24×10^{-4}
	4+1	1	Efficient SU2	2.41×10^{-4}	2.34×10^{-4}
	4+1	3	Real Amplitudes	6.49×10^{-4}	1.09×10^{-2}
1000 Samples - 32 Bins	4+1	1	Real Amplitudes	2.24×10^{-4}	2.21×10^{-4}
	4+1	1	Efficient SU2	2.28×10^{-4}	2.27×10^{-4}
	4+1	1	Real Amplitudes	7.88×10^{-4}	1.41×10^{-2}
5000 Samples - 32 Bins	4+1	1	Real Amplitudes	3.31×10^{-4}	3.32×10^{-4}
	4+1	1	Efficient SU2	2.32×10^{-4}	2.32×10^{-4}