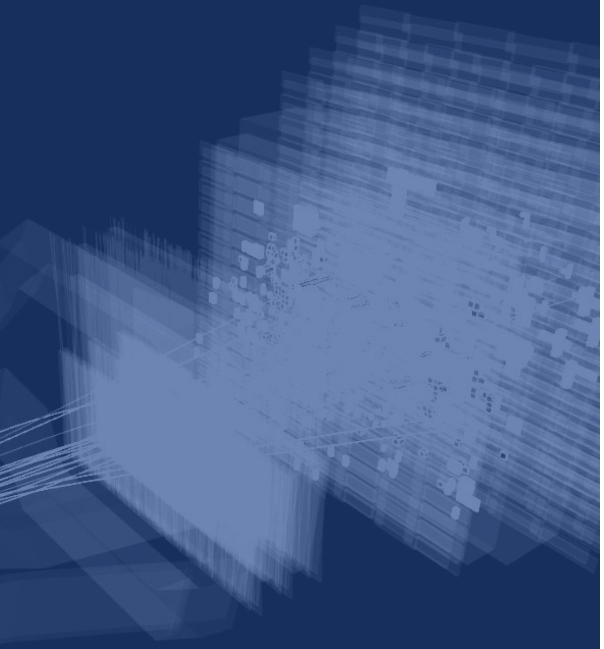
# First implementation of Quantum Machine Learning for b-jet tagging at LHCb

**IEEE Quantum Week 2021 Workshop: Quantum Computing for High Energy Physics**
**21 October 2021 (online)**

Alessio Gianelle, Davide Nicotra, Donatella Lucchesi, Lorenzo Sestini, Davide Zuliani*

*for questions and comments: davide.zuliani@cern.ch



# Outline

- Introduction
- QML for b-jet tagging
- Results and future perspectives

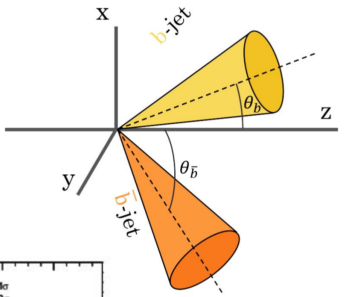
# Introduction

# Physics cases

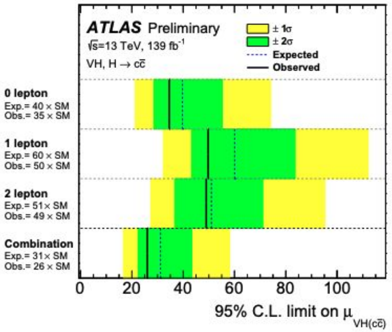
Jet flavor identification is mandatory for several physics cases

- **b-b̄ charge asymmetry**, interesting for **New Physics** searches

$$A_C^{b\bar{b}} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)} \qquad \Delta y = y_b - y_{\bar{b}}$$




- **identification** of **Higgs boson** decaying to:
  - bb̄ di-jets (observed @ ATLAS & CMS)
  - cc̄ di-jets (not yet observed)

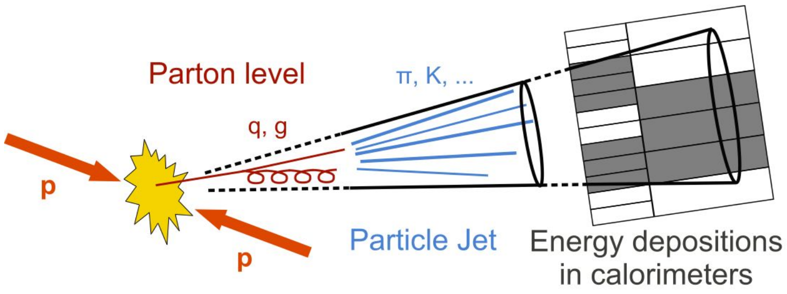


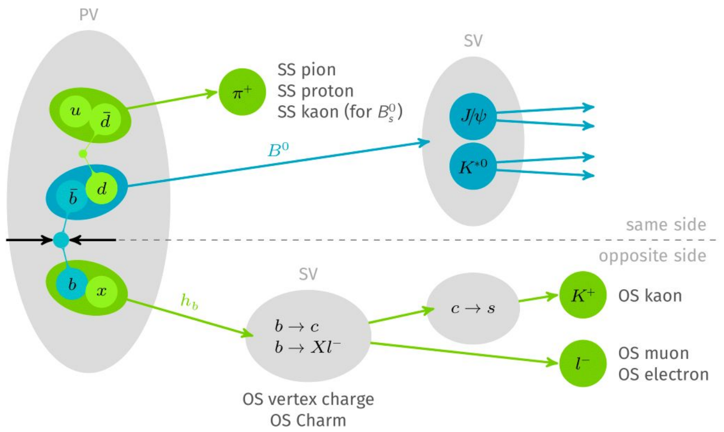

ATLAS-CONF-2021-021

- Final states detected by the experiment → **jets**

# b-jet tagging @ LHCb

@ LHC is **fundamental** to identify the **flavour** of the quark originating the jet → **jet tagging**



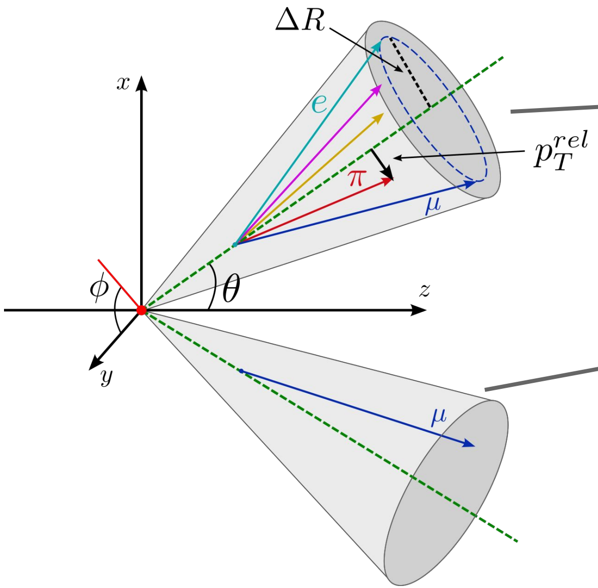

*"Jets are streams of particles produced by QCD processes in proton-proton collisions"*

In particular @ LHCb we are interested in studying jets generated by b and b̄ quarks

**b-jet tagging**

# b-jet (classical) tagging methods

There are two possible approaches to achieve b-jet tagging: **exclusive** and **inclusive** algorithms




**inclusive algorithms** use information from the **whole jet substructure**

Machine Learning (ML) algorithms
such as **Deep Neural Networks** (DNN)

**exclusive algorithms** use information from **a specific process** to infer the quark flavour

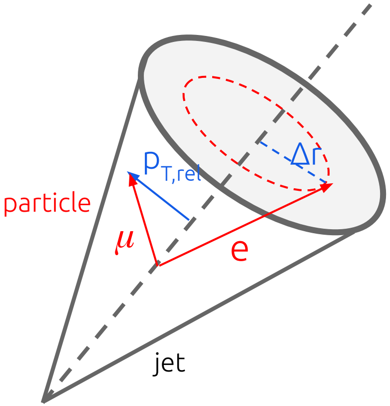
**muon tagging**: a $\mu$ coming from the b (or b-bar) semileptonic decay (P =10%) is used to tag the jet ($\mu$ and b charges are correlated)

# QML for b-jet tagging

LHCb Open Data

# Dataset

**LHCb simulation** of di-jet generated by b and b̄ quarks @ $E_{cm}$ = 13 TeV (Run 2 condition)
~ 700.000 jets (60% training, 40% testing & evaluation)




Inside each jet we consider 5 types of particles

**muon** **electron** **pion** **kaon** **proton**

and for each type we select 3 variables:

- $p_{T,rel}$: transverse momentum relative to the jet axis
- $\Delta r$: distance relative to jet axis
- q: charge of the particle

\+ 1 global variable → total jet charge

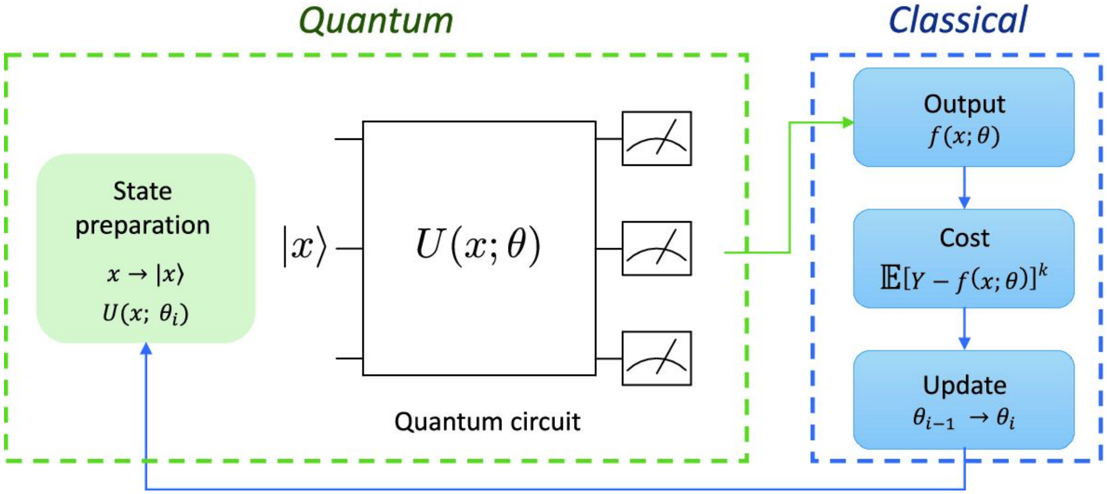
$\left.\right\}$ **16 input variables**

two datasets have been considered in this study:

- **"muon dataset":** only muon features + total jet charge (4 features)
- **"complete dataset":** all 16 features

# "New" approach: going to quantum

classification problem = **Variational Quantum Classifier**

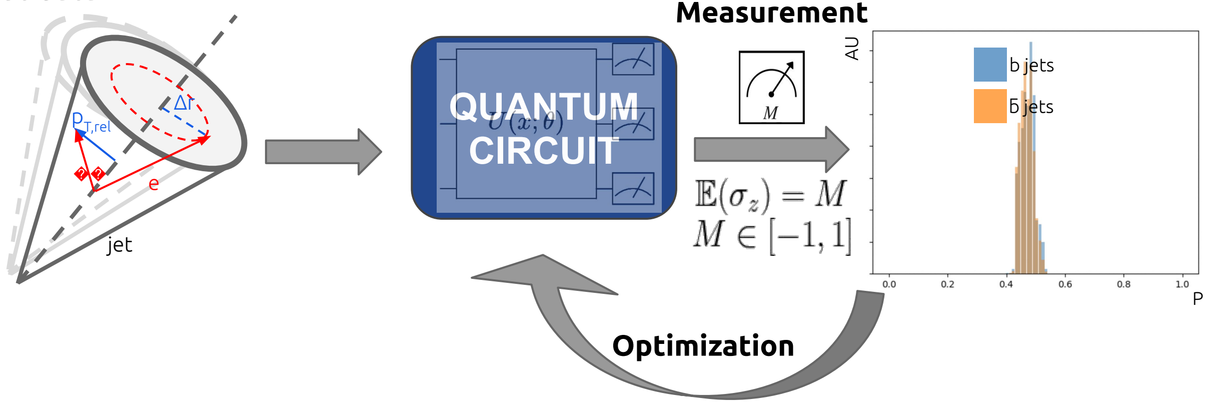



- data are fed into variational quantum circuit
- measurements of qubits are mapped to probabilities for labels
- probabilities are used to estimate a cost function which is optimized through a classical optimizer

- several quantum circuits geometries have been studied (more in the backup)
- quantum circuits are simulated with **noiseless simulators**, using PennyLane libraries

# Training procedure

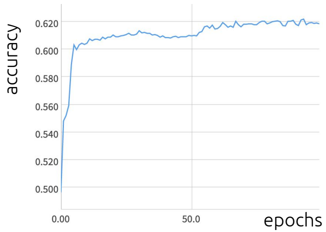
**Input data**

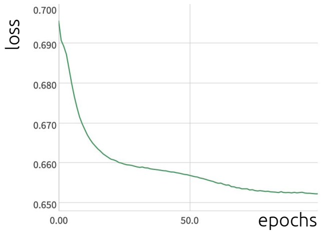



**Tagging probabilities**

$$\boxed{P_b = \frac{M+1}{2}}$$

$$\boxed{P_{\bar{b}} = 1 - P_b}$$

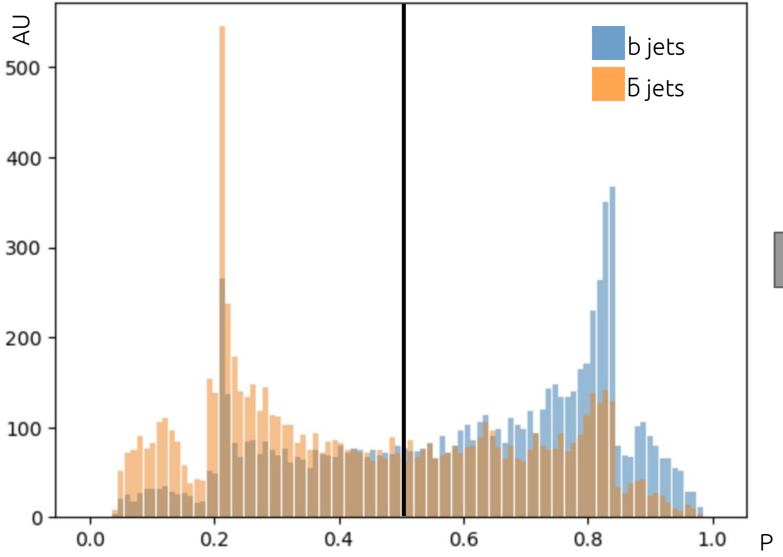



- cross-entropy $CE(\theta_i) = -\frac{1}{N} \sum_{dataset} \sum_{i \in \{b, \bar{b}\}} p_i \log q_i$
- ADAM optimizer

# Probability distribution and tagging power

The figure of merit for this task is the **tagging power**




$$\varepsilon_{tag} = \varepsilon_{eff} \cdot (1 - 2\omega)^2$$

$\dfrac{N_{tagged}}{N_{tot}}$

***"efficiency"***

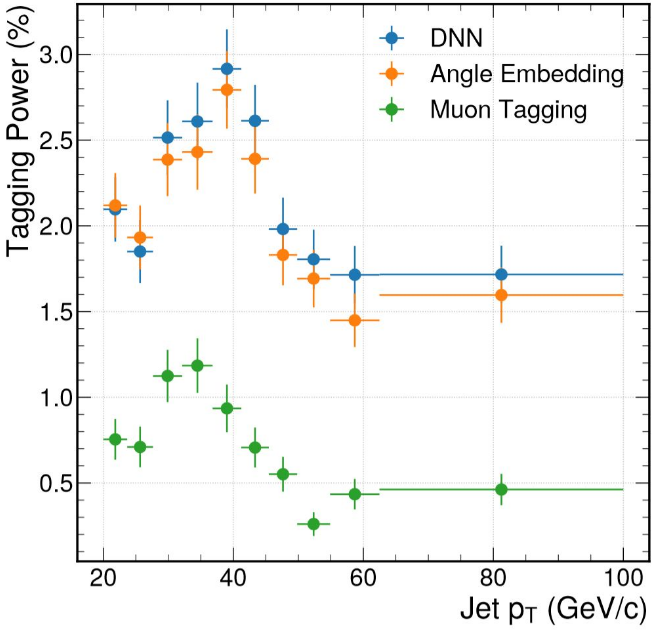
$\dfrac{N_{wrong}}{N_{tagged}}$

***"mistag"***

in the following **no optimization** of the tagging power has been performed

# Results and future perspectives

# Tagging power on the muon dataset



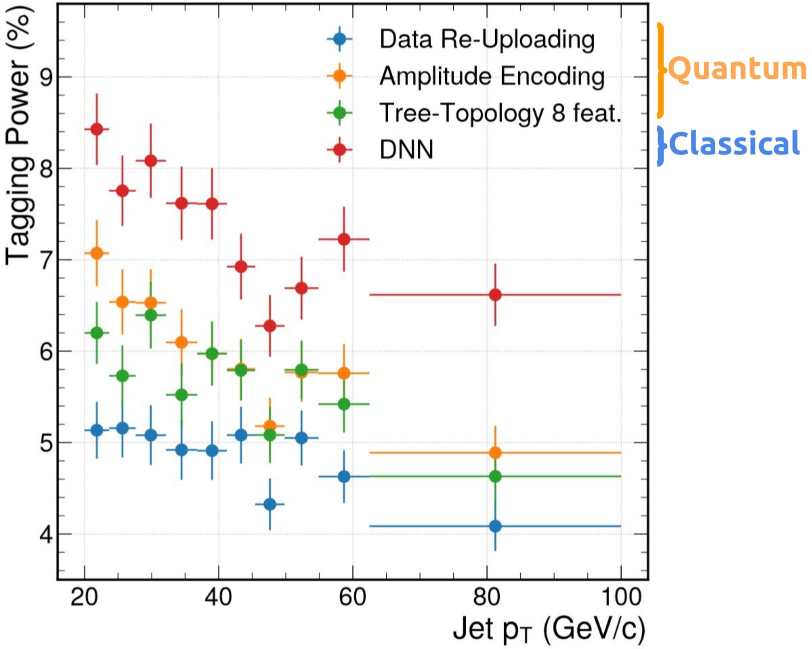

} **Classical**
} **Quantum**

- DNN and QML are trained and tested on the same number of events
- **Angle Embedding** is used as quantum circuit

| Model | $N_{feat.}$ | $N_{par}$ |
|---|---|---|
| Angle Embedd. | 4 | 48 |
| DNN | 4 | 181 931 |
| Muon Tagging | 1 | N.D. |

Classical ML and Quantum ML show **equal performance within the uncertainty**

# Tagging power on the complete dataset




- QML algorithms are trained on 1/50th of events w.r.t. DNN
- testing is performed on the same number of events

| Model | $N_{feat.}$ | $N_{par}$ |
|---|---|---|
| Amplitude Embedd. | 16 | 72 |
| Data re-uploading | 16 | 216 |
| Tree-topology | 8 | 104 |
| DNN | 16 | 163 307 |

The DNN shows better performance with respect to all quantum models (1-2 % better)

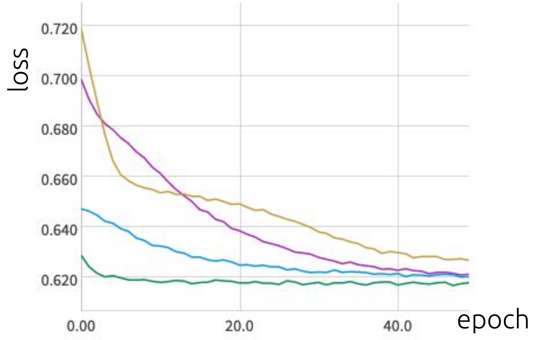
# Tests on noisy simulators

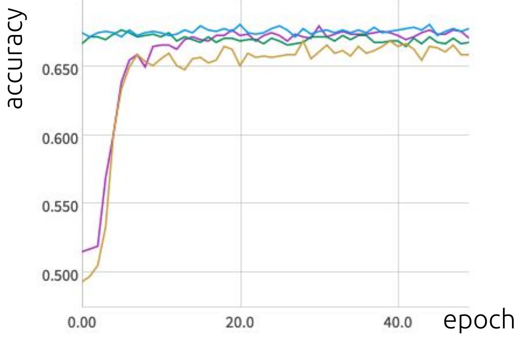
**future perspectives**

- Up to now everything was done using **noiseless** simulators
- `qiskit.aer` allows to simulate noise of a specific device:

```
noise_model = NoiseModel.from_backend(provider.get_backend('ibmq_'+city))
```

- computational cost is demanding
- up to now training on 1000 events with muon (a.k.a. “muon dataset”)
- took ~ 30 h to train and validate on 5 CPU cores for 50 epochs




**no clear difference when considering noise!**

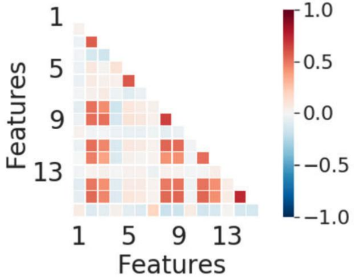
going to more qubits and different geometries

# Search for correlations

future perspectives

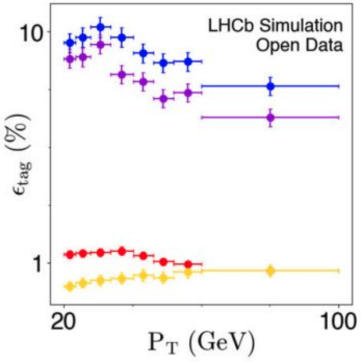
the main idea is to exploit the quantum properties of the circuits to measure correlations between qubits (see https://doi.org/10.1038/s41534-021-00443-w)

$$C^{l}_{i,j} = \frac{\langle \psi_l | \sigma^z_i \sigma^z_j | \psi_l \rangle}{\langle \psi_l | \psi_l \rangle}$$




- understand which features are important for the classification
- select only the relevant features and check classification performance

- possibly have a better understanding of jet physics!



# Conclusions

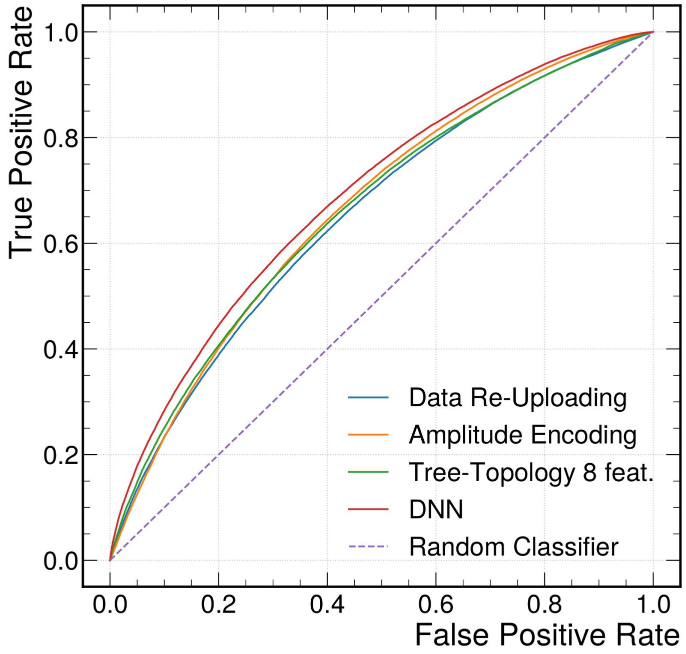
- Jet flavor identification is mandatory for several physics cases:
  - charge asymmetry for b- and c-jets
  - identification of Higgs boson decaying to b and c di-jets
- LHCb could play an interesting role in studying jet tagging with QML
- Results obtained so far show that **it's possible to apply QML to jet tagging**:
  - for the muon dataset **performance between DNN and QML are comparable**
  - for the complete dataset the DNN performs slightly better than QML

- Still room for improvement and for **new studies**:
  - noisy simulation
  - correlations between variables
  - (hardware application)

# Thank you for your attention!

# Backup slides

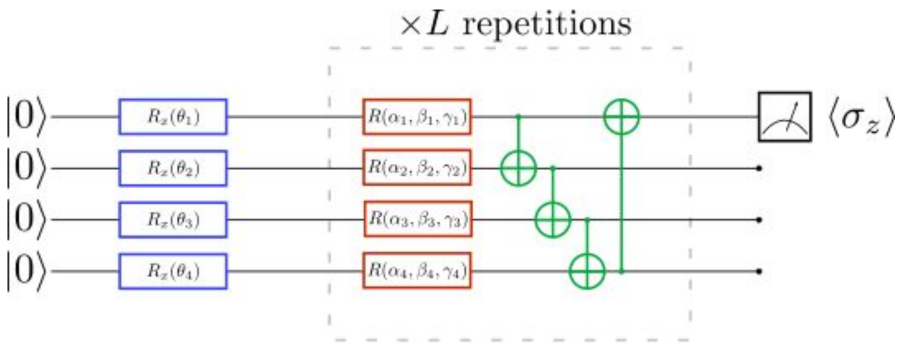
# ROC curves for complete dataset




| Dataset | Model | AUC |
|---|---|---|
| muon dataset | DNN | 0.75 |
| | Angle embedding | 0.75 |
| complete dataset | DNN | 0.69 |
| | Amplitude encoding | 0.67 |
| | Data re-uploading | 0.65 |
| | Tree-Topology (8 features) | 0.66 |

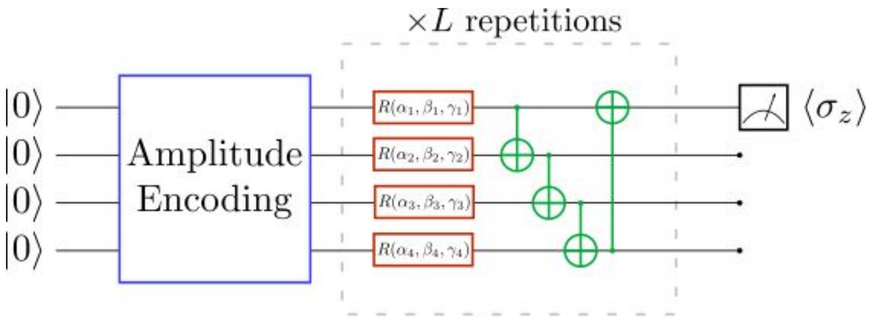
- for the muon dataset, DNN and Angle embedding performance are comparable
- for the complete dataset the DNN performs slightly better than the Amplitude encoding

# Circuit geometries



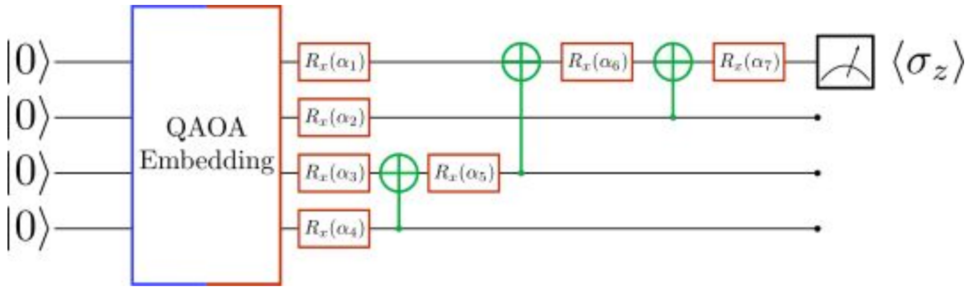

Angle Embedding

Amplitude Encoding

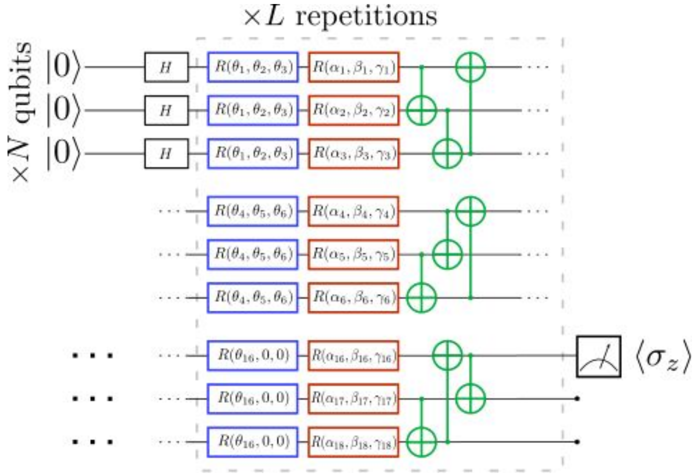


# Circuit geometries

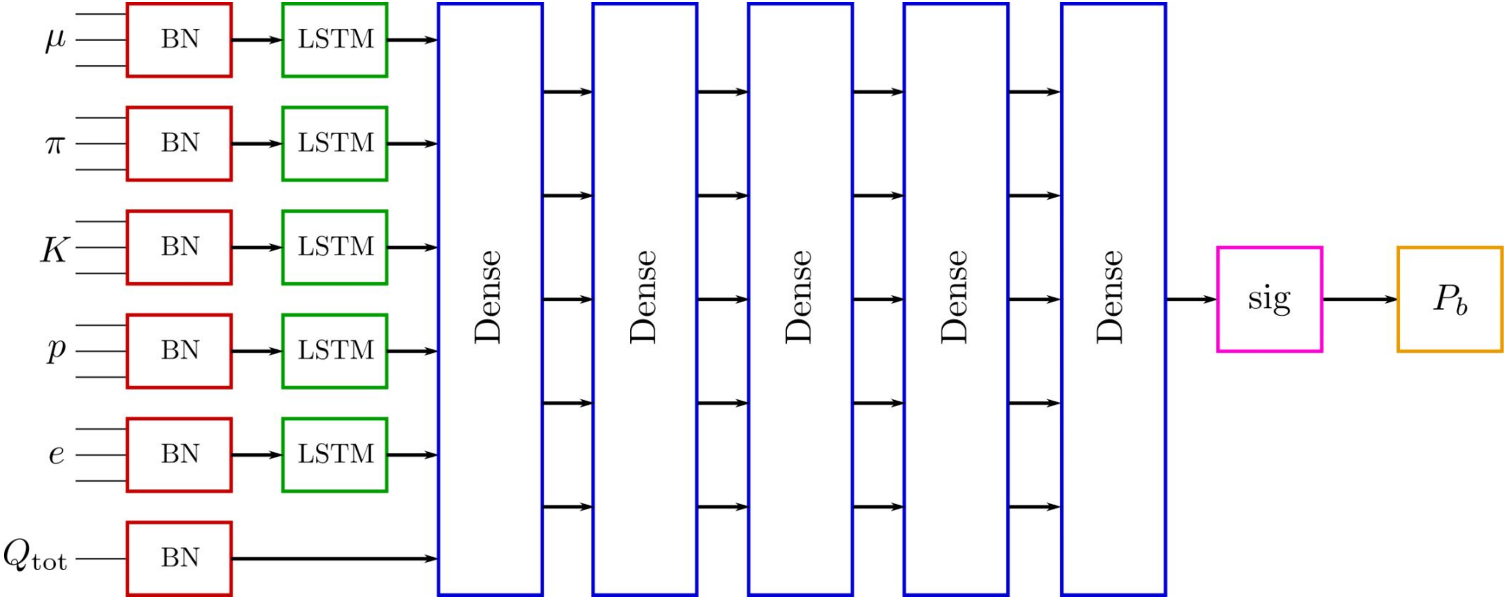
VQC with “Tree-Topology”




Data re-uploading

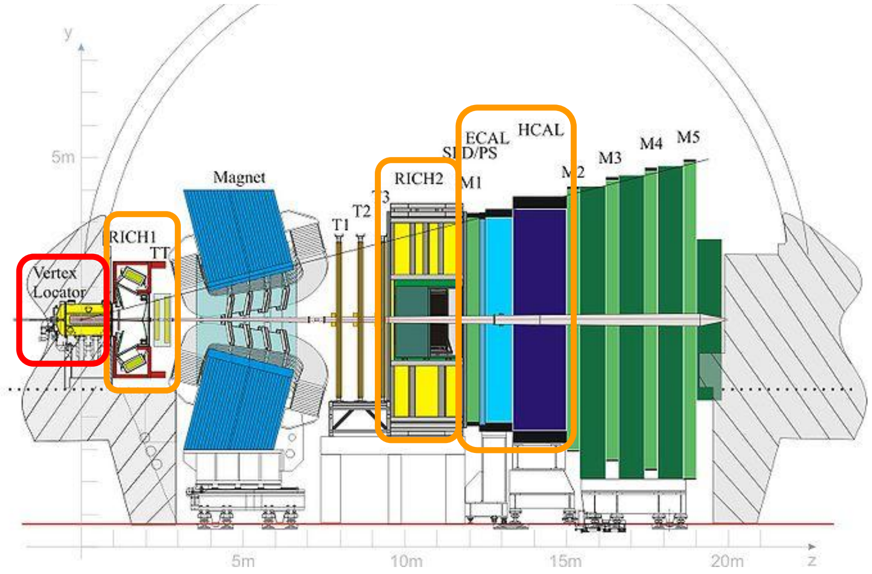


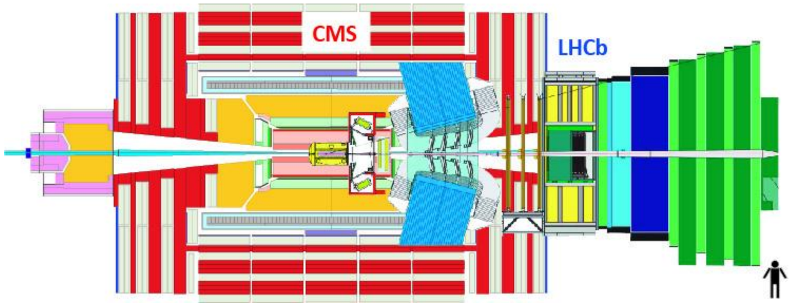
# DNN structure



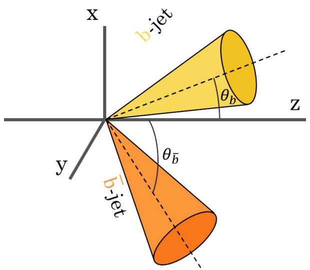
# LHCb

LHCb is a **forward spectrometer** designed to study **flavour physics**



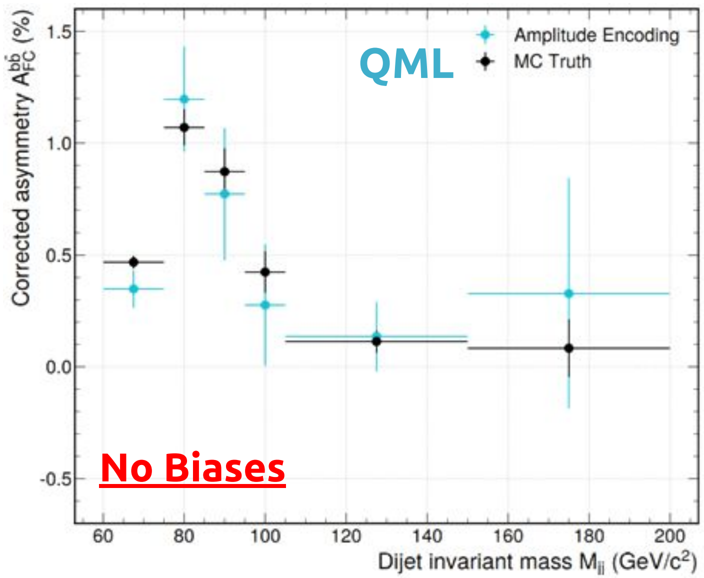

- complementary phase-space region w.r.t. ATLAS & CMS
- excellent vertex reconstruction
- excellent Particle Identification (PID)

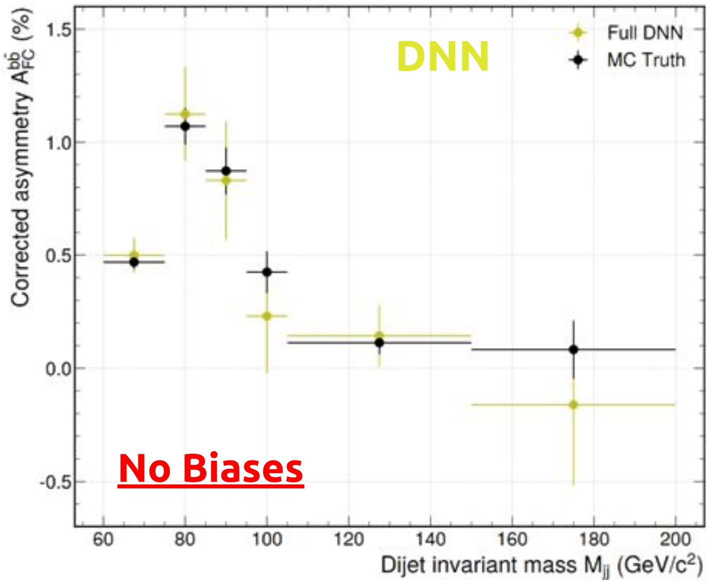



# Forward-central asymmetry estimation

$$A^{FC}_{b\bar{b}} = \frac{N(\Delta|y|>0) - N(\Delta|y|<0)}{N(\Delta|y|>0) + N(\Delta|y|<0)}$$

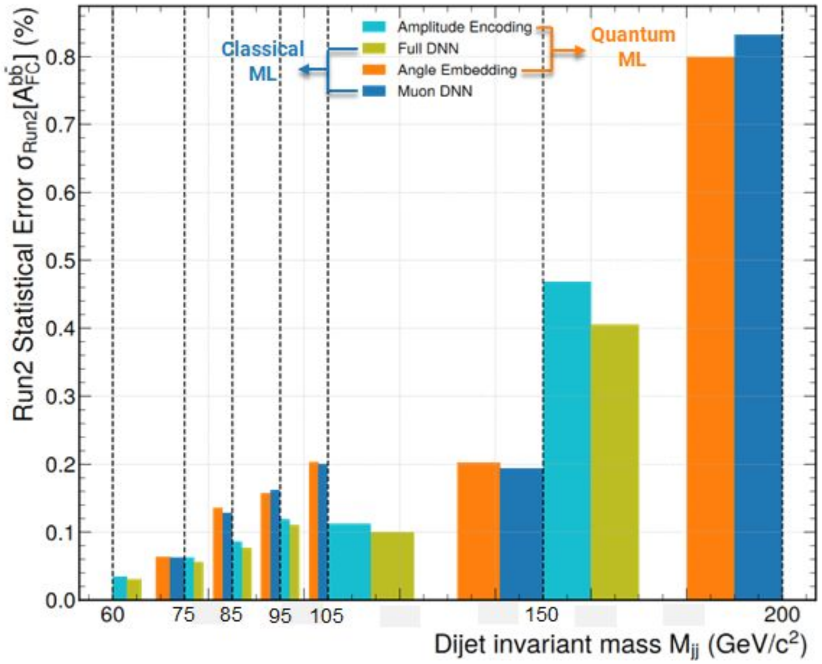
$$\Delta|y| = |y_b| - |y_{\bar{b}}|$$



# Run 2 statistical uncertainty on the FC asymmetry

$\sigma[A_{b\bar{b}}^{FC}]$




$$\sigma[A_{b\bar{b}}^{FC}] \propto \frac{1}{\sqrt{\epsilon_{tag}}}$$

on the complete dataset the classical DNN shows better performance w.r.t. QML