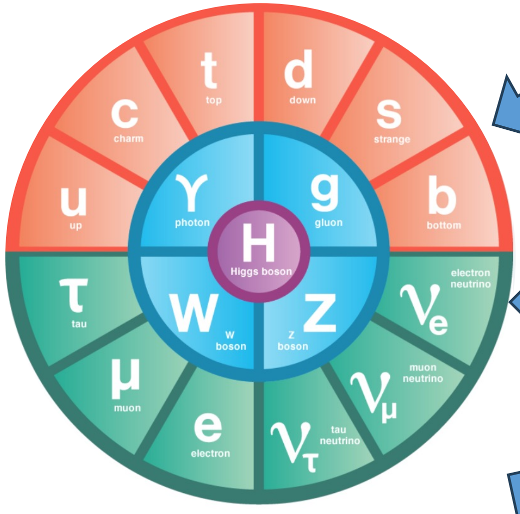
Quantum Machine Learning in CEPC

Yaquan Fang (IHEP)

15th France China Particle Physics Network/Laboratory Workshop, 2024
Bordeaux, June 10 – 14, 2024

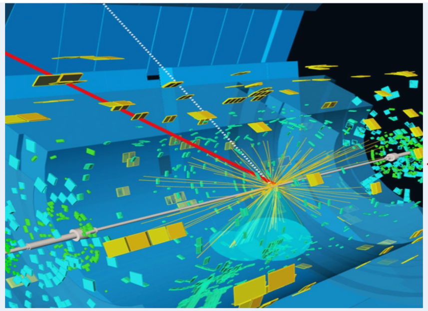
Quantum Computing in HEP

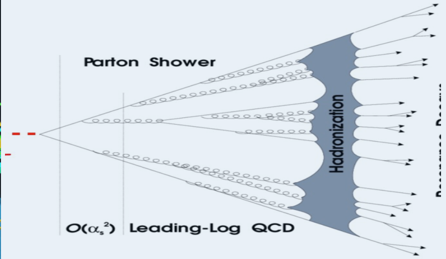
HEP



Particle Collision Calculation

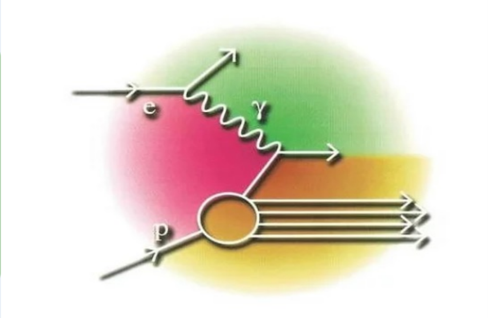
e.g. parton shower computation

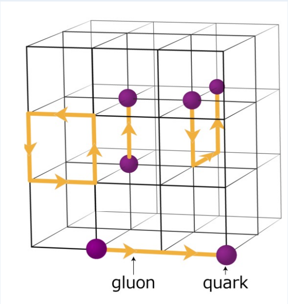




Quantum Field Theory

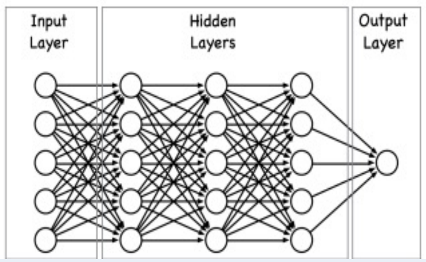
Lattice QCD

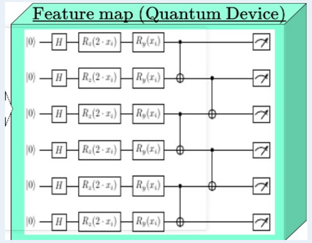




Data Analysis

Quantum Machine Learning



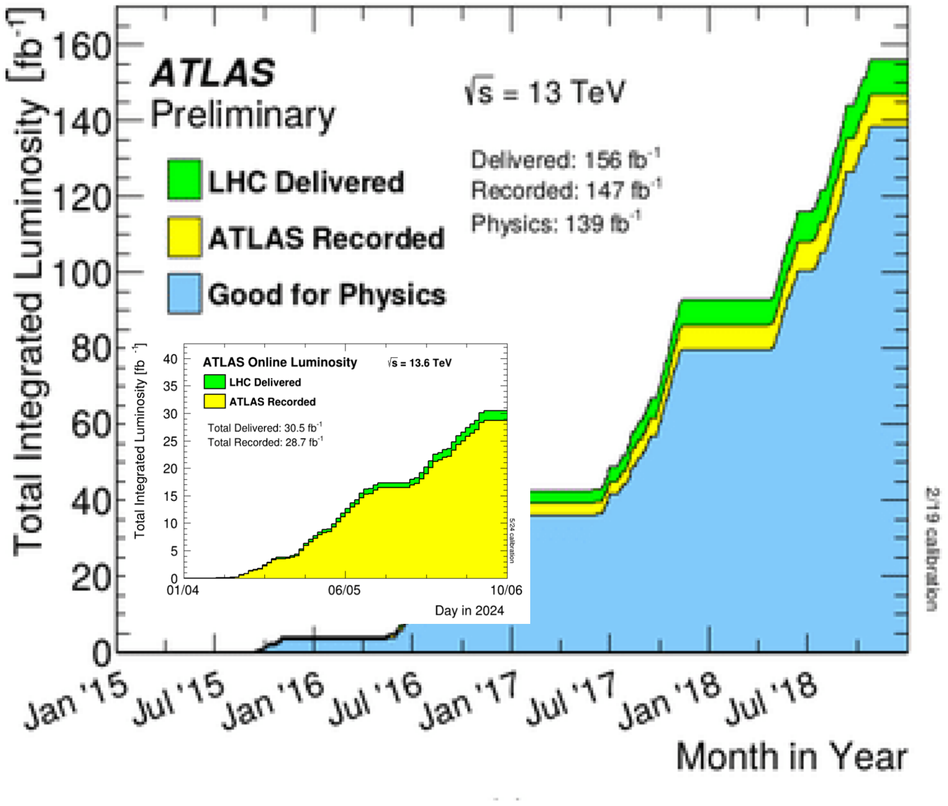


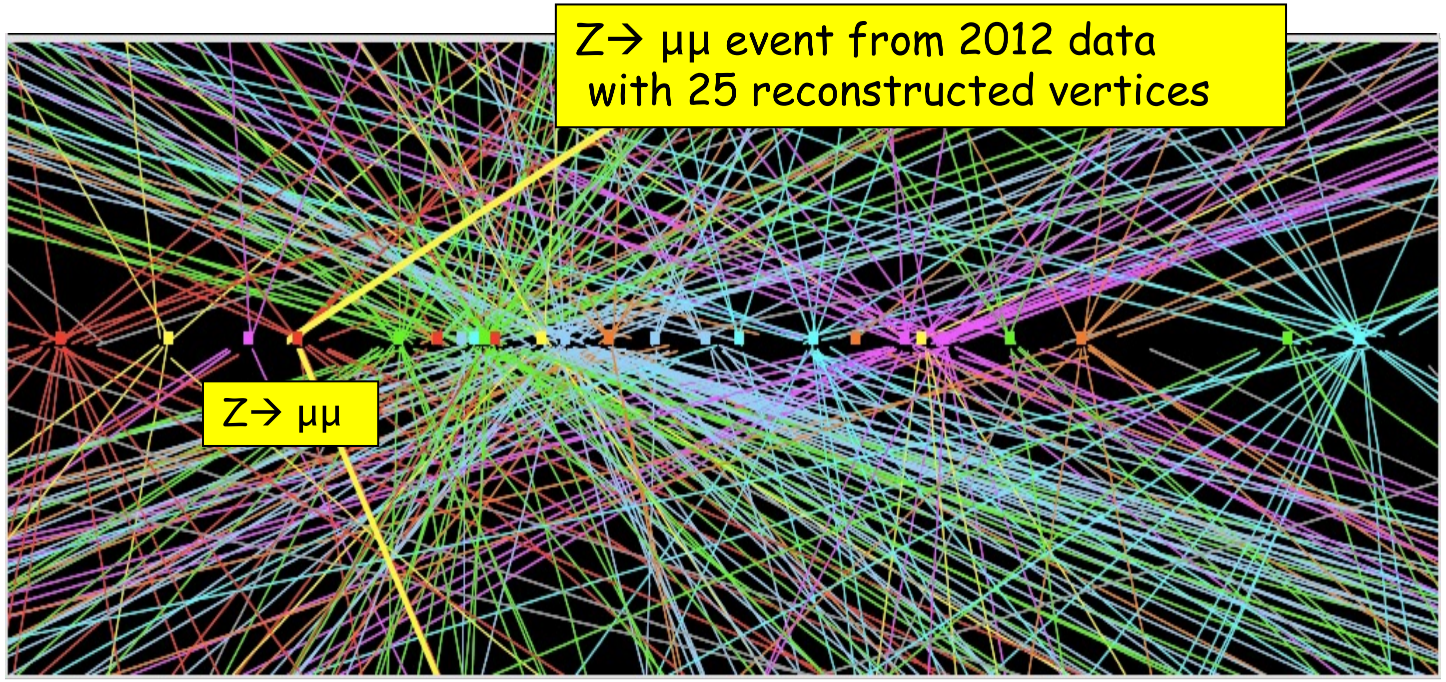
Big Data in HEP (LHC as an example)

ATLAS/CMS 200 events/s passing triggers

ATLAS/CMS 2PB/year of data

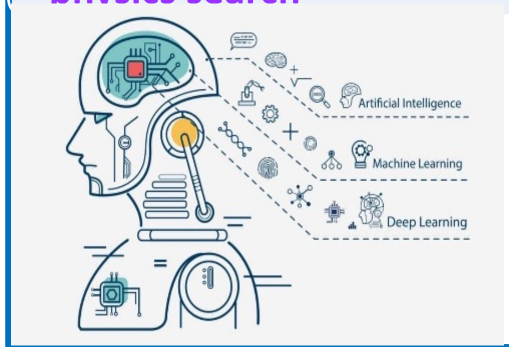
Tremendous amount of highly complex data



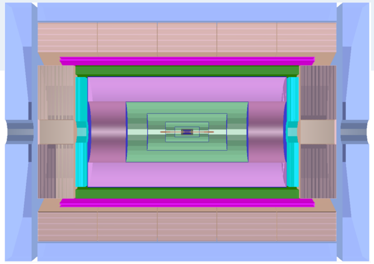


Big Data possibly leads to the Large Model

AI HEP detector simulation/new physics search

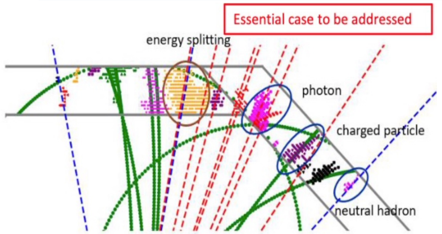


Detector design



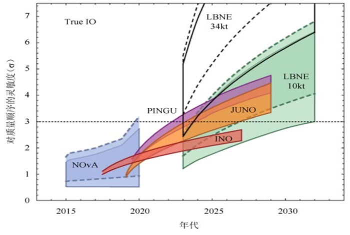
+

Detector Simulation



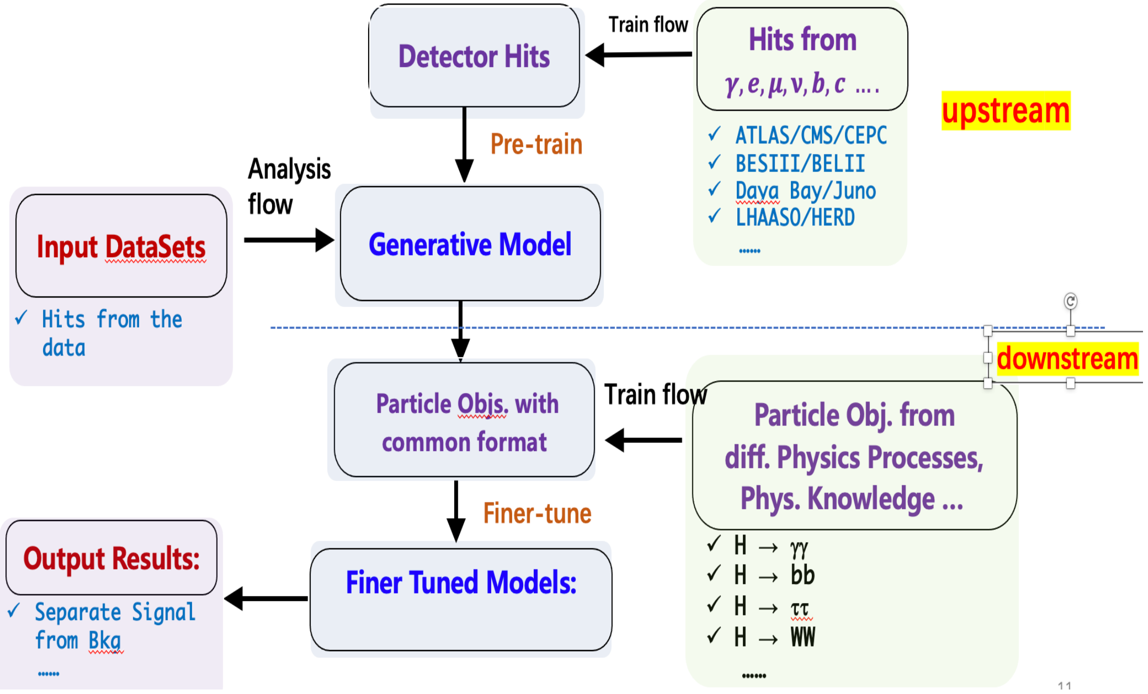
+......+

New physics search



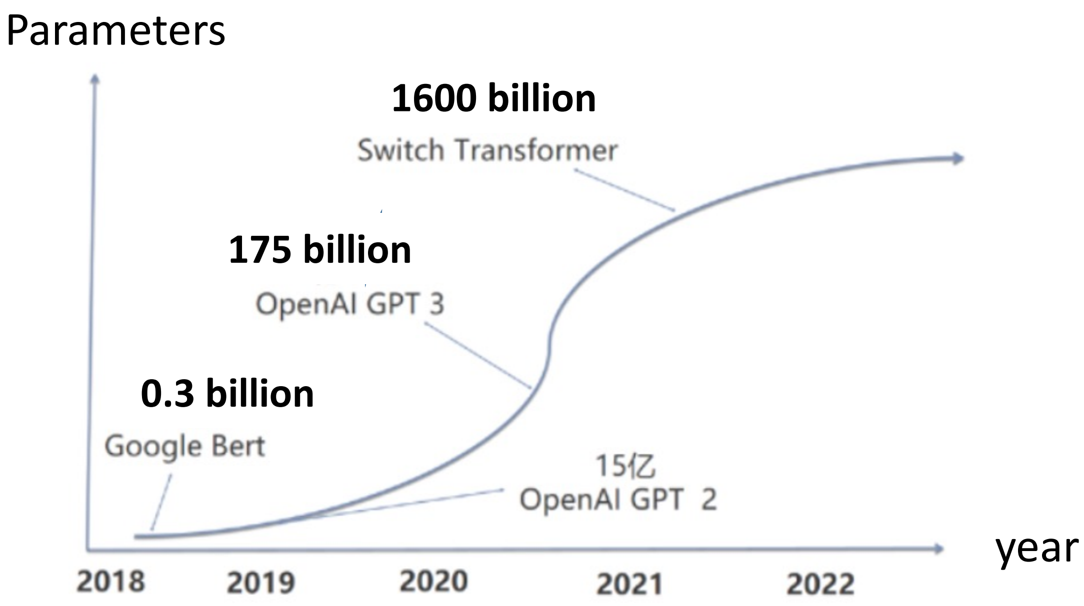
Elements:
Transformer： self-attention
Generative model
End-to-end (Hit)

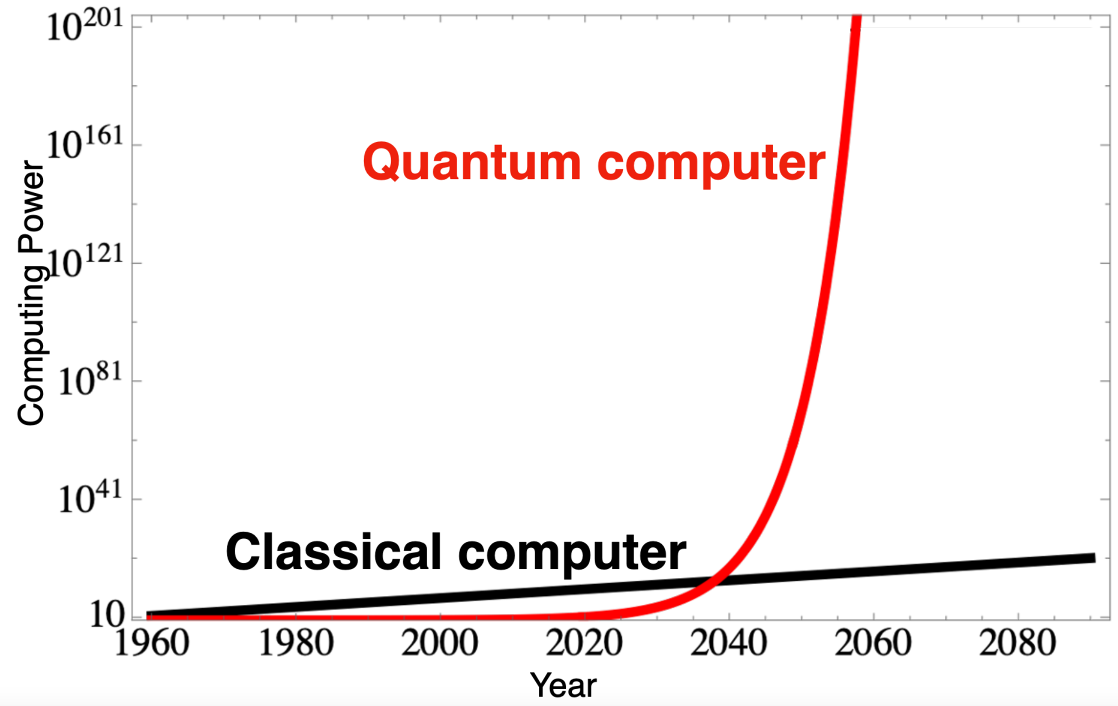
Generative model for physics analysis



Following the idea of the Generative model, 2 - photon separation in CEPC based on end-to-end (hit) will need several Million parameters

Quantum AI： the exponential demand of the computing power

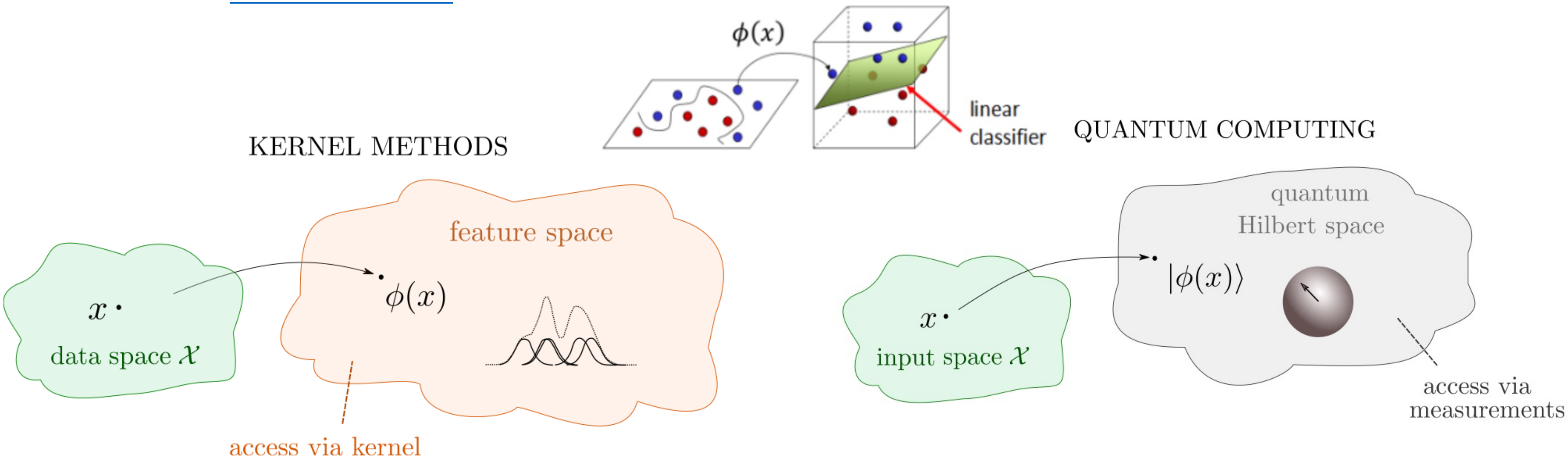




- **The increasing computing power from classical computer doesn’t match with the exponential demand from the development of AI.**
- **Potentially Quantum computer can resolve this challenge.**

Machine Learning (ML) & QML

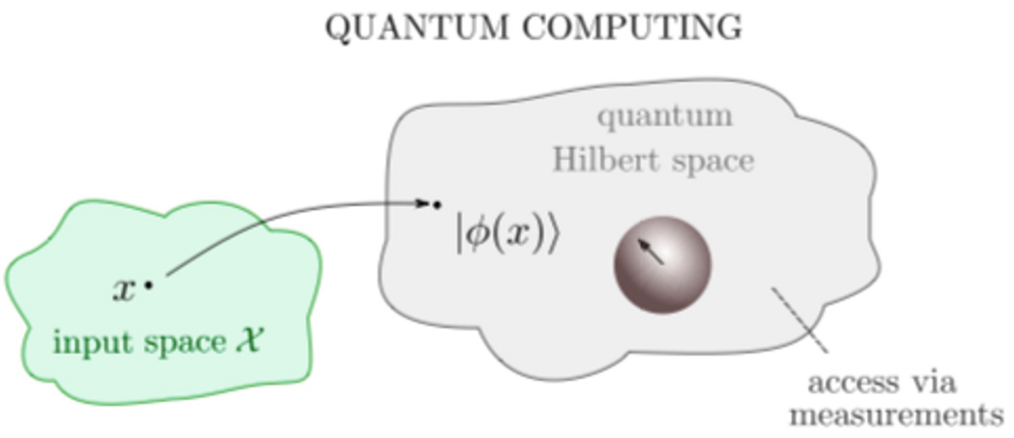
Maria Schuld [arXiv:2101.11020]

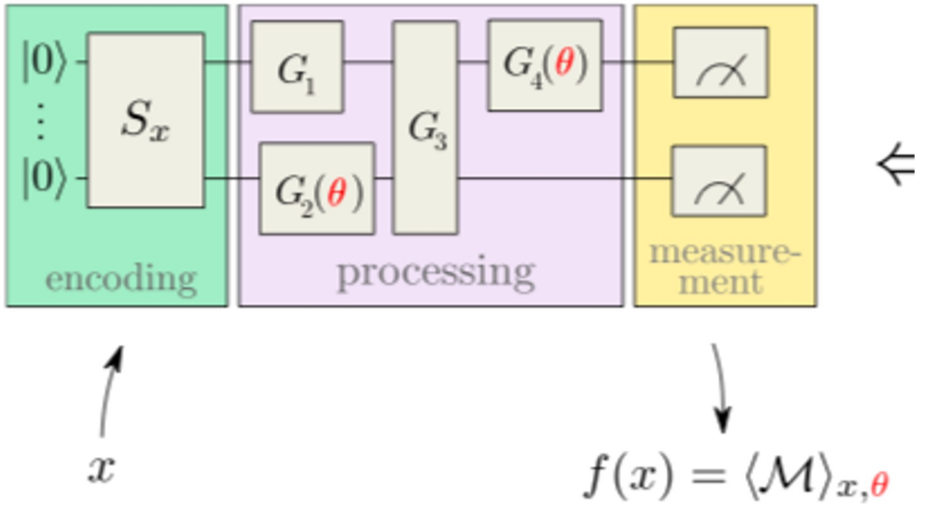


- ✓ **Both ML and QML map the information to high-dimensional spaces**
- ✓ **QML maps to quantum Hilbert space.**

What Quantum Machine Learning does

Maria Schuld [arXiv:2101.11020]

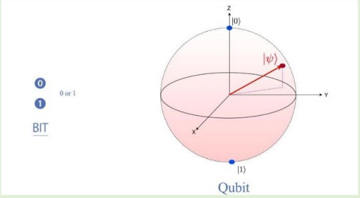


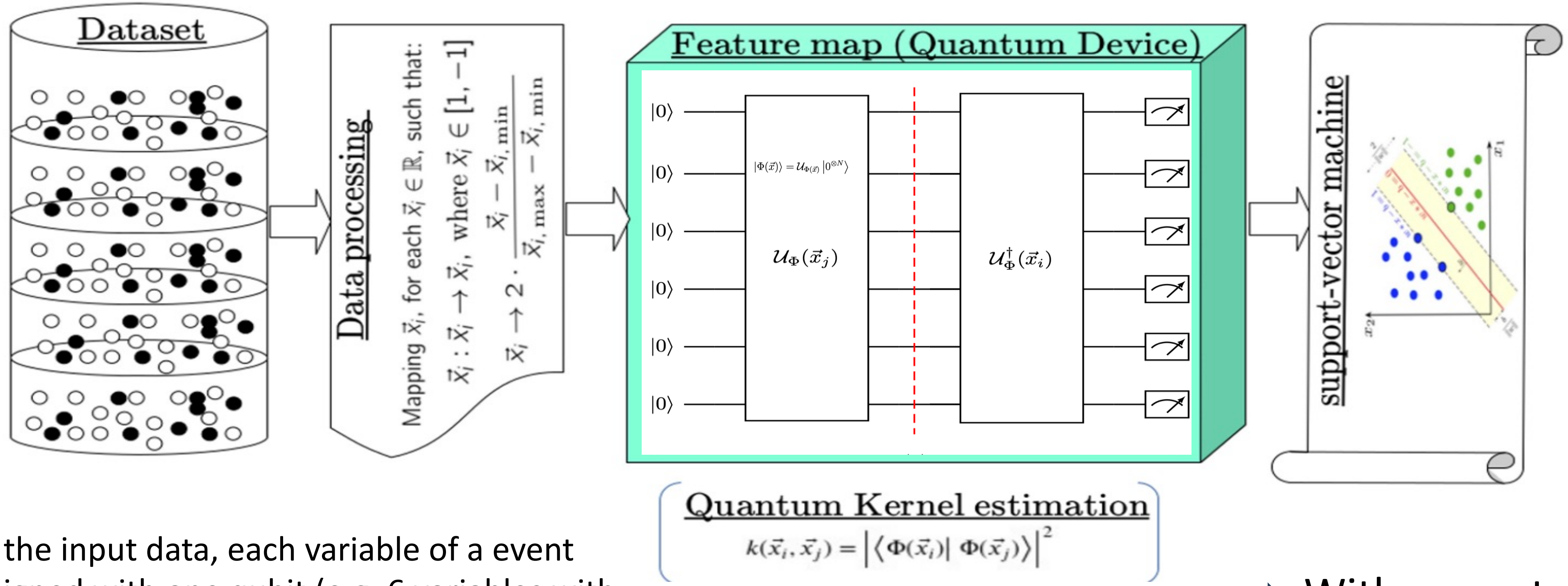


Encode input data to a quantum state

"process" (transform) the quantum state, and access the quantum state via measurements

QSVM





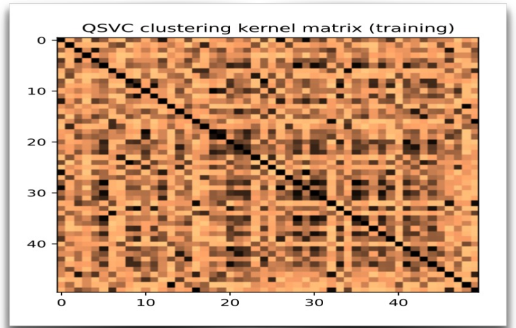
✓ For the input data, each variable of a event is assigned with one qubit (e.g. 6 variables with 6 qubits)

✓ Process with the designed circuits (rotation) to obtain "$|\phi(\vec{x_j})> = \mathrm{U}_{\phi(\vec{x_j})}|0^{\otimes N}>$" and "$<\phi(\vec{x_i})|$"

✓ The similarity of event i and j can be obtained by $k(\vec{x_i}, \vec{x_j})$

- 50 events → 50 X 50 matrix elements

With space to optimize:



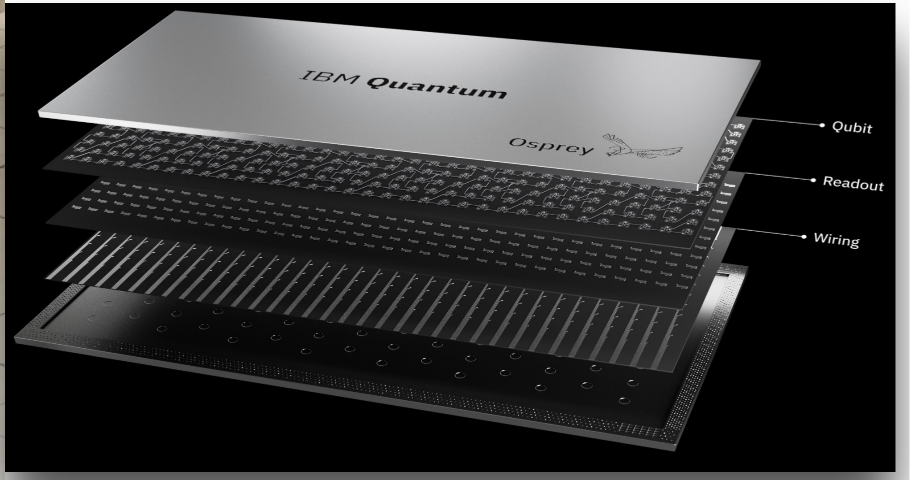
IBM quantum computer

Credited to Thomas Prior for TIME

- Slide bullet text

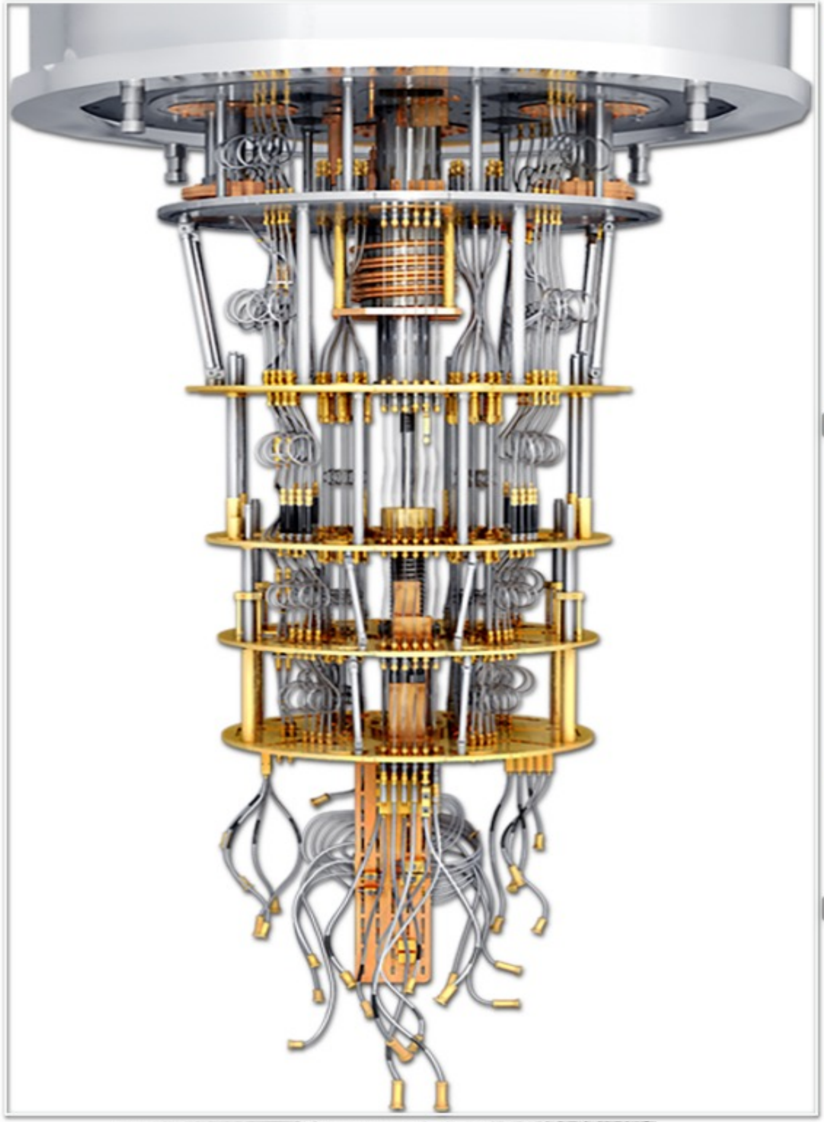
□ Taking quantum computing out of the lab:

- NY Computing Data Centre.
- It provides over 20 computers.
- Scales the processor's availability.
- It provides over 20 comput



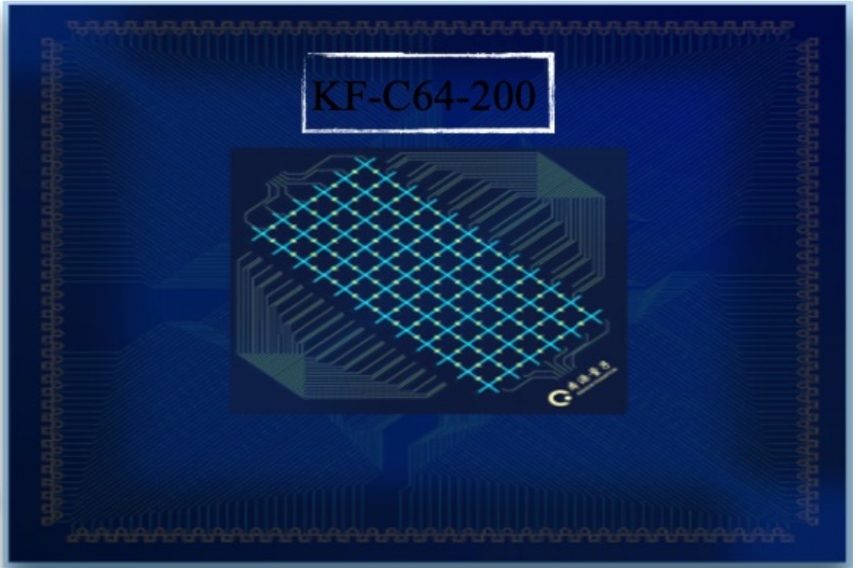
Origin quantum computer

China Domestic Quantum Computing Hardware



[Taken from Origin web page](#)

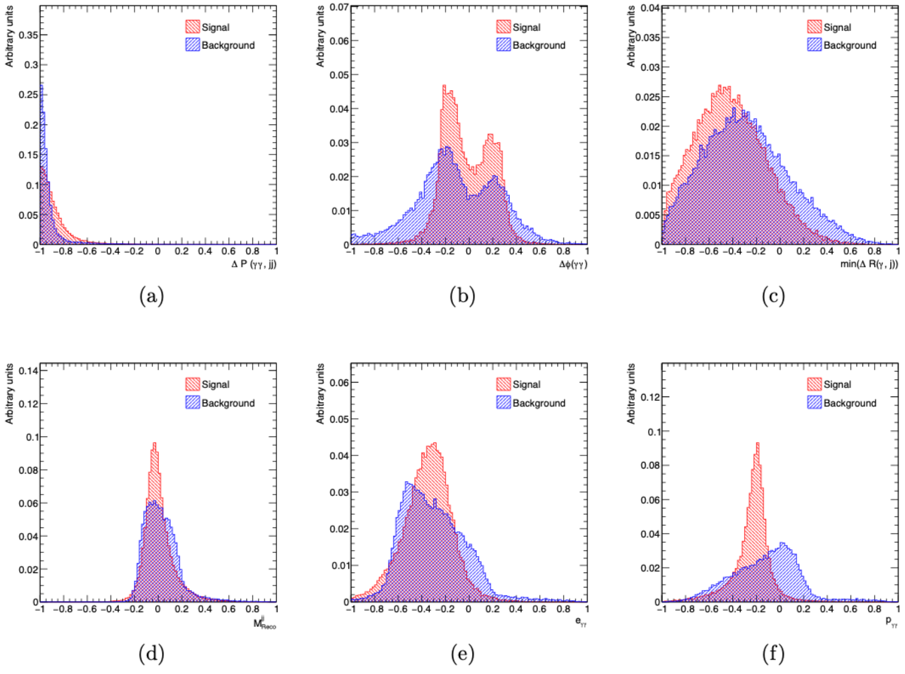
- Origin launched 64-qubit QPU:
  - Single-qubit gate fidelity > 99.9%.
  - Double-qubit gate fidelities > 98%.
  - Readout fidelities > 96%.

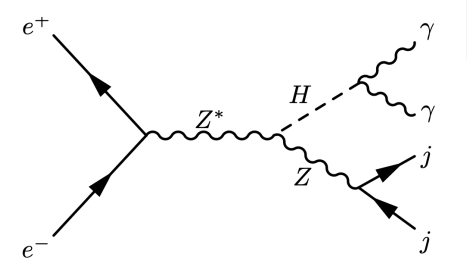


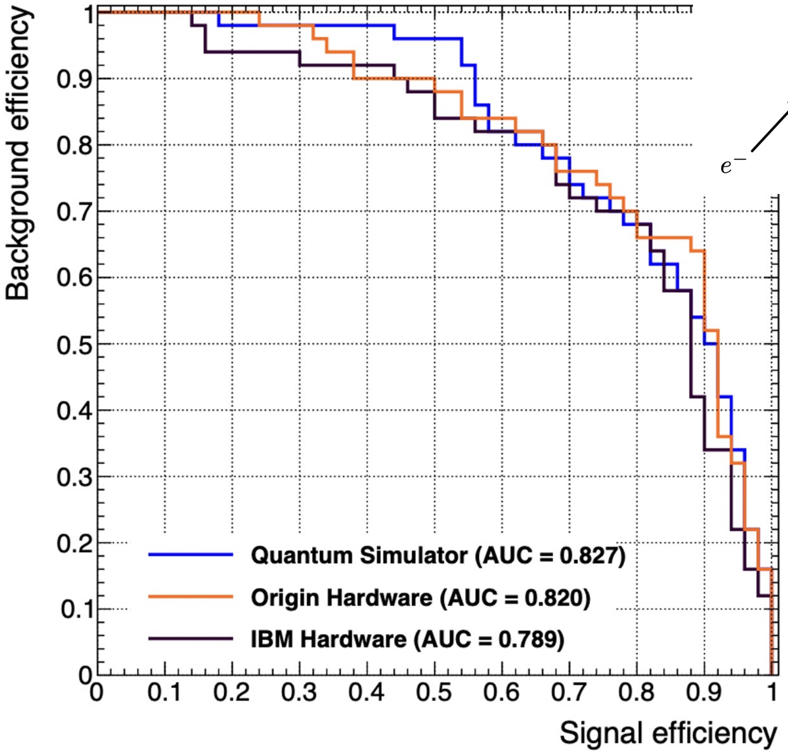
- Origin quantum computer provides up to 6 qubits for free. However, another hardware called Quafu provides up to 136 qubits. The Beijing Academy of Quantum Information Sciences maintains it.

QSVM Kernel with IBM and OriginQ hardware for CEPC ZH ($H \rightarrow \gamma\gamma$) analysis

International Journal of Modern Physics A, 2024, 2450007



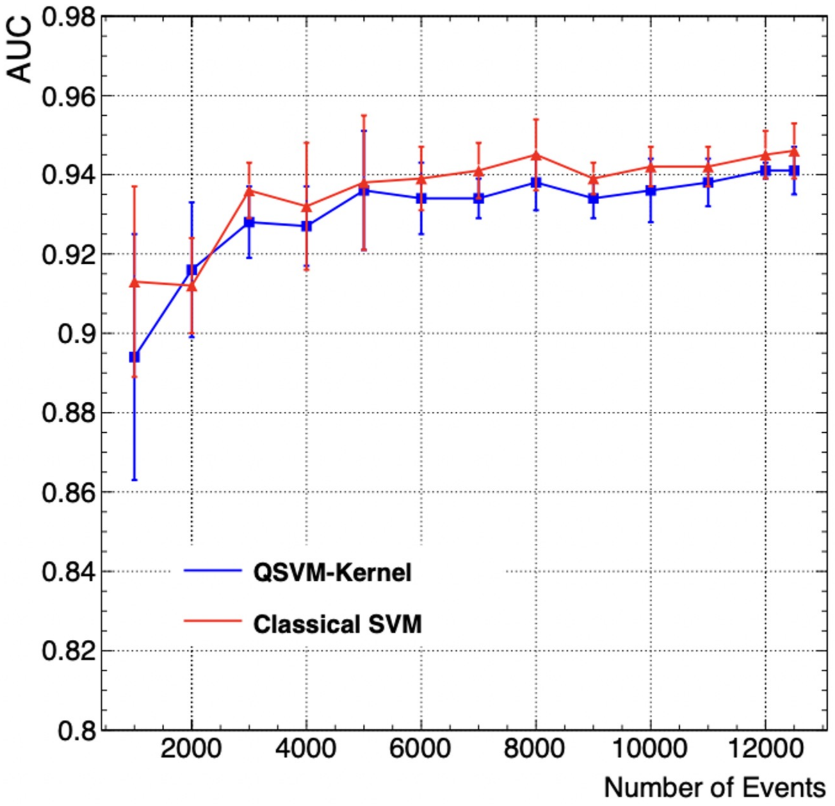




- ➢ Using ZH analysis dataset (100 events, 6 variables), the QSVM Kernel results on the Quantum Hardware (6 qubits) are promising
- ➢ It is the first attempt using Chinese domestic quantum hardware resource to implement QML

Employing Quantum SVM Kernel method with quantum simulators for ZH ($H \rightarrow \gamma\gamma$) analysis

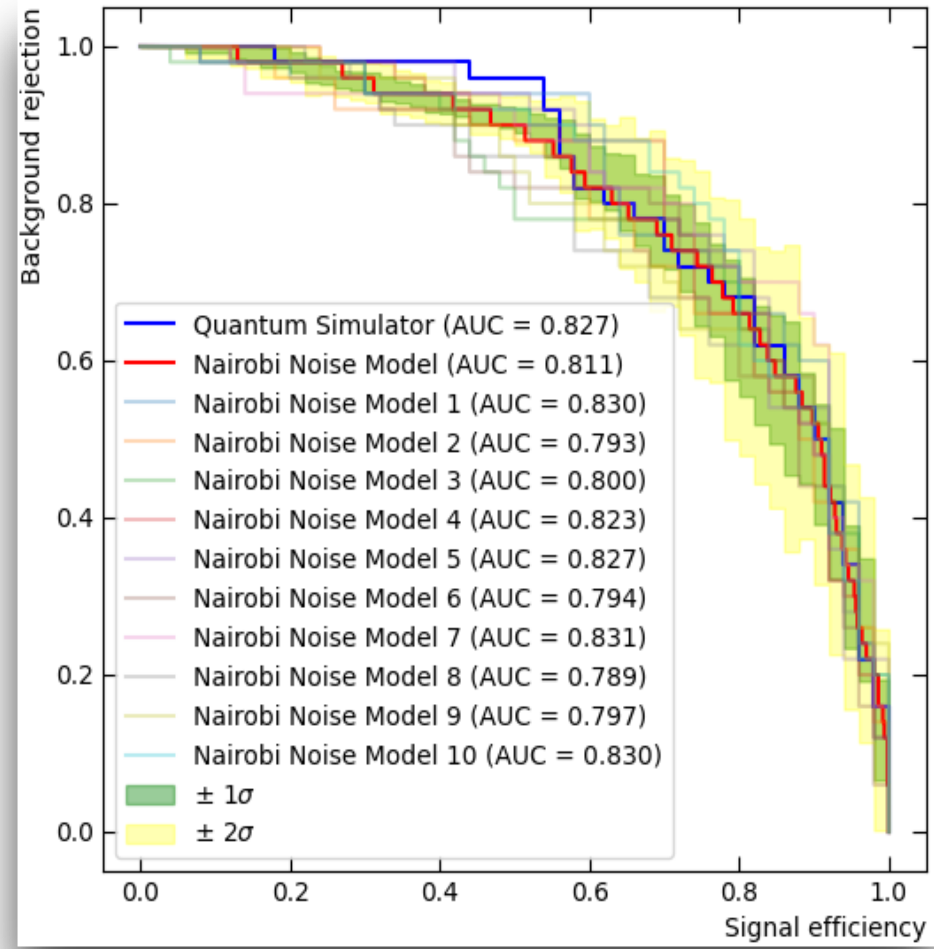
AUC vs number of events



- **QSVM Kernel method and noiseless simulators enable us to work with a larger number of events.**

Noise modelling in the IBM Nairobi

- ✓ <u>Noise in quantum computers:</u> Quantum computers are susceptible to noise.
  - ✓ An electromagnetic signal coming from a WiFi
  - ✓ A disturbance in the Earth's magnetic field
- ✓ The model used automatically generates a simplified noise model for a real device.
- ✓ It takes into account the following:
  - ✓ The gate error probability of each basis gate
  - ✓ The gate length of each basis gate
  - ✓ $T_1 and T_2$ relaxation time constant
  - ✓ The readout error probability
- ✓ The estimated noise in the IBM Nairobi computer is 0.017.

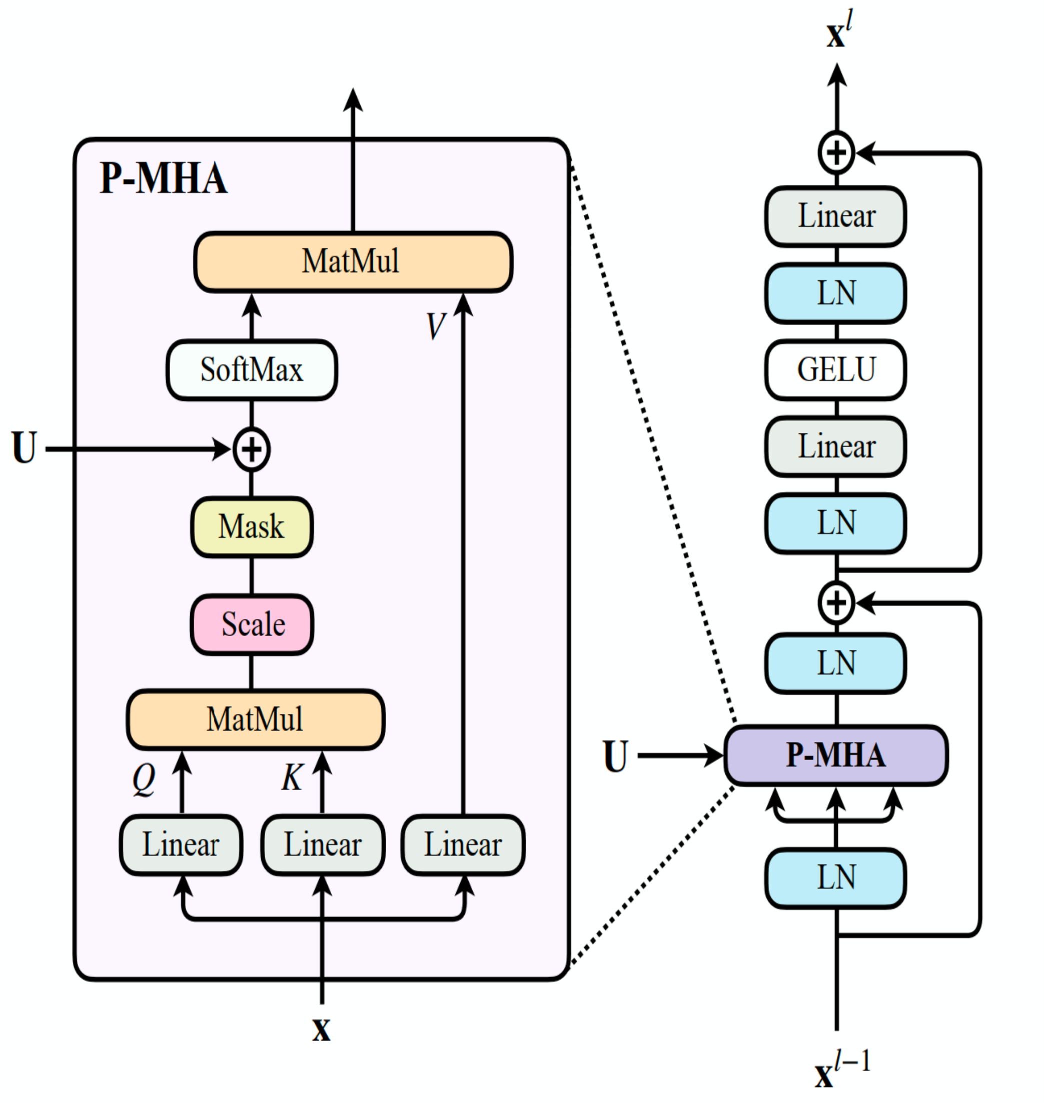


The Quantum Transformer

- ➢ Simplified version of the transformer
  - ✓ Multi-head Attention based on PyTorch
  - ✓ Three different linear transformations:
    - $W_Q, W_K, \textbf{\textit{and}}, W_V$

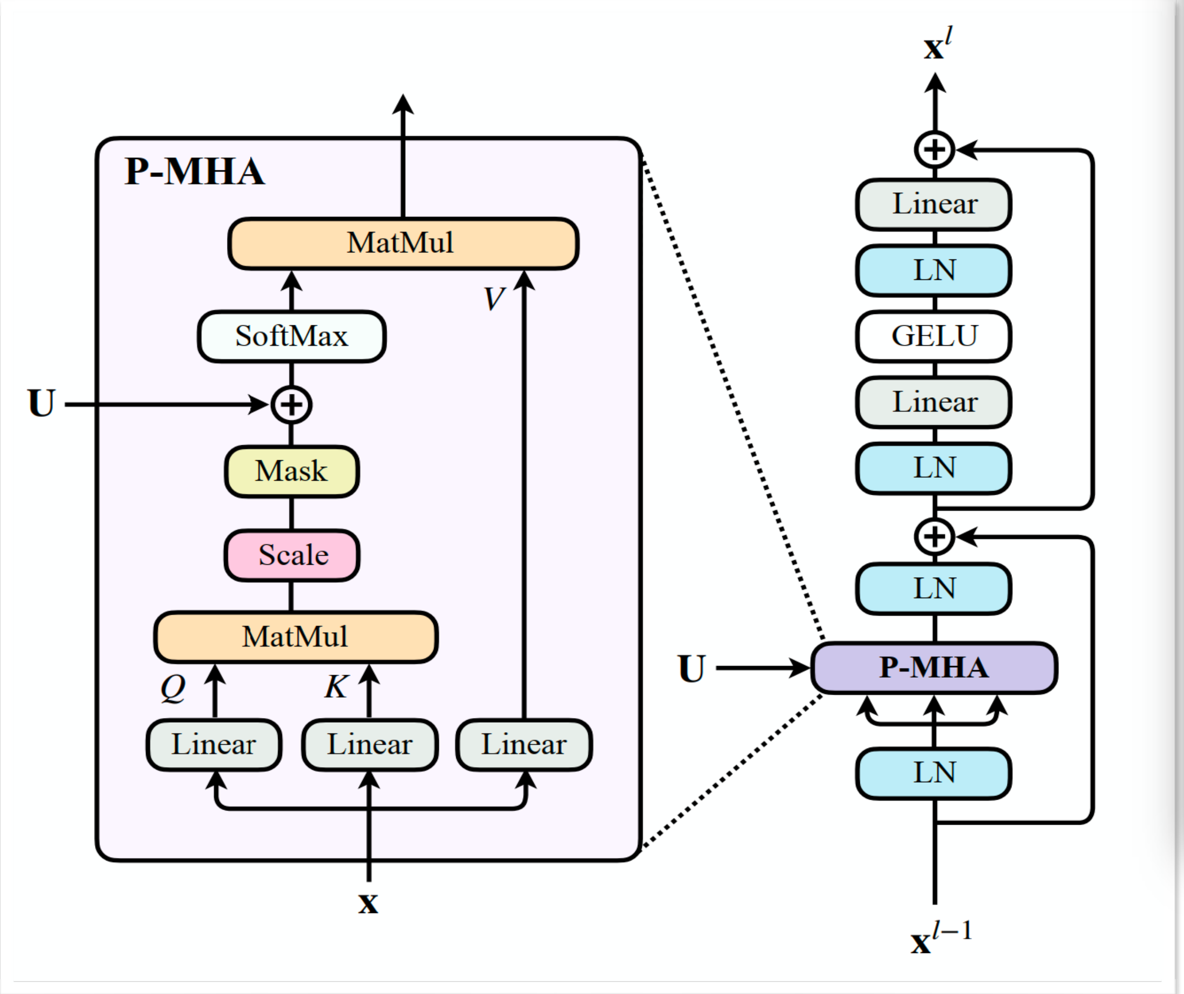
$$\boldsymbol{P\text{–}MHA}(Q,K,V) = \boldsymbol{SoftMax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \cdot V$$

  - ✓ **Where Q, K, and V are linear projections of the input.**
- ➢ Going to the quantum version, one could replace the linear transformation with a quantum one.

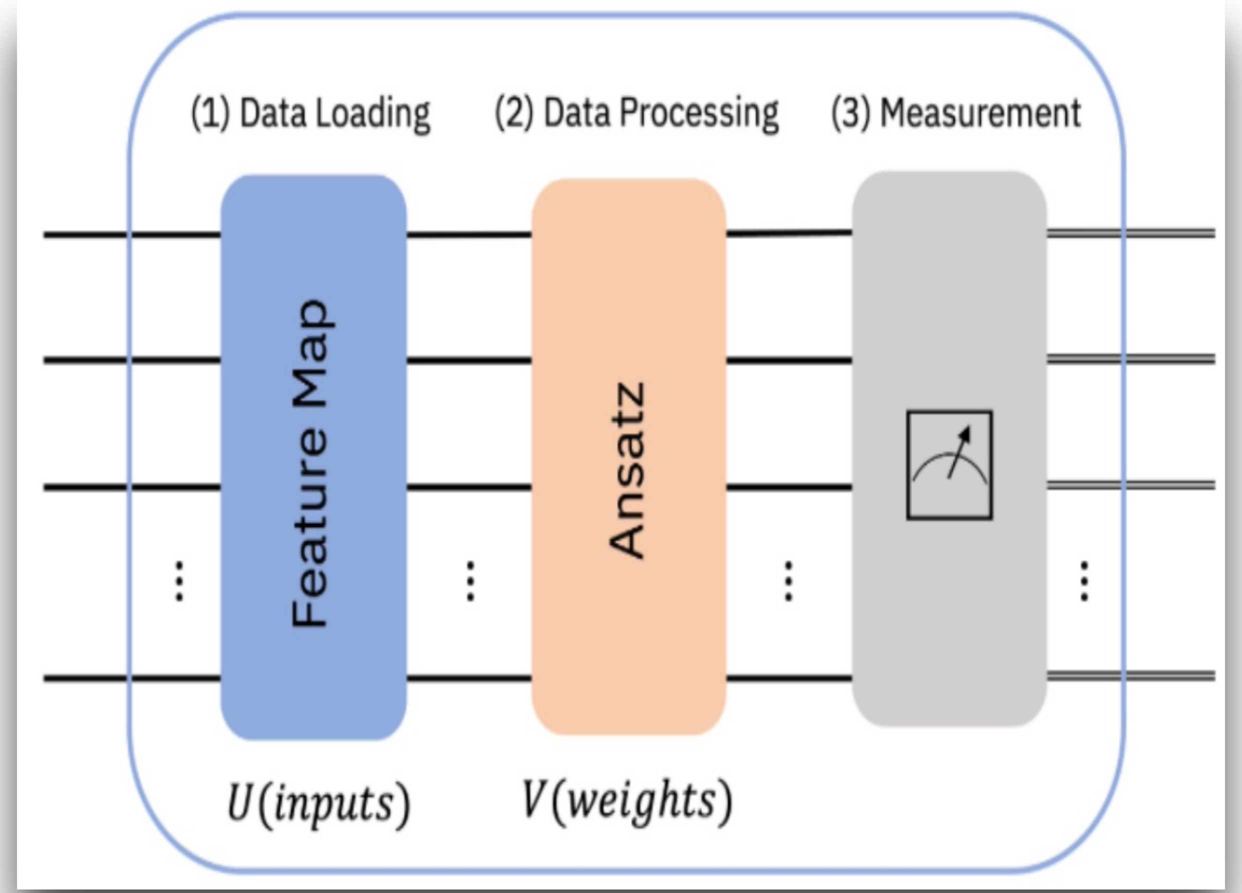


The Quantum Transformer

- ☐ Implementing quantum layer instead



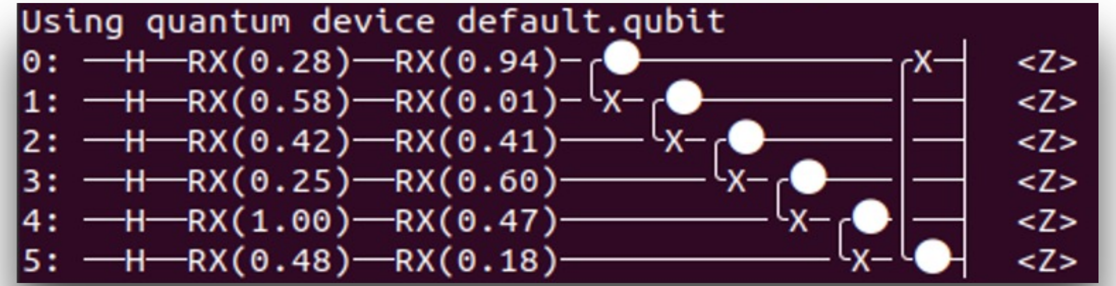
Quantum Neural Network



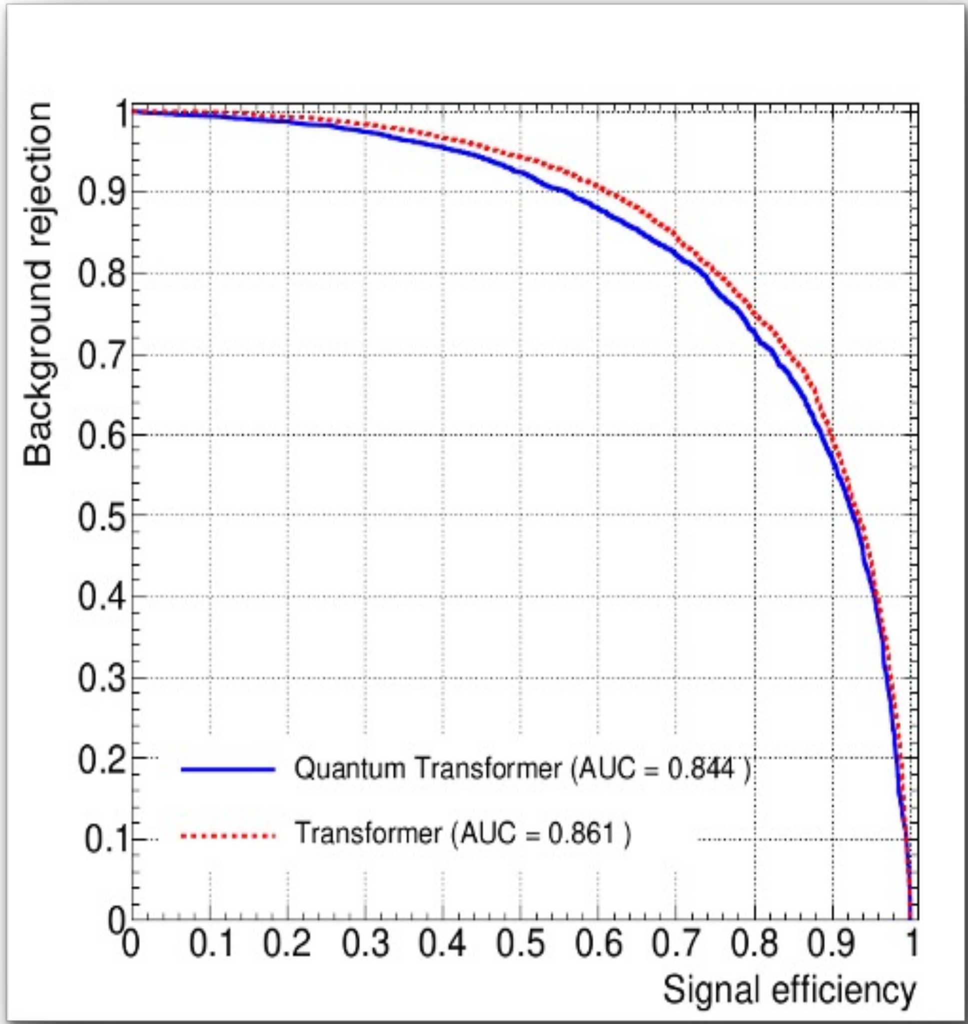
The trainable parameters are added using the Anzatz with a feature-map that acts as an encoder.

The Quantum Transformer

- We are using PennyLane, which was developed by Xanadu.



- Running on CPUs: about 76% accuracy on validation dataset both in Quantum transformer and classical transformer in 20k dataset (10k train, 10k val).
- Time-consuming: about 80 mins for a 10k dataset with one epoch and one block(Quantum layer).



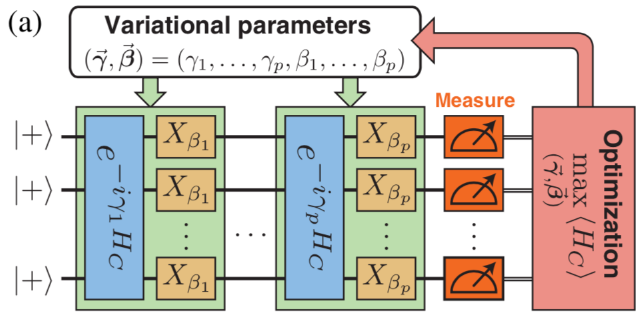
Application of Quantum Approximate Optimization Algorithm (QAOA) on Jet Clustering in CEPC

QAOA: a hybrid quantum-classical algorithm

1, cost Hamiltonian: $\hat{H}_C = -\frac{1}{2}\sum_{(i,j)\in E} w_{ij}(I - Z_i Z_j)$, **mixer Hamiltonian:** $\hat{H}_M = \sum_{j\in V} X_j$,

initial the quantum circuit in the state: $|s\rangle = |+\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}}\sum_{x\in(0,1)^n}|x\rangle$,

I**: identity operator,** $Z_i(X_j)$**: Pauli-Z(-X) operator acting on the i(j)-th qubit,**
n: the number of qubits

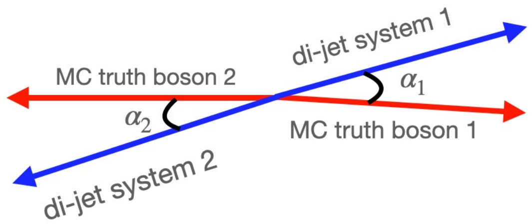


2, output state by the circuit: $|\psi_P(\gamma,\beta)\rangle = e^{-i\beta_P \hat{H}_M} e^{-i\gamma_P \hat{H}_C} \ldots e^{-i\beta_1 \hat{H}_M} e^{-i\gamma_1 \hat{H}_C}|s\rangle$

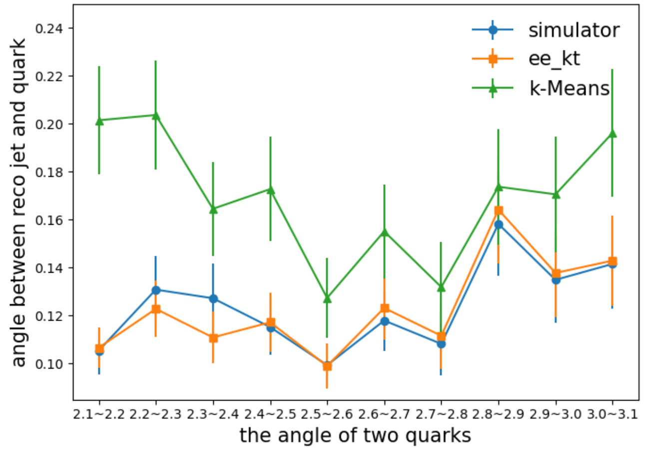
and calculate $F_P(\gamma,\beta) = \langle\psi_P(\gamma,\beta)|\hat{H}_C|\psi_P(\gamma,\beta)\rangle$.

3, A classical optimization algorithm is employed to $(\gamma^*,\beta^*) = \underset{\gamma,\beta}{argmax} F_P(\gamma,\beta)$.

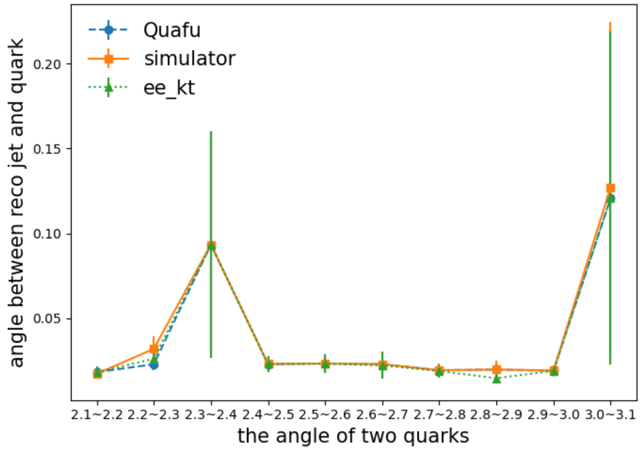
4, The state $|\psi_P(\gamma^*,\beta^*)\rangle$ **encodes the solution to the optimization problem.**



Performance, criterion is $\alpha = \alpha_1 + \alpha_2$



$H \to s\bar{s}$ with 30 final state particles
compare quantum simulator, ee_kt, and k-Means algorithms.



$H \to s\bar{s}$ with 6 final state particles
compare quantum hardware, simulator, and ee_kt

Conclusion

- The rapid developments of the large model in AI industry demand the corresponding increase of the computing powers.
- Quantum computer implemented on QAI could provide the potential matching the need:
  - ✓ **Superposition could help**
- We did the study of the implementing QSVM in CEPC physics analysis with domestic quantum computer resources.
  - ✓ **Comparable performances as SVM have been seen**
- The QTransformer has been developed and implemented in phsics analysis in CEPC as well.
  - ✓ **Will try the generative model with QTransformer in CEPC physics analysis.**
- Look forward to having some collaborations with colleagues in France.