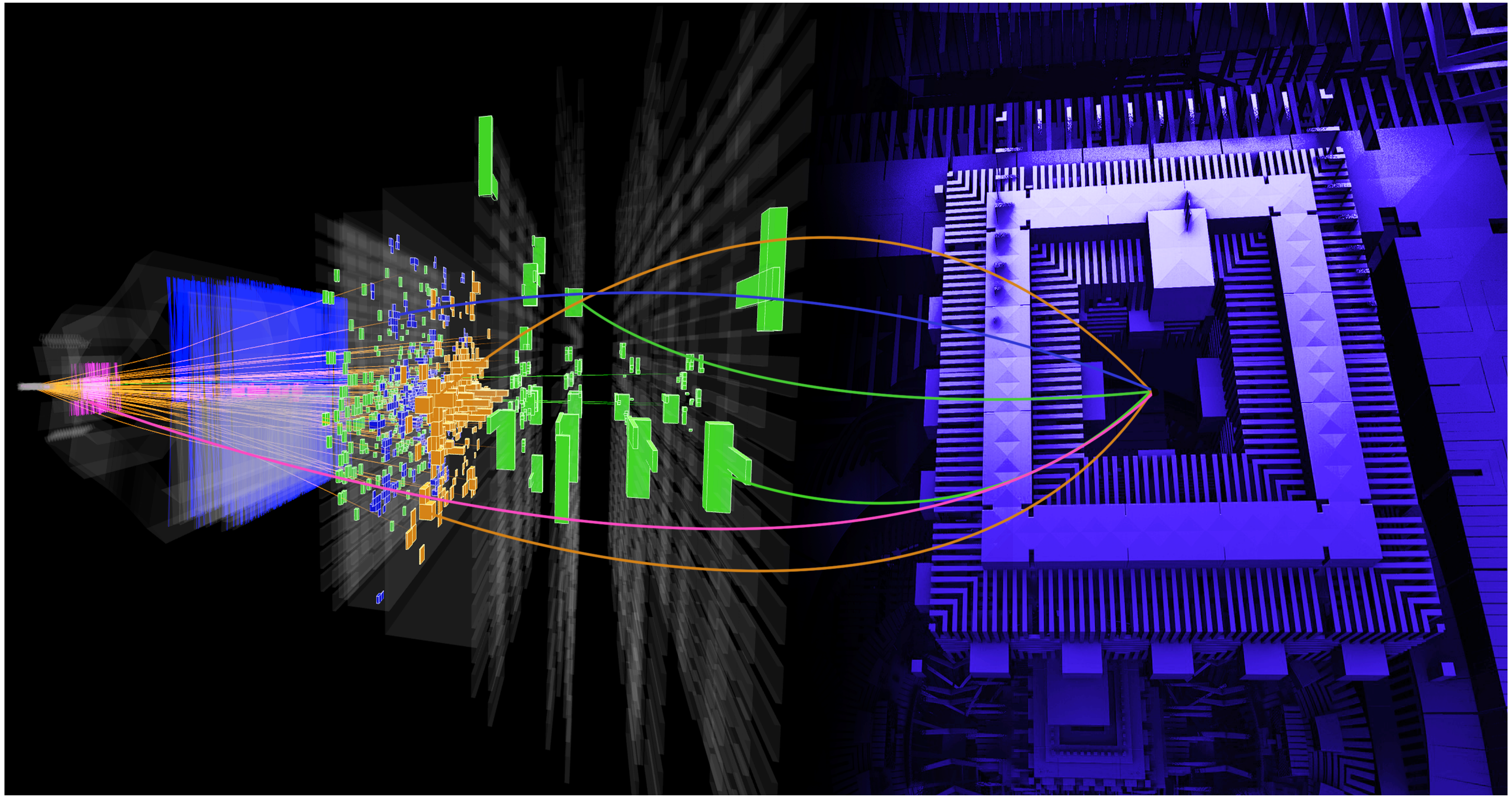
Jet identification with quantum machine learning at LHCb

Alessio Gianelle[a], Donatella Lucchesi[a,b], Lorenzo Sestini[a], Davide Zuliani[a,b]
[a]INFN Padova
[b]University of Padova

Quantum Computing @INFN workshop - Bologna - 15/11/2022

Quantum computing at LHCb



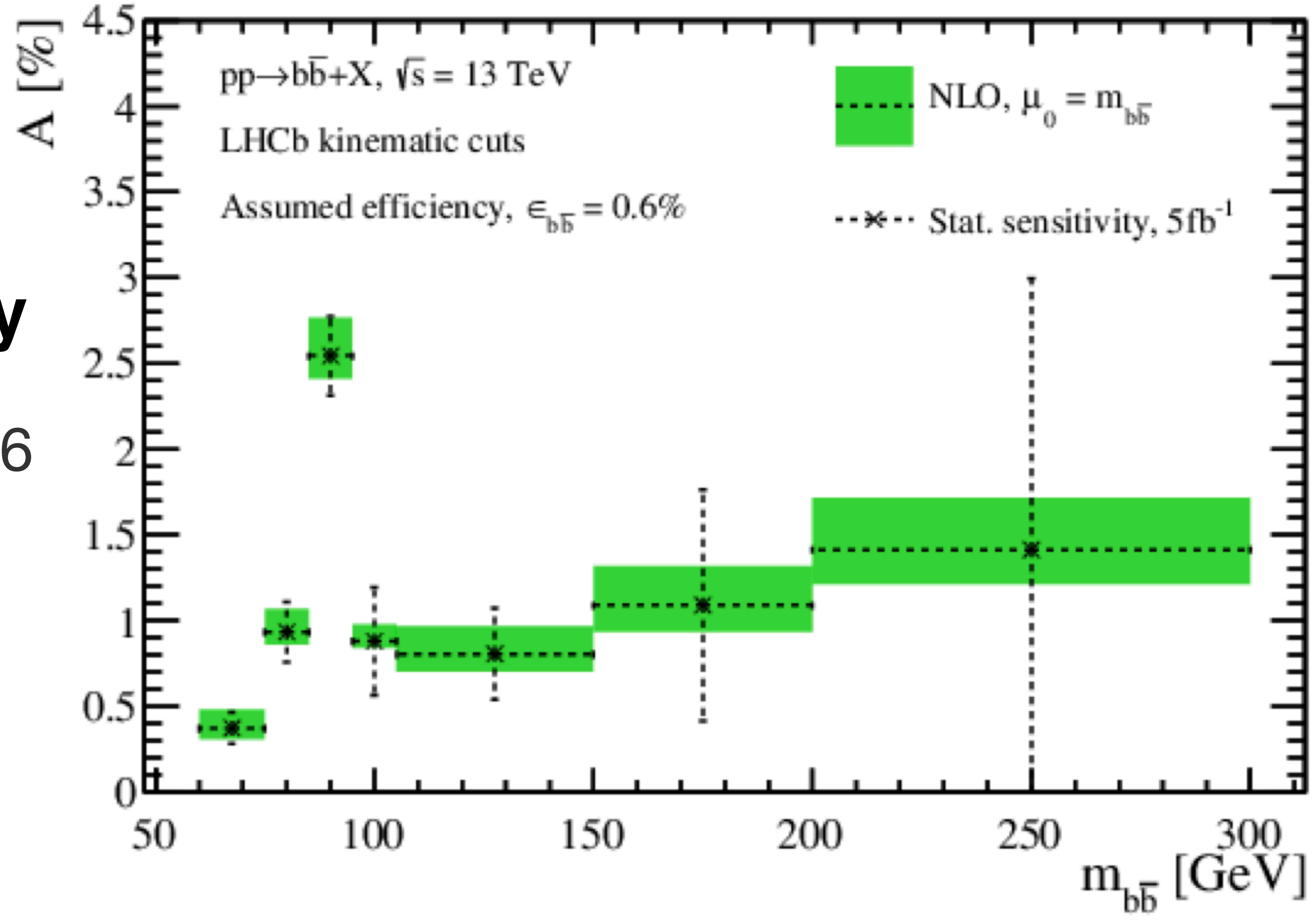
- Up to now two main activities in the collaboration:
  - **Jet identification**
  - **Tracking**
- **One LHCb paper published**: *"Quantum Machine Learning for b-jet charge at LHCb"* JHEP08(2022)014
- A dedicated work package in the **Data Processing & Analysis (DPA) project**: https://lhcb-dpa.web.cern.ch/lhcb-dpa/

Jet identification

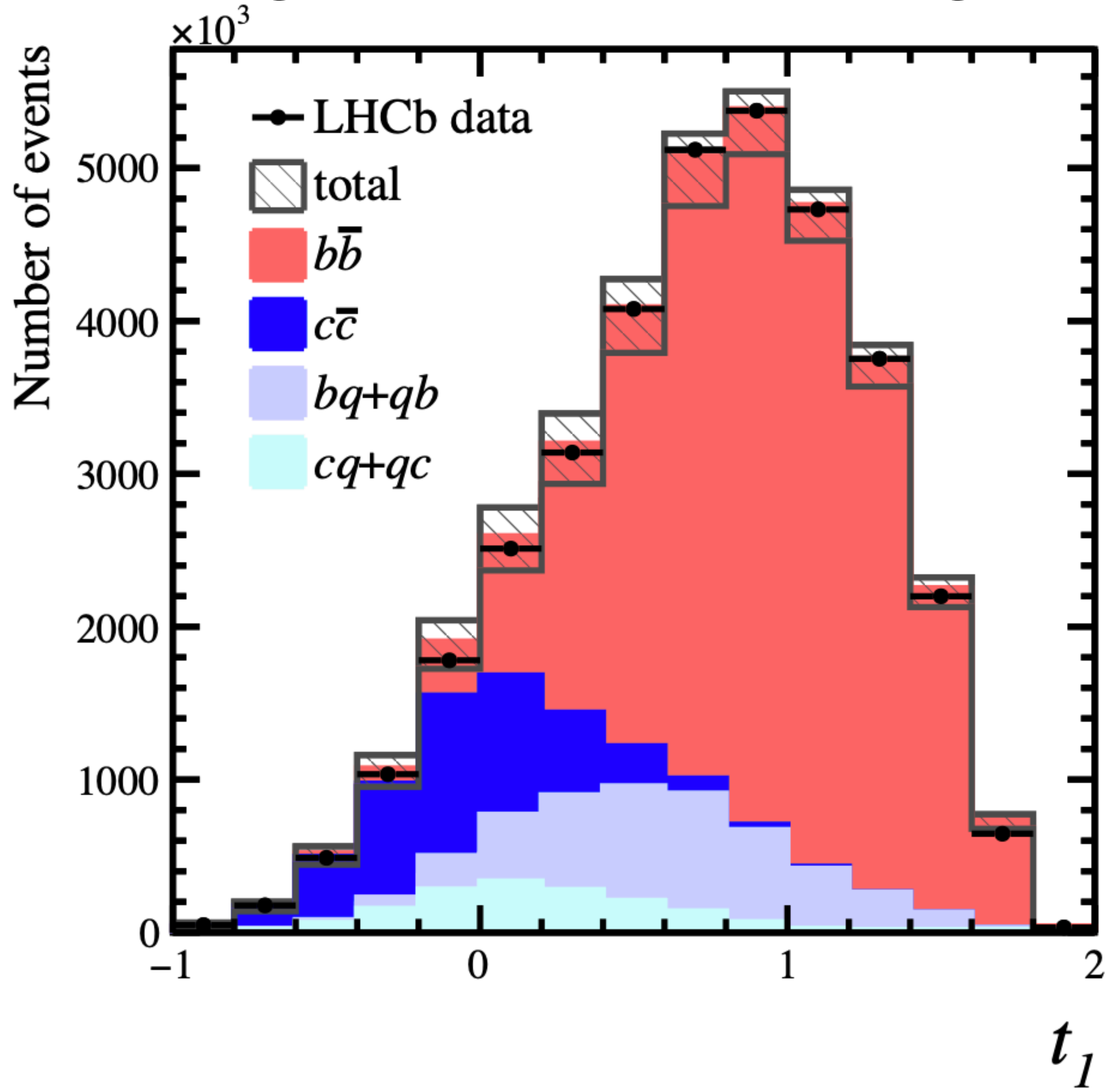
- Jets: spray of particles produced by quark hadronization and fragmentation
- Classification problem that involves many particles/features
- **LHC experiments heavily use machine learning** to improve the performance on jet identification
- **Many use cases at LHCb!**

Measure the $b\bar{b}$ asymmetry

JHEP03(2019)166



Disentangle $b\bar{b}$, $c\bar{c}$ and backgrounds



JHEP02(2021)023

Quantum machine learning

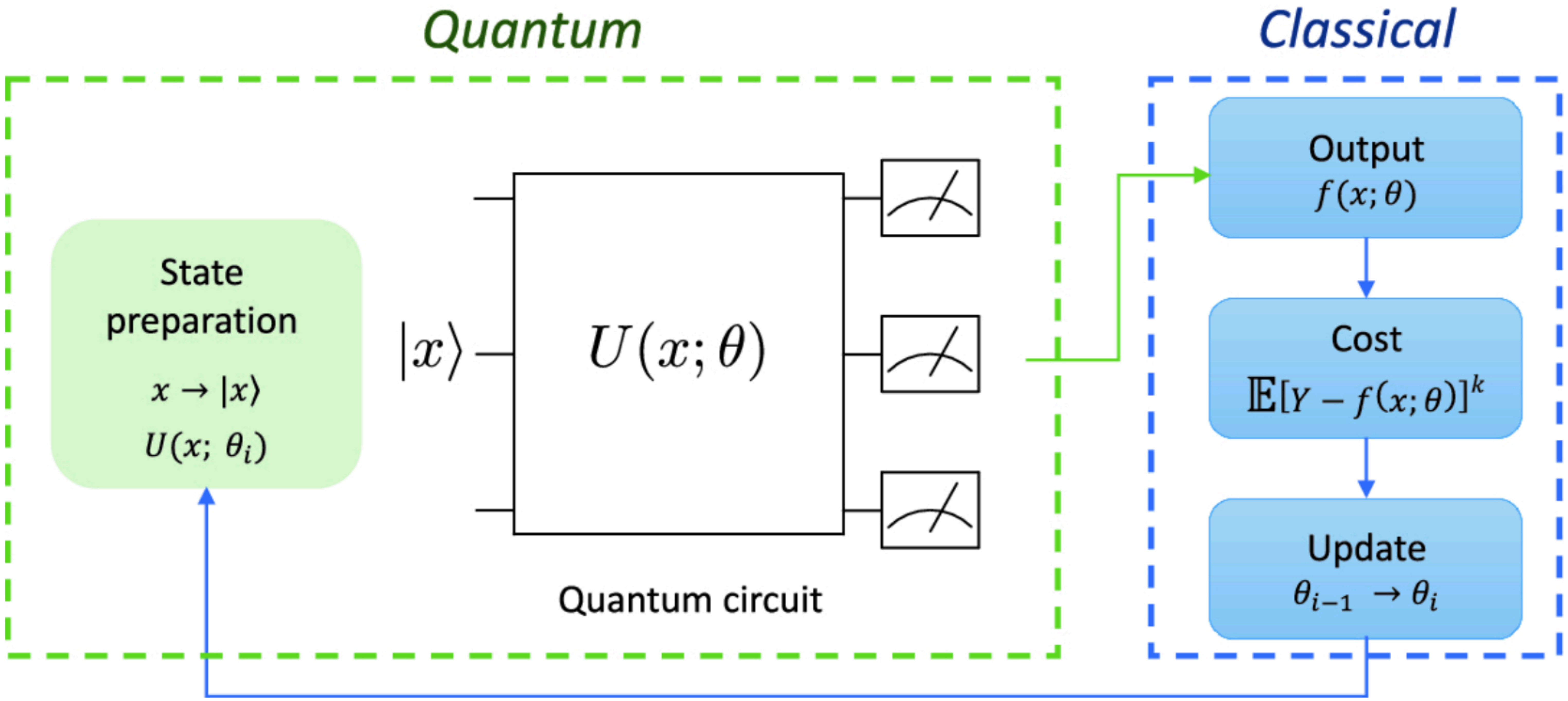
- Jet identification is the ideal task to test Quantum Machine Learning algorithms
- In all our studies we use the **full simulation provided by the LHCb experiments**: our samples resemble the real data
- We employ the **hybrid approach**: the quantum circuit with tunable parameters (Variational Quantum Circuit) is trained by using a classical loss function

Quantum Circuits implemented with Pennylane/Qiskit

In particular **Qiskit** is used for tests on IBM hardwares

PENNYLANE

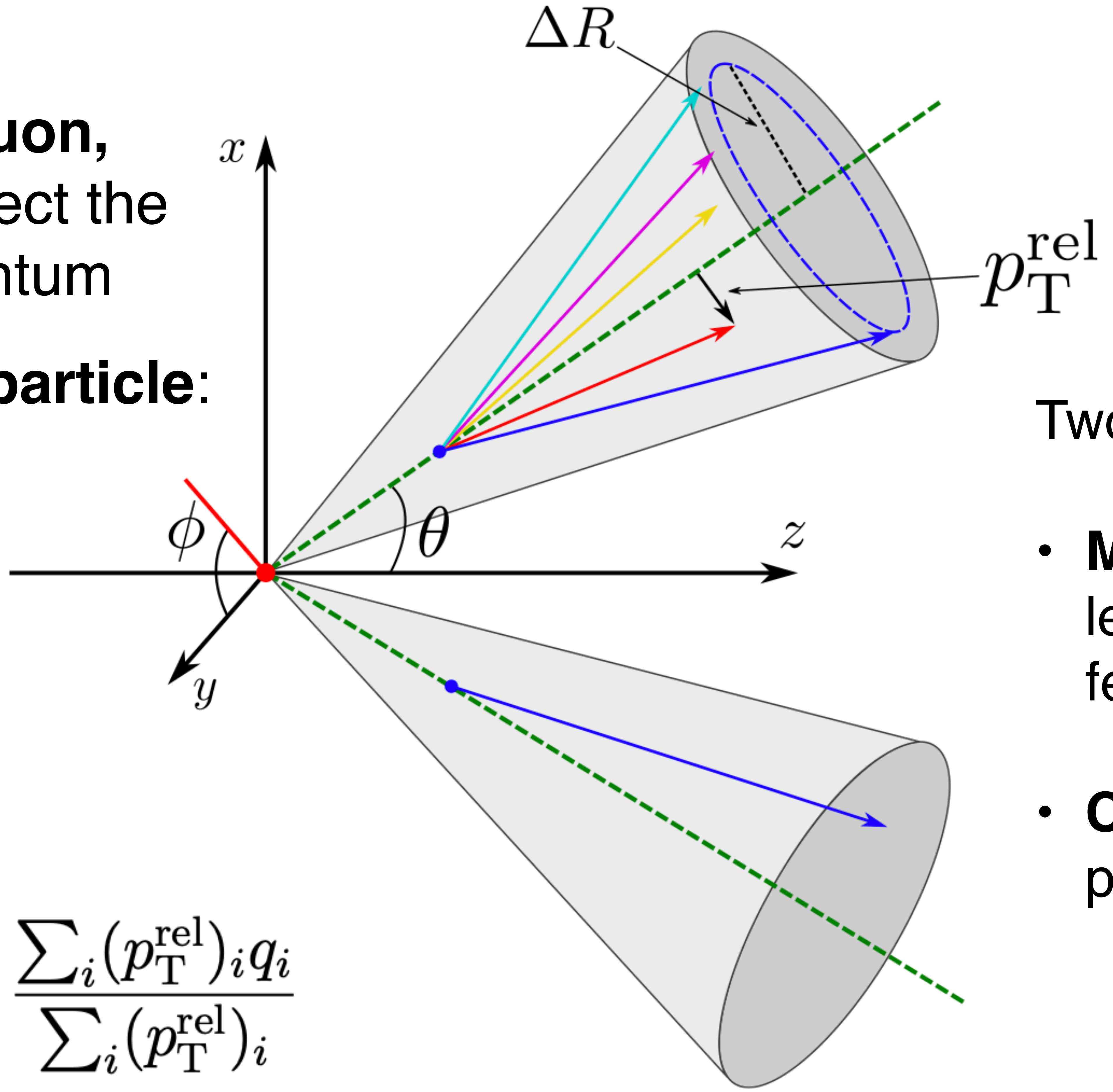
Qiskit



b-jets vs $\bar{b}$-jets: features

- We take profit of the **Particle Identification** capabilities of LHCb
- For each identified type of particle (**muon, electron, kaon, pion, proton**) we select the one with the higher transverse momentum
- We consider **three observables per particle**:
  - ΔR (distance in η-ϕ space) between the particle momentum and the jet axis
  - p_T^{rel} with respect to jet axis
  - Charge (+1 or -1)
- We include also the jet charge: $Q = \frac{\sum_i (p_T^{rel})_i q_i}{\sum_i (p_T^{rel})_i}$

A total of 16 features are considered to distinguish jets produced by b and $\bar{b}$ quarks

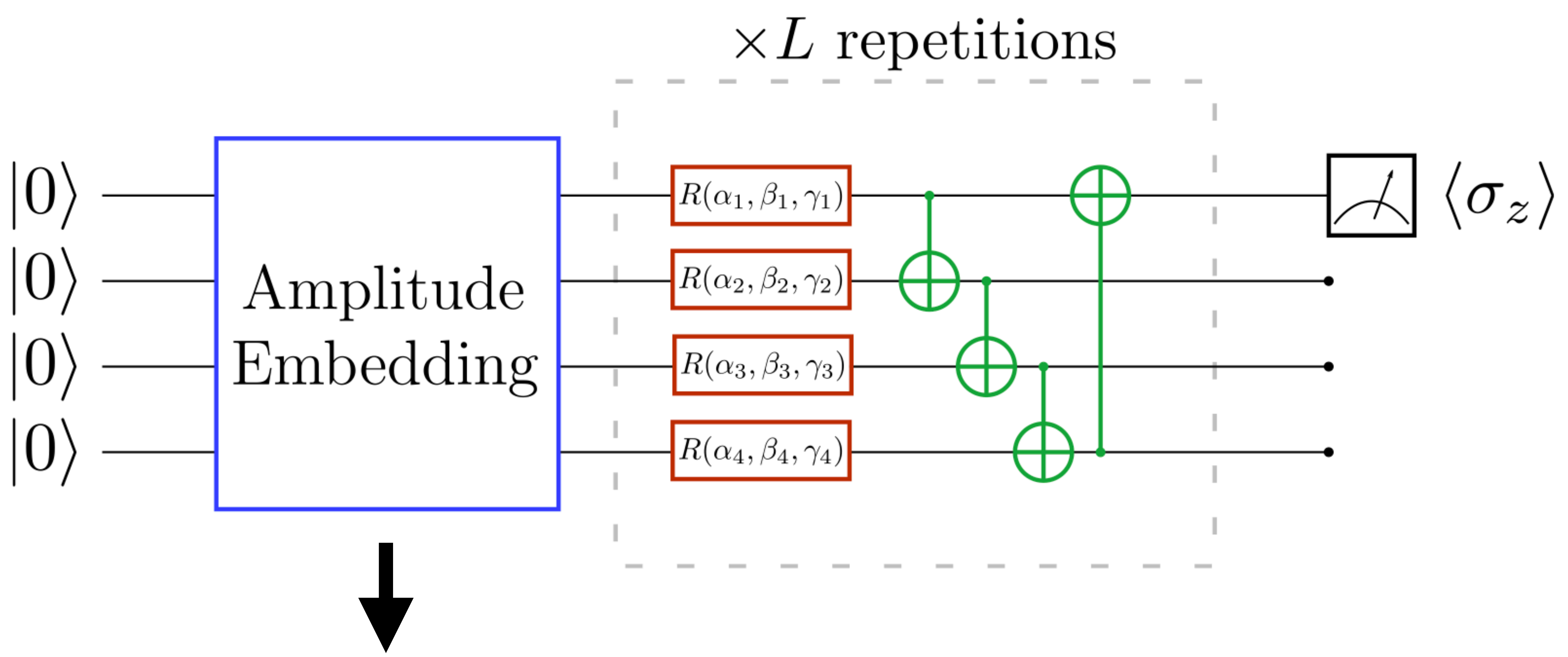


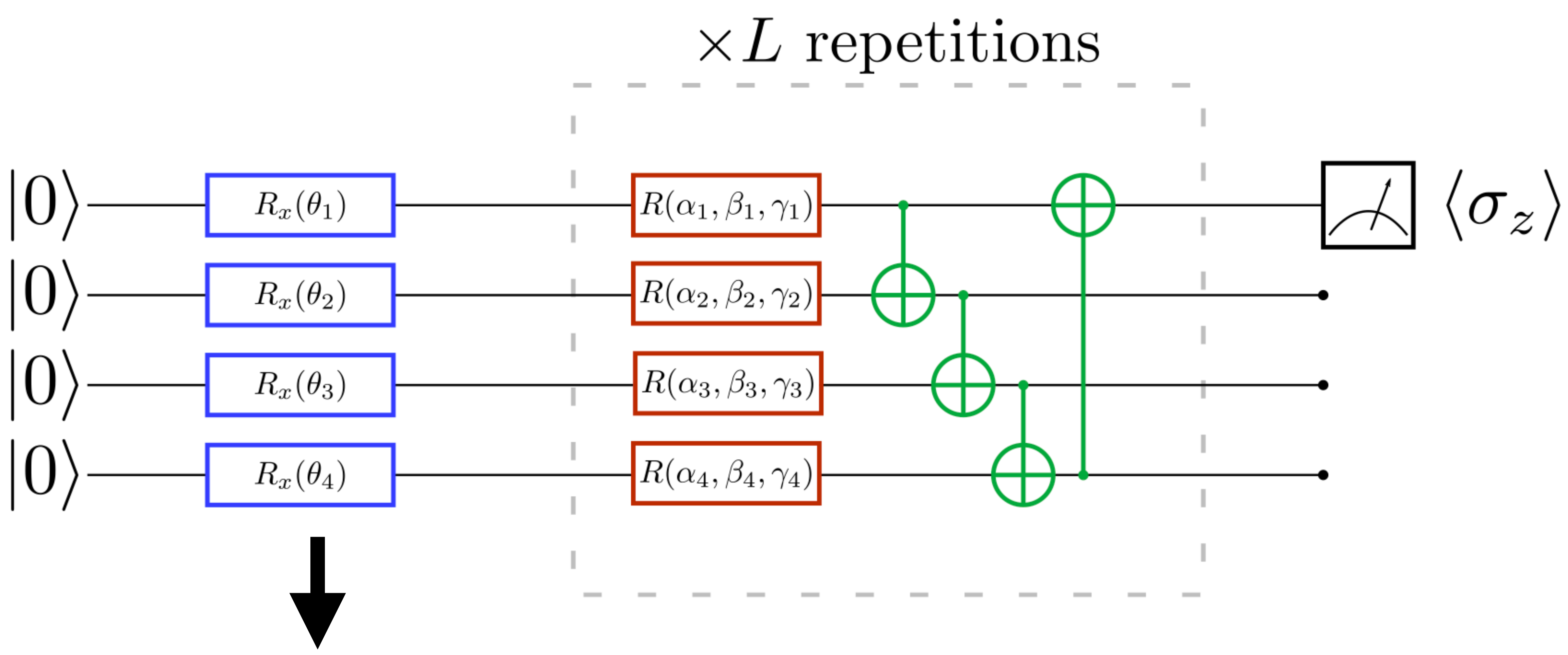
Two datasets/set of features:

- **Muon dataset**: jets with at least one muon, 3 muon features+jet charge
- **Complete dataset**: all jets, 15 particle features+jet charge

Quantum Circuits

- Embedding + *L* layers of rotational gates
- Two types of embedding tested: **Amplitude Embedding** and **Angle Embedding**



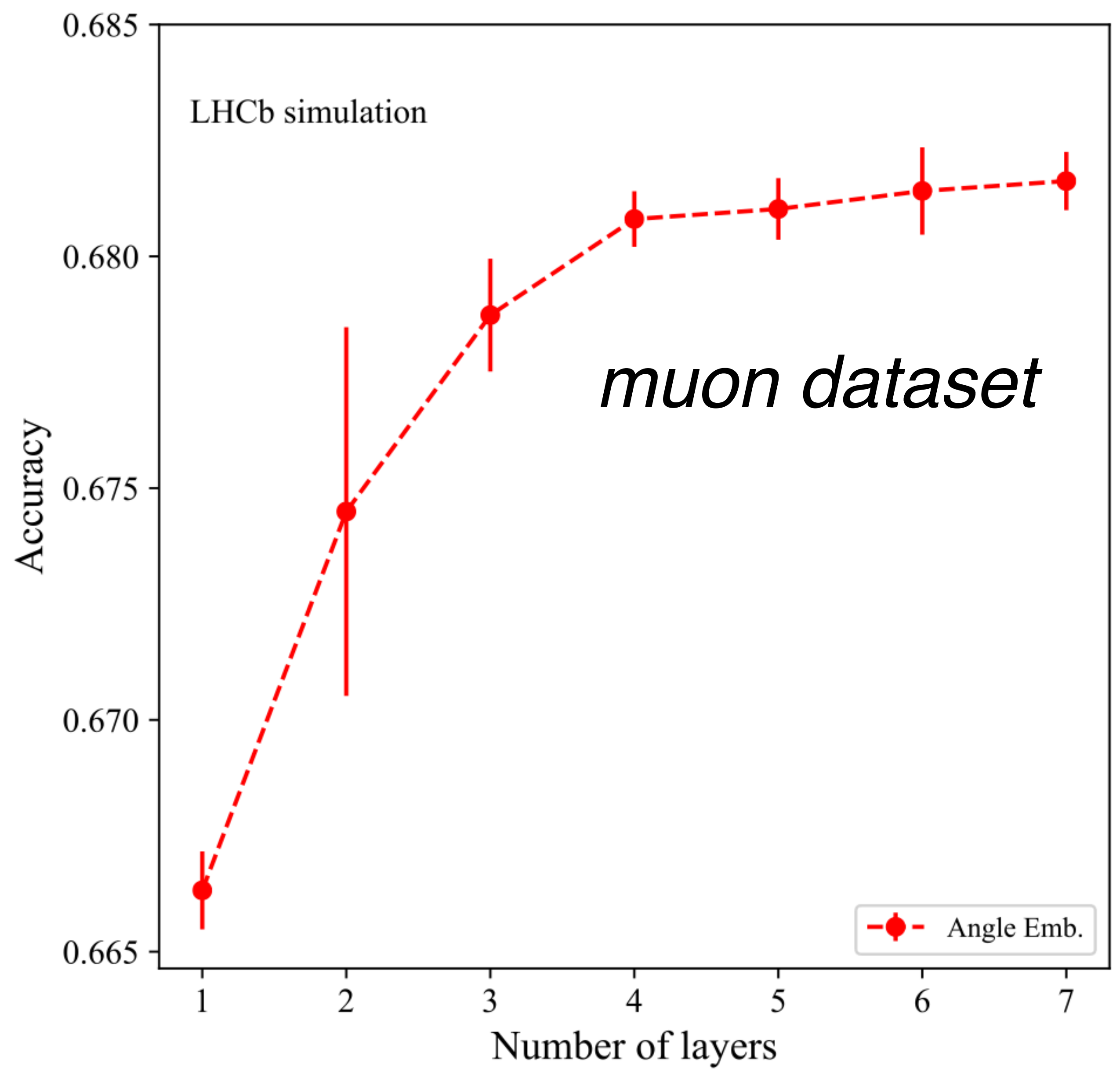


Amplitude encoder: 2^n features in n qubits

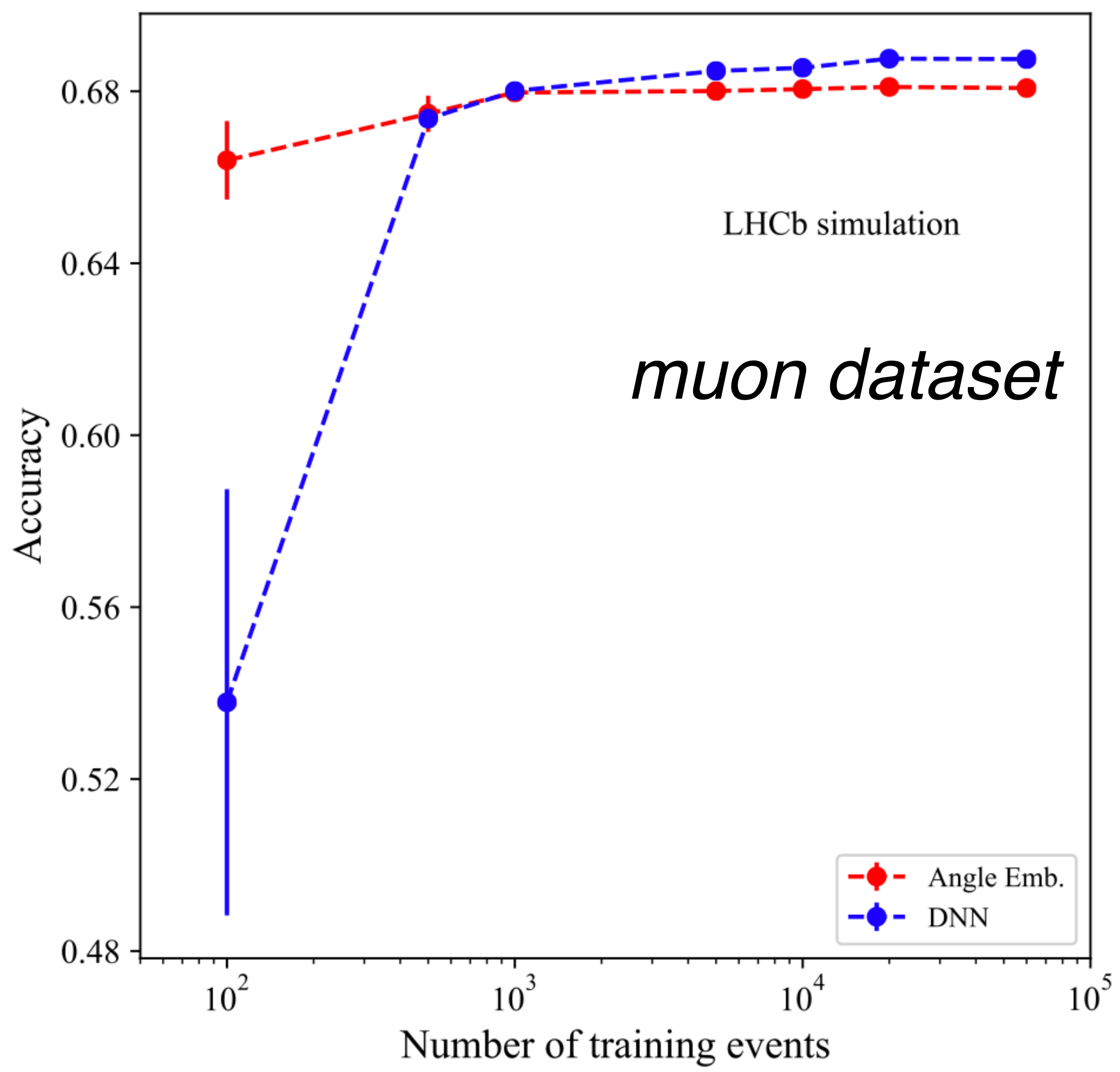
$$|x\rangle = \sum_{i=1}^{2^n} x_i \, |n_i\rangle$$

Angle embedding: one rotational gate per feature (#features=#qubits)

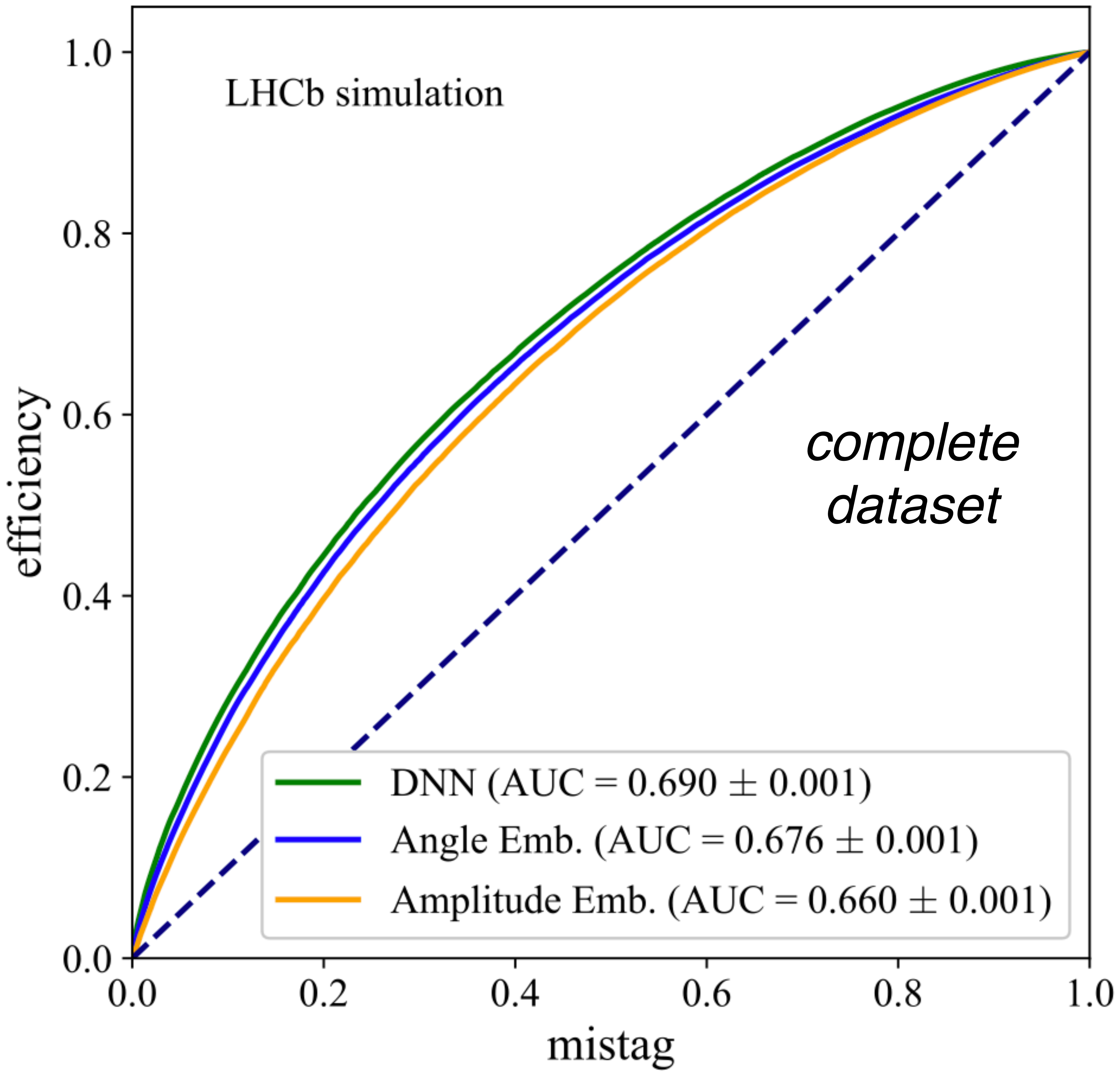
Results on simulator



Different number of rotational layers tested: **the accuracy saturates after few layers**



Compared to a classical DNN, the quantum classifier requires less training events to achieve the same accuracy

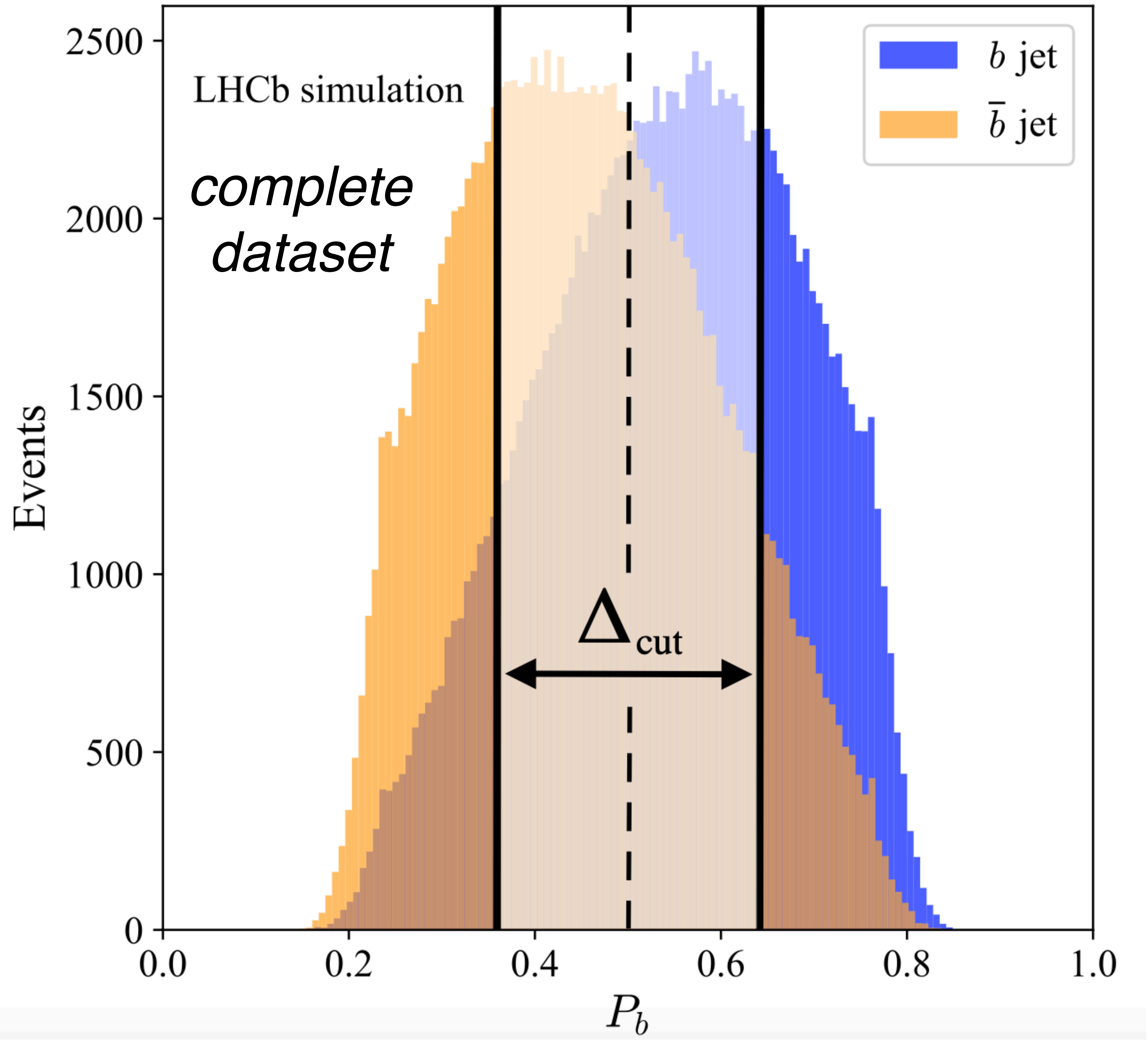


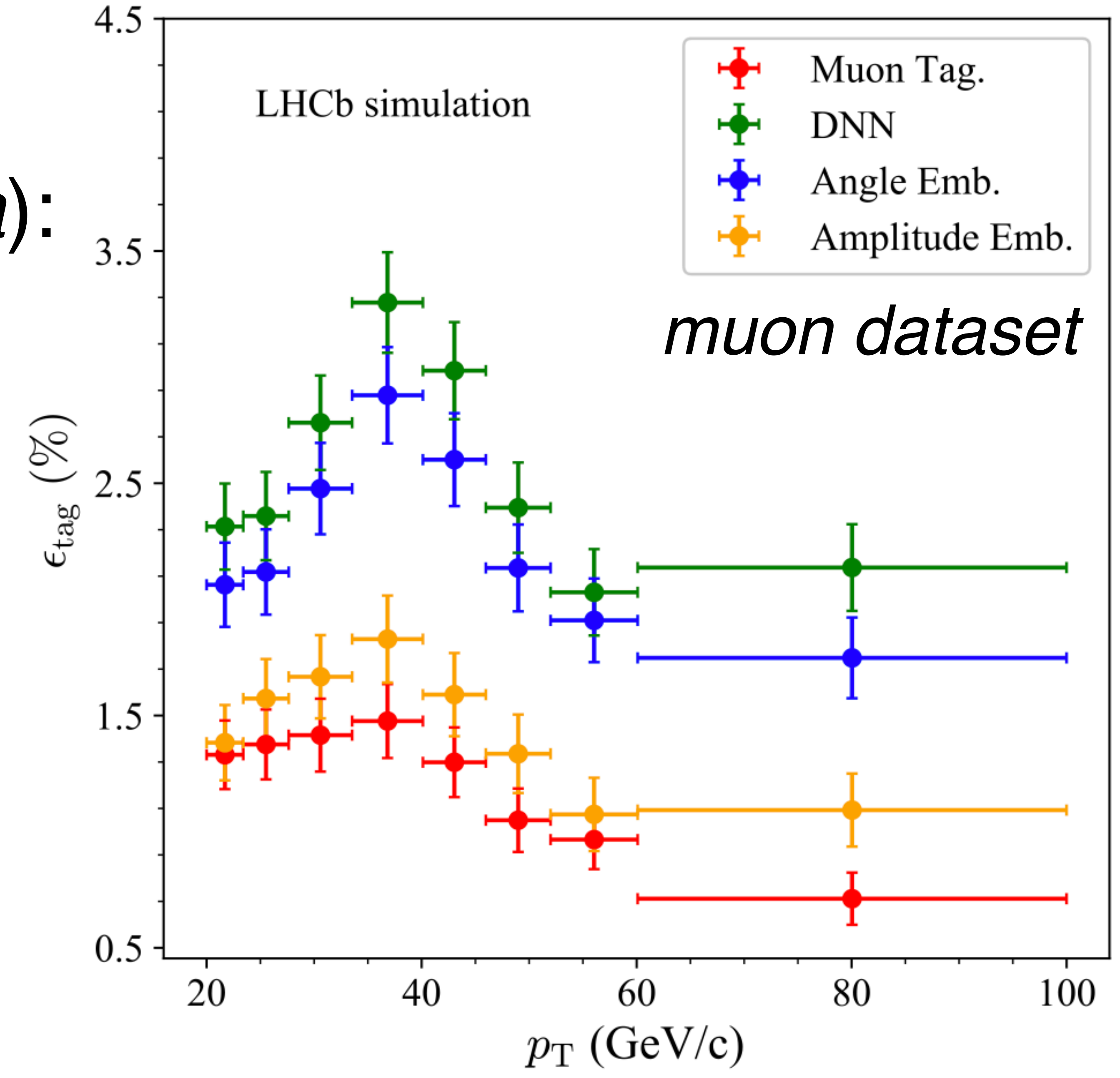
The DNN and the quantum circuits show similar ROC areas

Performance on simulator

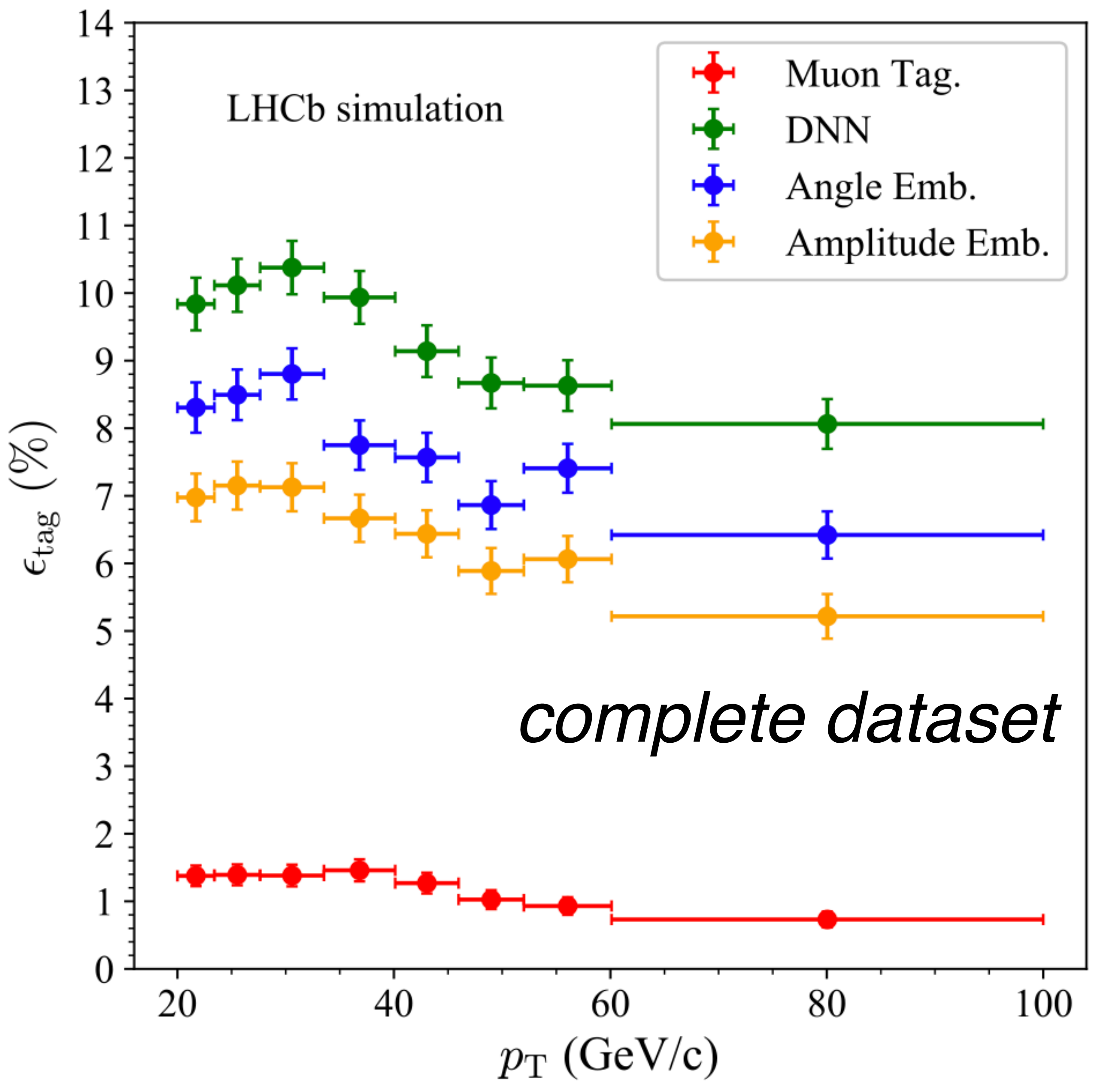
A requirement is applied on the probability output to maximize the **tagging power** (combination of efficiency, ε_{eff}, and accuracy, a):

$$\epsilon_{\rm tag} = \epsilon_{\rm eff}(2a-1)^2$$





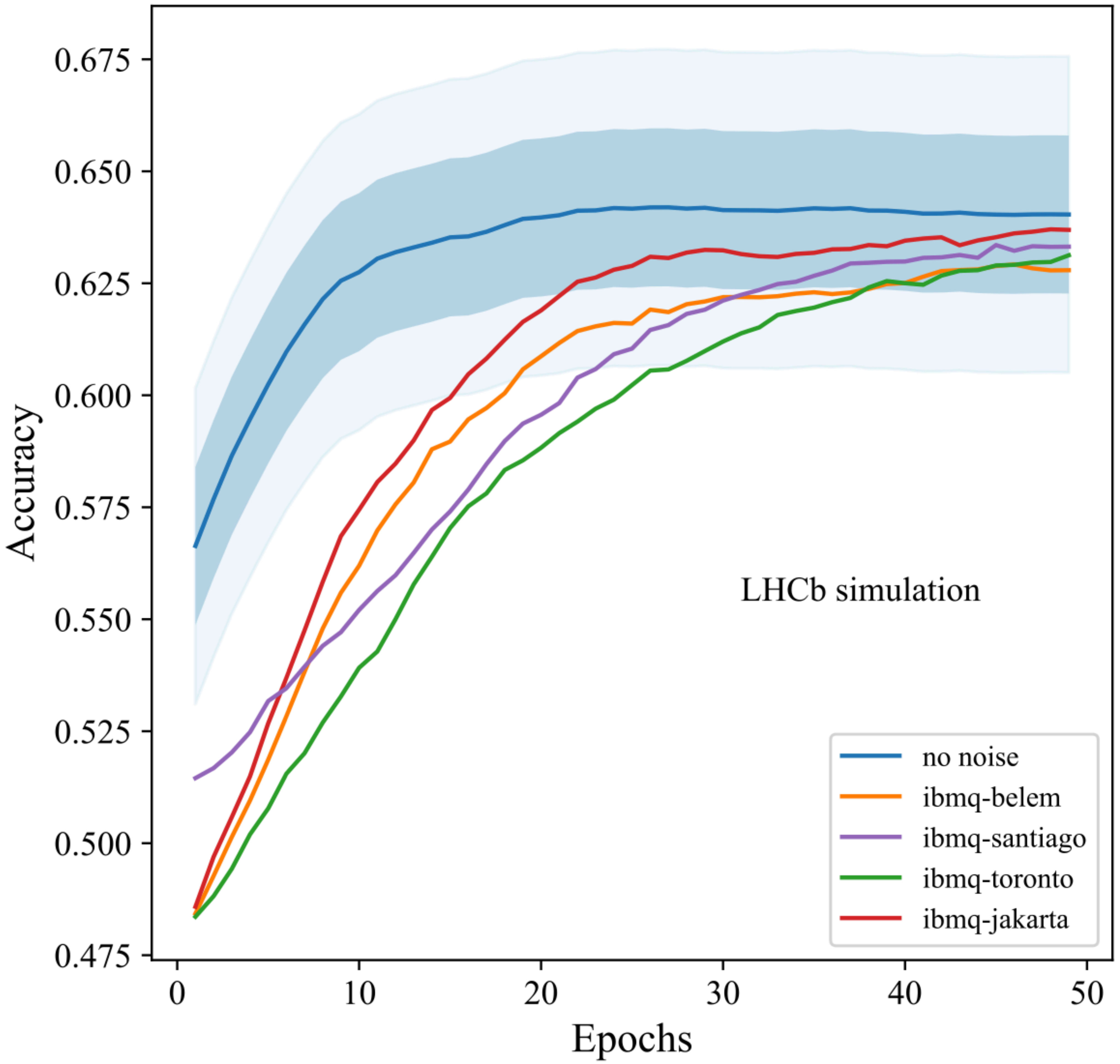
In the ***muon dataset***, the DNN and the Angle Embedding circuit have a **similar performance**



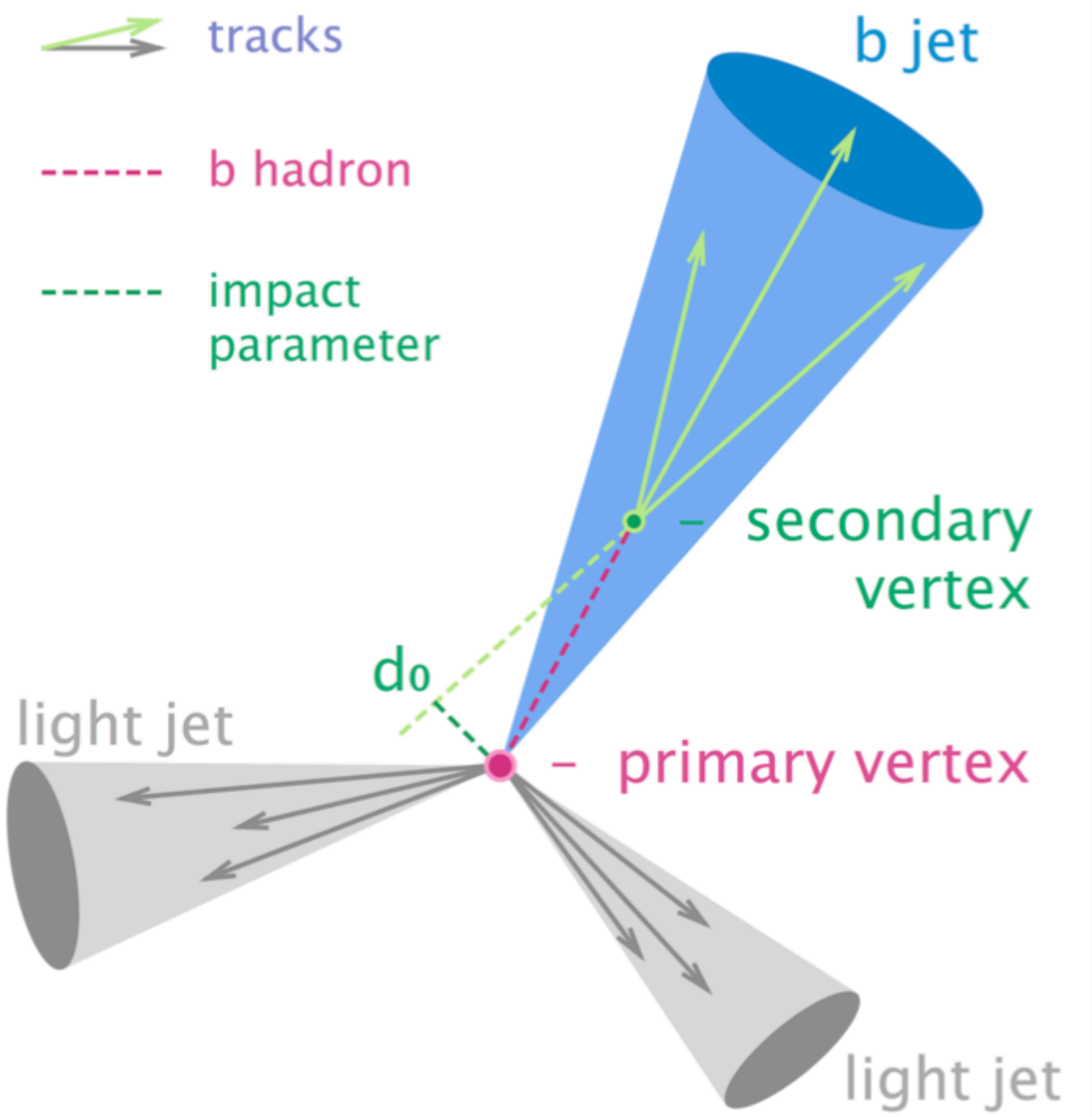
In the ***complete dataset***, the **Angle Embedding shows a lower tagging power than the DNN** (2% absolute difference)

Quantum noise

- Several **noise models** have been applied to the simulator in order to study its impact
- With the noise, **a higher number of training epochs is necessary to achieve the best accuracy**
- With a sufficiently high number of epochs, **the accuracy obtained with the noise is of the same order of the accuracy obtained without noise**

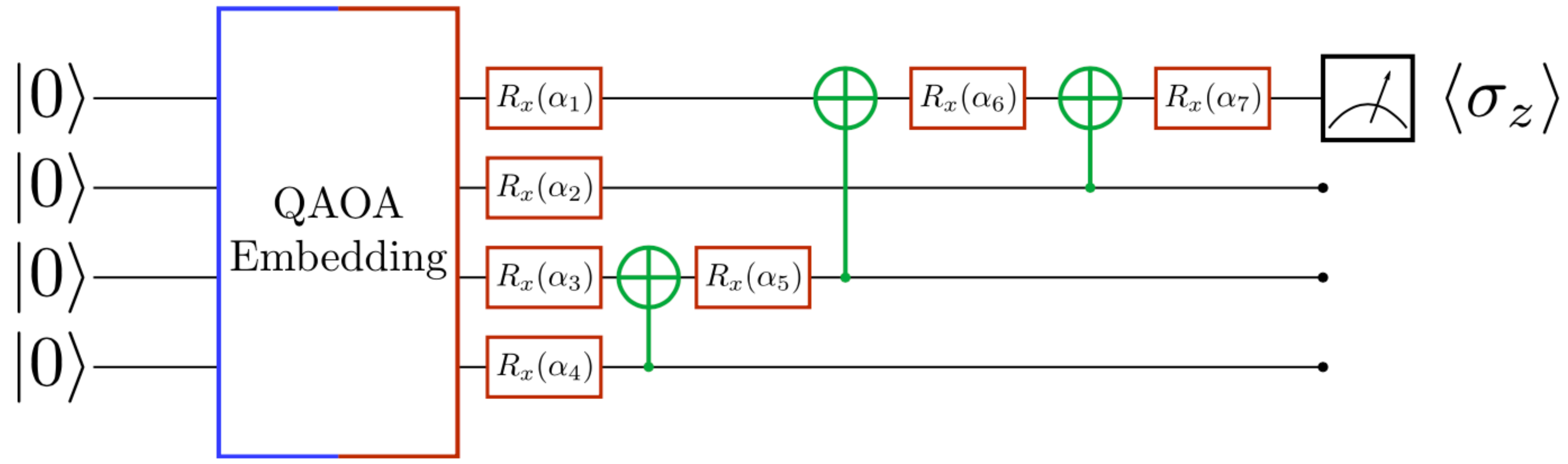


Classification of b- vs c-jets

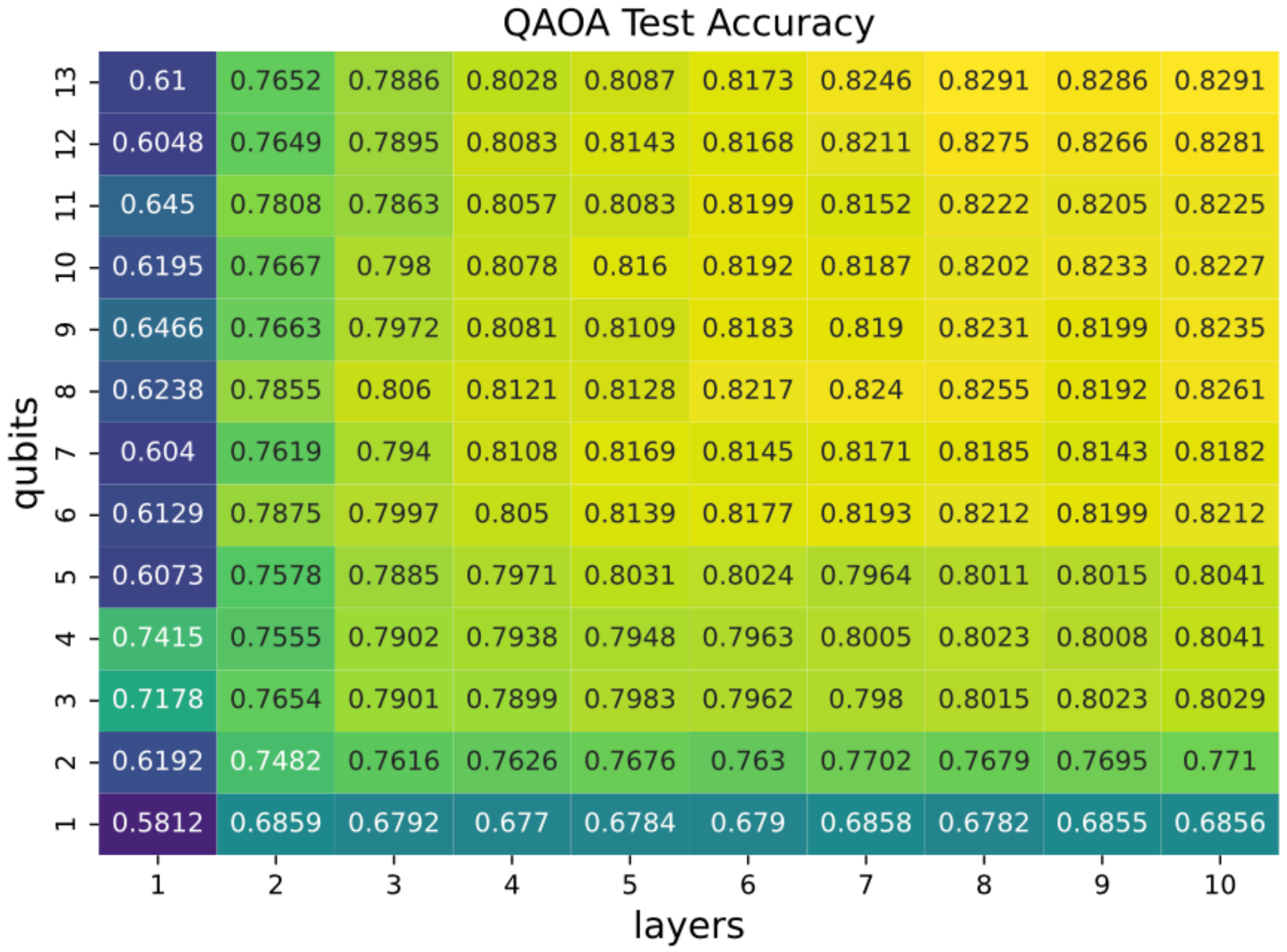


- For this task, features related to the reconstructed **Secondary Vertex (SV)**, formed by particle tracks and matched with the jet, are used
- Most important features:
  - SV mass
  - SV corrected mass
  - Fraction of jet momentum taken by the SV
  - Delta R distance of SV with respect to jet axis
- **From 4 to 13 features are used**

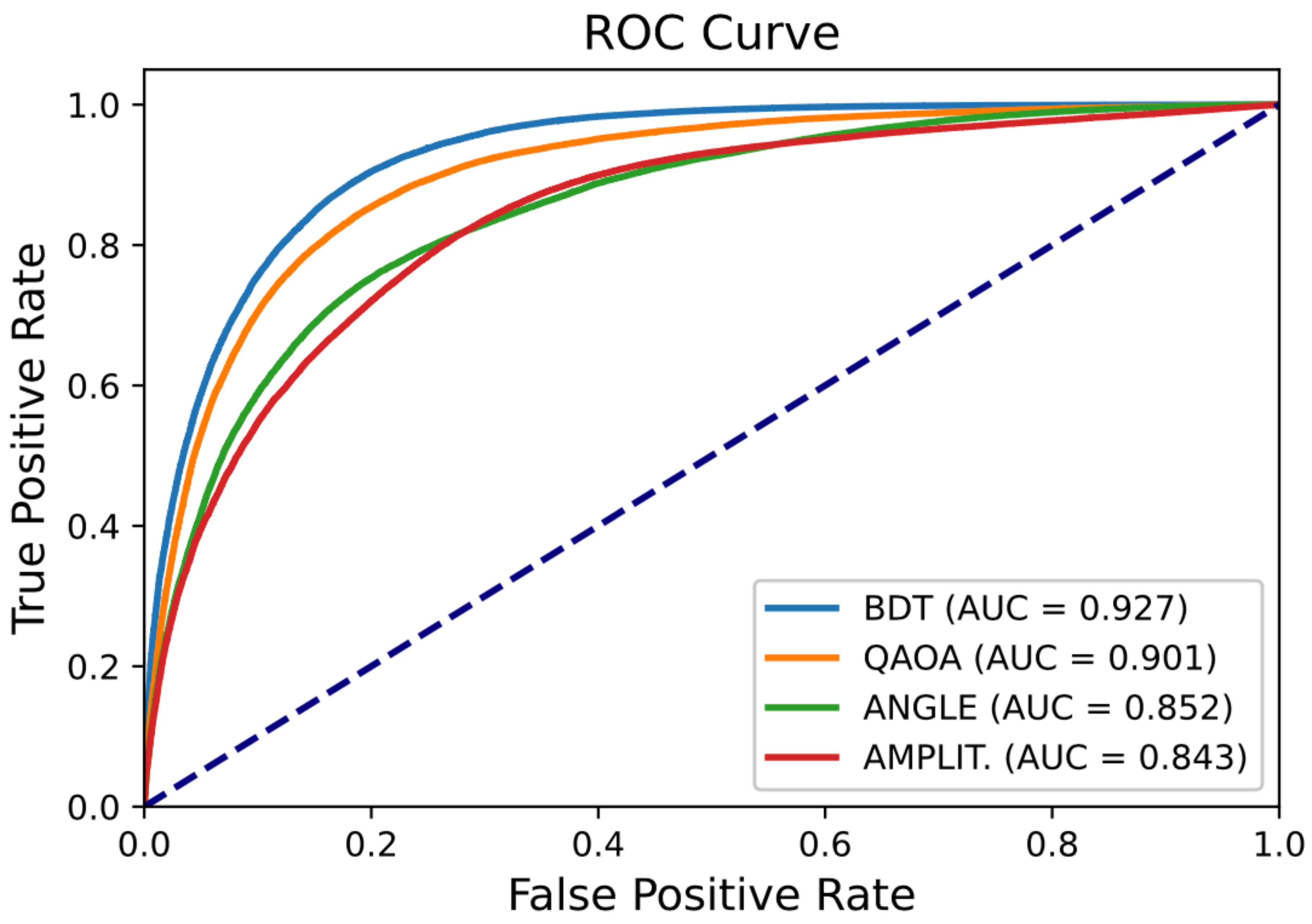
Classification of b- vs c-jets



Quantum Approximate Optimization Algorithm encoding (QAOA), a variational circuit for the embedding



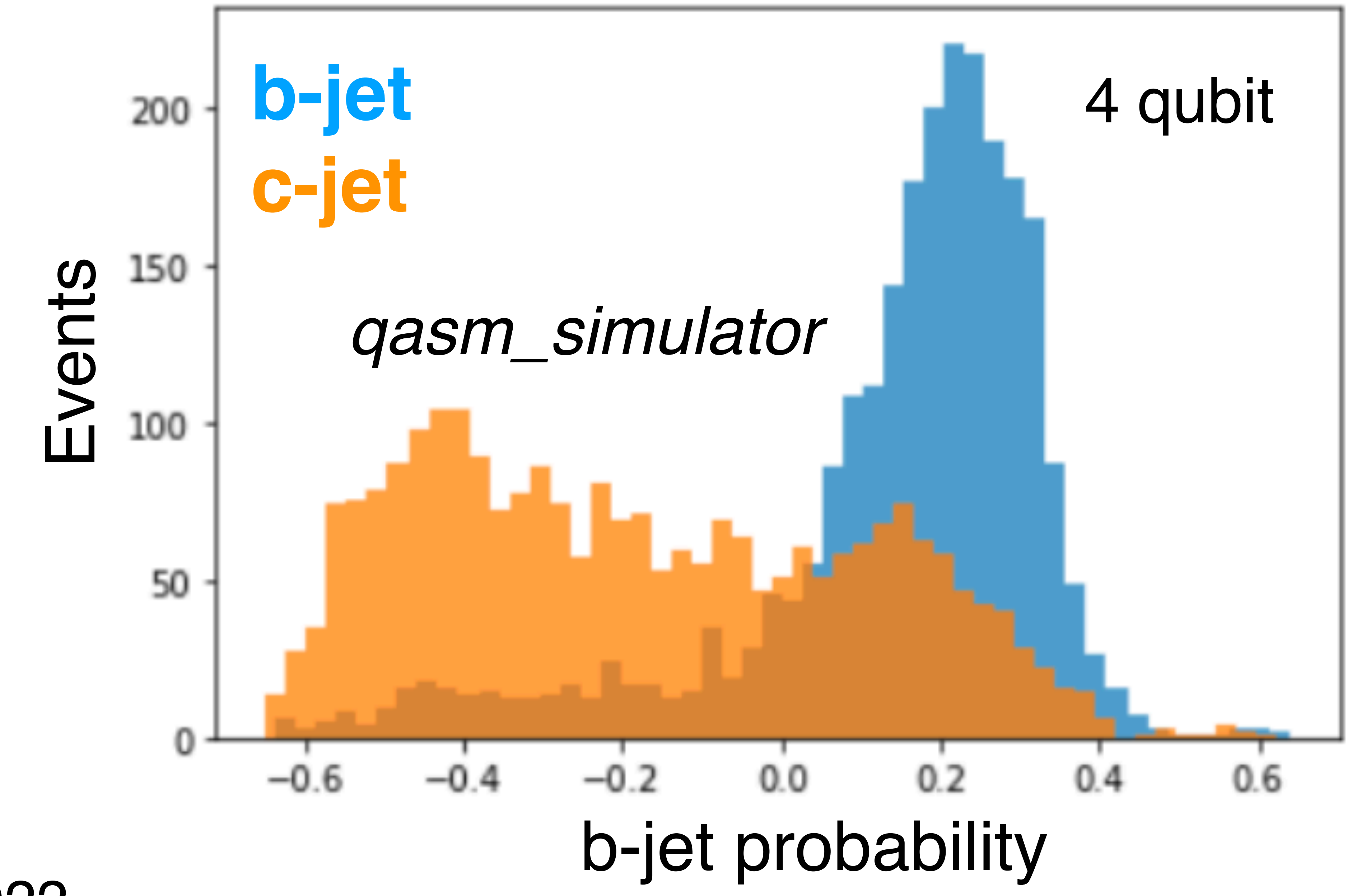
Several tests have been performed by varying the number of features (=#qubits) and the number of layers

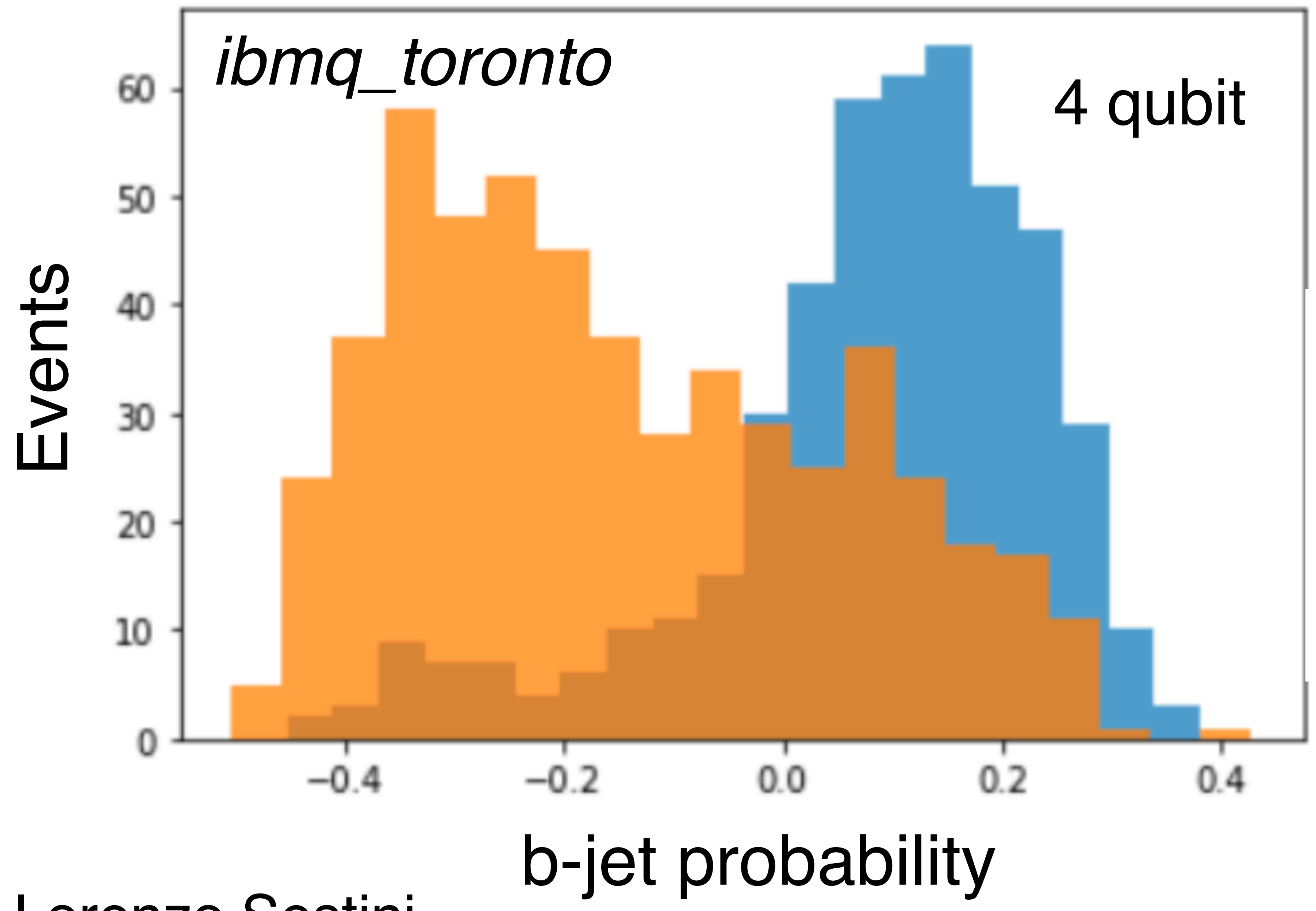


With the QAOA algorithm the performance is close to the one obtained with a classical BDT

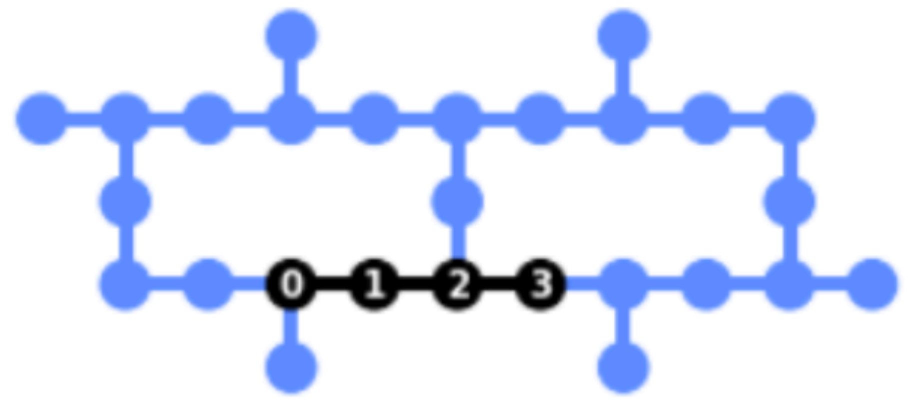
Validation on hardware

- **The evaluation of the pre-trained quantum circuit for b vs c has been performed on IBM hardware**
- The goal is to check if there are differences in the output between hardware and simulator
- For this task the circuit has been implemented using the **Qiskit** library, (angle embedding is considered)
- **The probability distributions show some differences, but the discriminating power is similar**





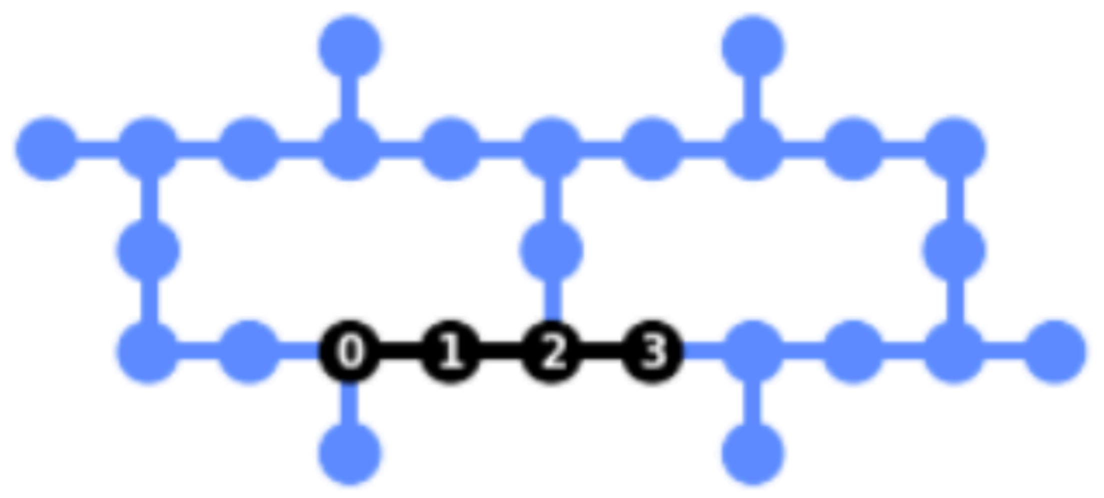
ibmq_toronto 27 qubits



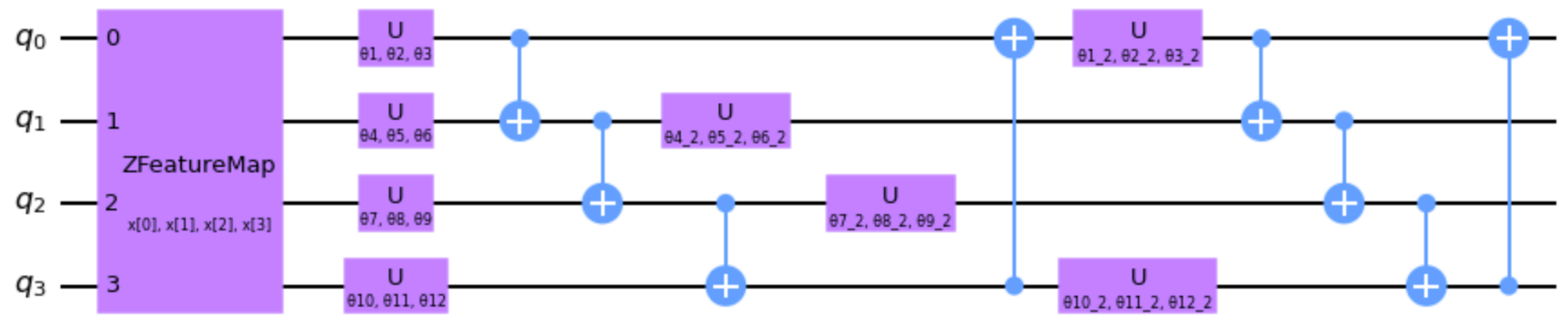
Prospects: circuit optimization

- **When circuits are ported to the hardware, they look very different from the original design**: the implementation depends on the qubit connections, geometry and native gates
- The optimization is done with the **transpiler**
- However we should try to perform an accurate circuit design to improve the **timing performance, impact of the noise** etc.

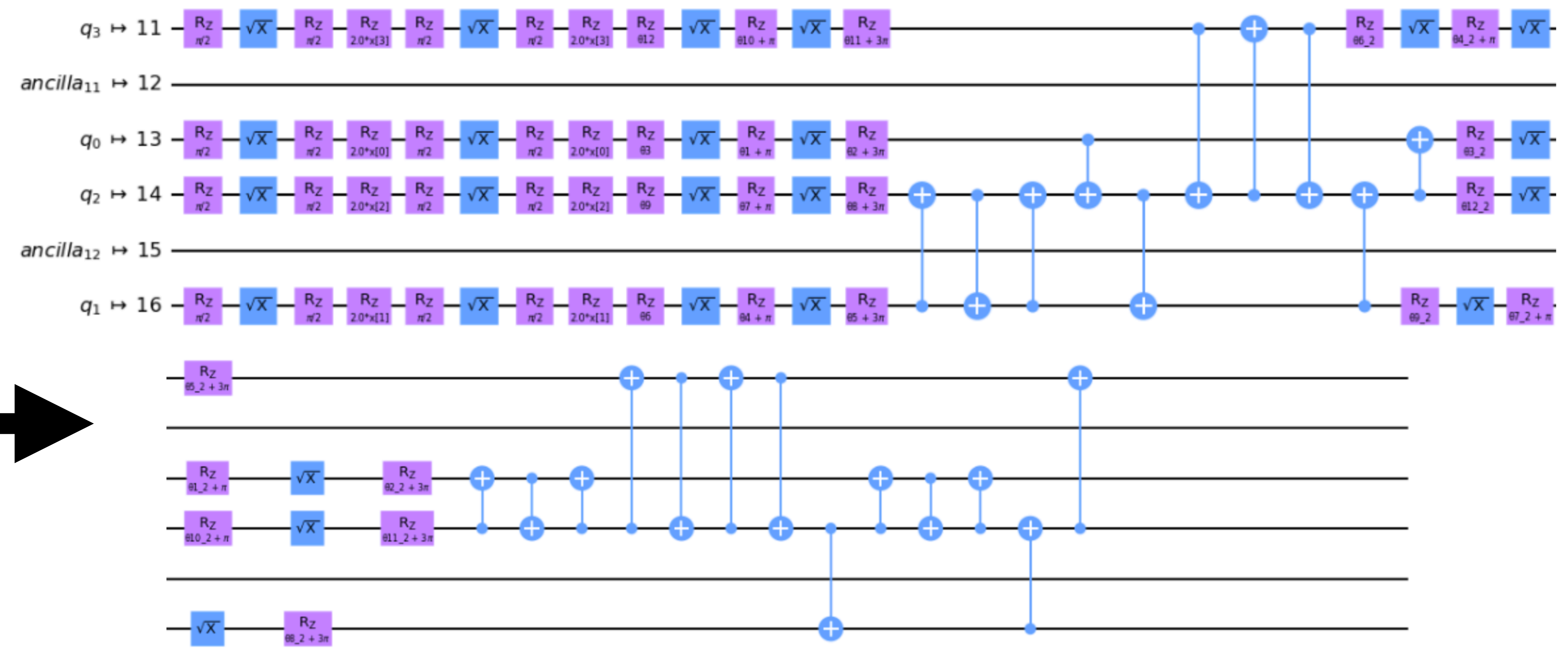
ibmq_toronto 27 qubits



- We are also studying the impact of **noise mitigation techniques**



4-qubit angle embedding circuit

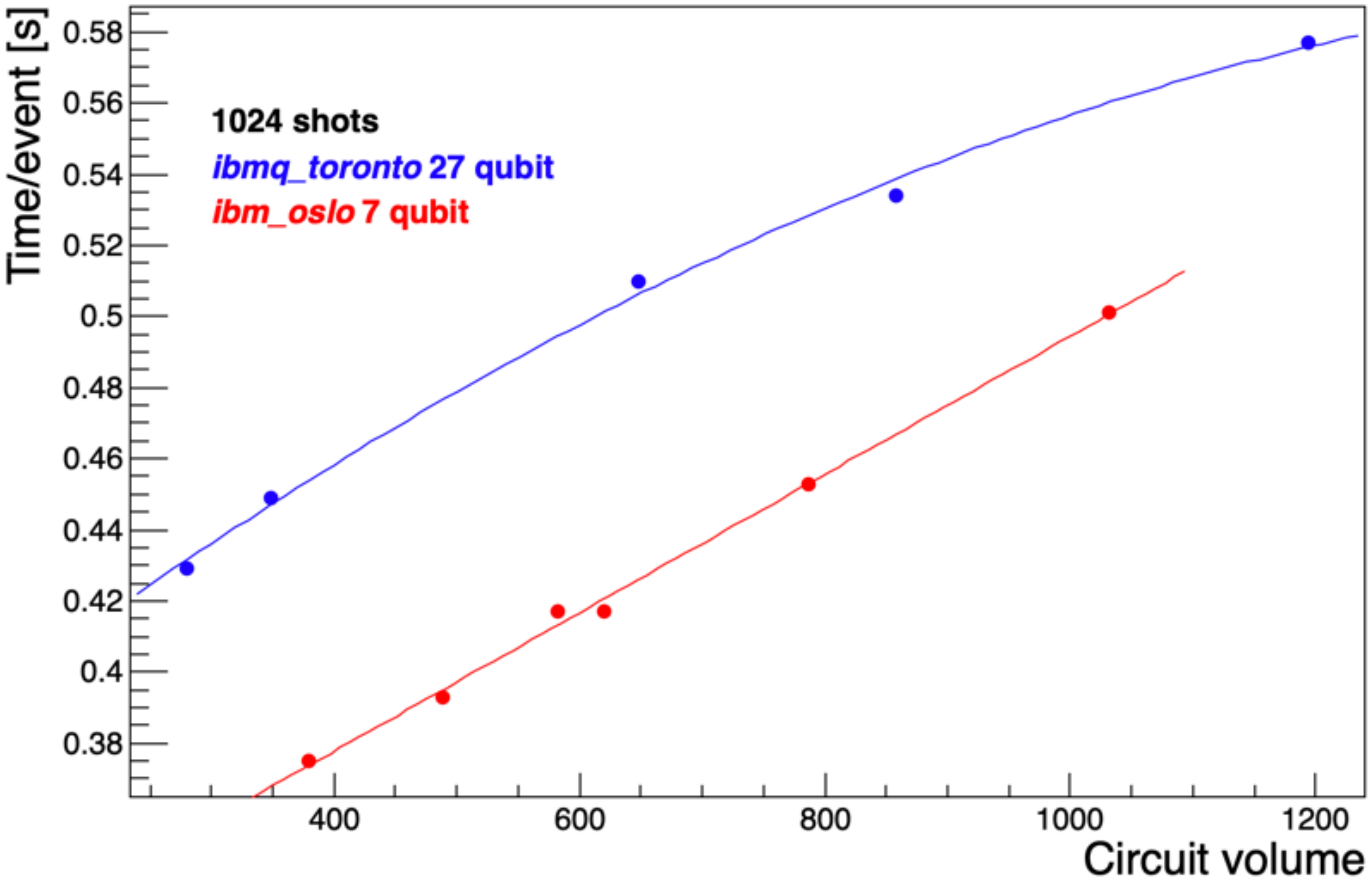


Same circuit on the ibmq_toronto hardware

Prospects: timing performance

- We have measured the **job time on IBM hardware**
- The queue time should be already subtracted
- There is a **dependence of the time from the Circuit Volume**
- However we have several questions: **how this time is divided in quantum and classical operations? How much time is needed for data upload?**
- An accurate analysis and comparison with simulations can help in **scaling the performance to larger Circuit Volumes**

Circuit depth = maximum number of gates applied to the same qubit
Circuit Volume: depth*number of effectively used qubits

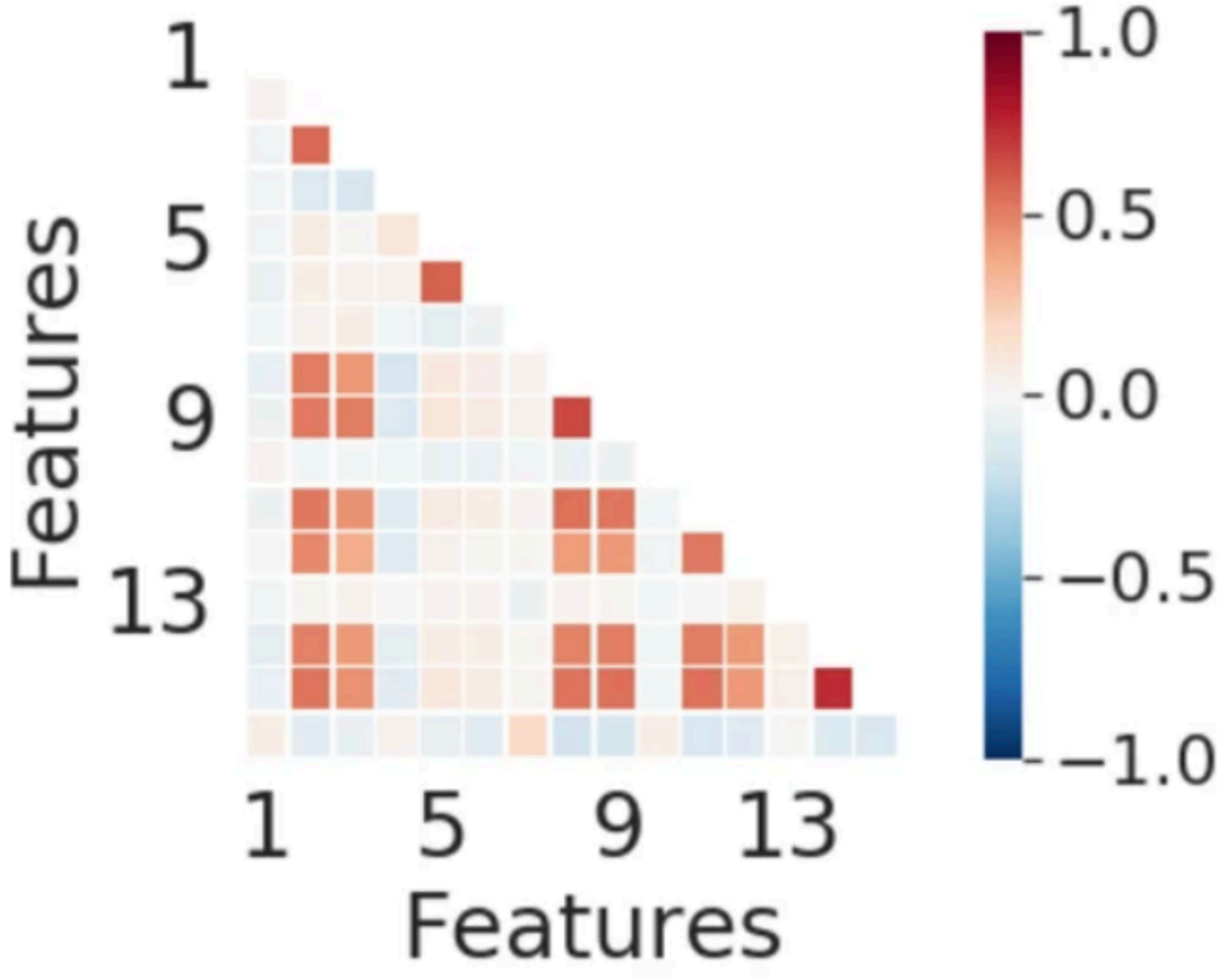


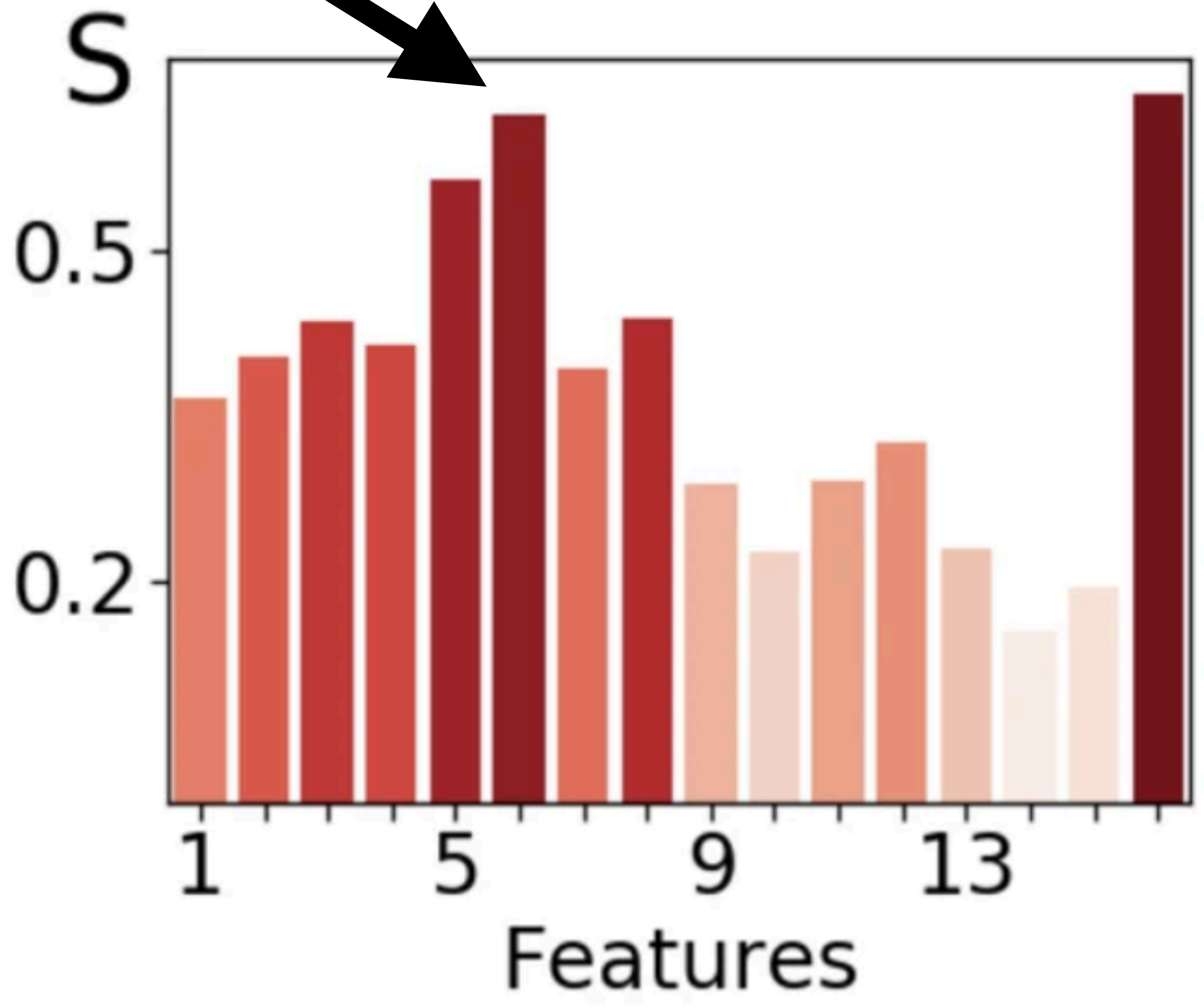
Prospects: entropy and correlations

- **Quantum circuits could give us more information on data than classical machine learning**, by measuring **entanglement correlations and entropy** between qubits (features)
- A proof of principle on the b vs $\bar{b}$ task at LHCb has been given in (npj Quantum Inf **7**, 111 (2021)), for a quantum-inspired method: **the entropy and correlations have been used to determine a ranking of the features**

The same side Kaon algorithm has been re-discovered

- Could the quantum entropy and correlation give us a deeper insight on data?
- Could be useful to measure these quantities on real data? **Could they be used to improve our simulation**?
- A more general question: **do we have quantum data in our experiments**?





Conclusions

- **The LHCb collaboration is lively working on Quantum Computing algorithms**
- The **Jet Identification** has been considered as first problem
- **The Quantum Algorithms have shown a similar performance to the “classical” machine learning algorithms, but they are not yet able to surpass Deep Neural Networks/Boosted Decision Trees**
- We are working on several promising aspects:
    - Circuit optimization in the **hardware**, noise mitigation strategies
    - Measurement of **timing performance**, scaling to larger Circuit Volumes
    - Study **Entanglement Correlations and Entropy** to learn something new from our data

Thanks for your attention!

Backup