Application of Quantum Machine Learning to HEP Analysis at LHC using IBM Quantum Computer Simulators and Hardware

Sau Lan Wu and Chen Zhou

Jay Chan, Wen Guan, Shaojun Sun, Alex Wang, Sau Lan Wu, Chen Zhou
Physics Department, University of Wisconsin-Madison

Miron Livny
Computer Sciences Department, University of Wisconsin-Madison

Federico Carminati, Alberto Di Meglio
CERN Openlab, IT Department, CERN

Panagiotis Barkoutsos, Ivano Tavernelli, Stefan Woerner, Jennifer Glick
IBM Research Zurich and IBM T.J. Watson Research Center

Andy Li, Joseph Lykken, Panagiotis Spentzouris
Quantum Institute, Fermilab

Samuel Yen-Chi Chen, Shinjae Yoo
Computational Science Initiative, BNL

Tzu-Chieh Wei
C.N. Yang Institute for Theoretical Physics, State University of New York at Stony Brook

July 28, 2020 - Prague, Czech Republic
40th International Conference on High Energy Physics (ICHEP)

Machine learning and quantum computing

- Machine Learning has become one of the most popular and powerful techniques and tools for High Energy Physics (HEP) data analysis
  - It greatly enhances our ability to identify rare signal against immense backgrounds (important for discovery of new physics)
- Issues raised by machine learning
  - Heavy CPU time is needed to train complex models
    - With more data, the training time increases very quickly
  - May lead to local optimization, instead of global optimization
- Quantum computing
  - A way of parallel execution of multiple processes using Qubits
  - Can speed up certain types of problems effectively
  - It is possible that quantum computing can find a different, and perhaps better, way to perform machine learning.

Our program with IBM Qiskit

Our Goal:

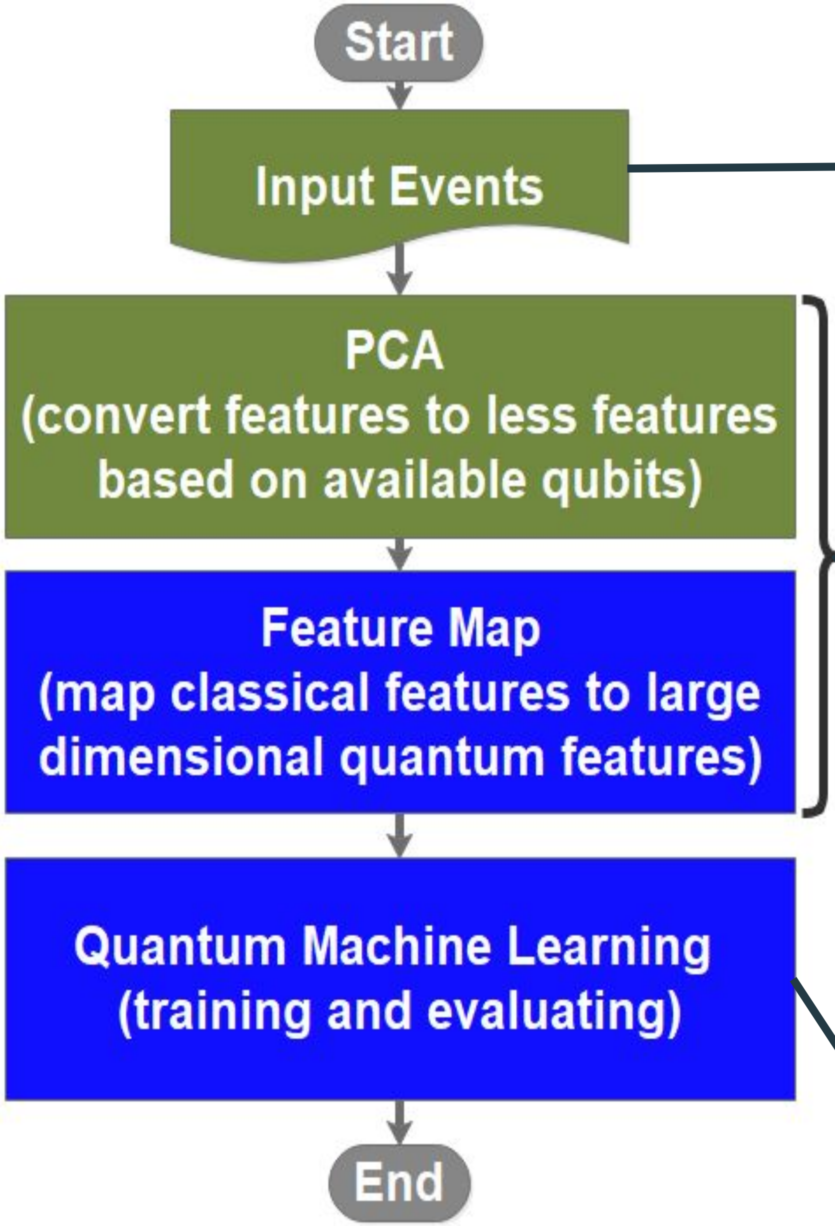
Perform LHC High Energy Physics analysis with Quantum Machine Learning, to explore and to demonstrate the potential of quantum computers can be a new computational paradigm for big data analysis in HEP

Our preliminary program is to:

Employ the Quantum Support Vector Machine (QSVM) method with the IBM gate-model quantum computers and the IBM Qiskit environment to LHC High Energy Physics analysis, for example ttH ($H \rightarrow \gamma\gamma$) and $H \rightarrow \mu\mu$ (two LHC flagship analyses).

*** IBM Qiskit = IBM Quantum Information Science Kit**

Our Workflow for Quantum Machine Learning



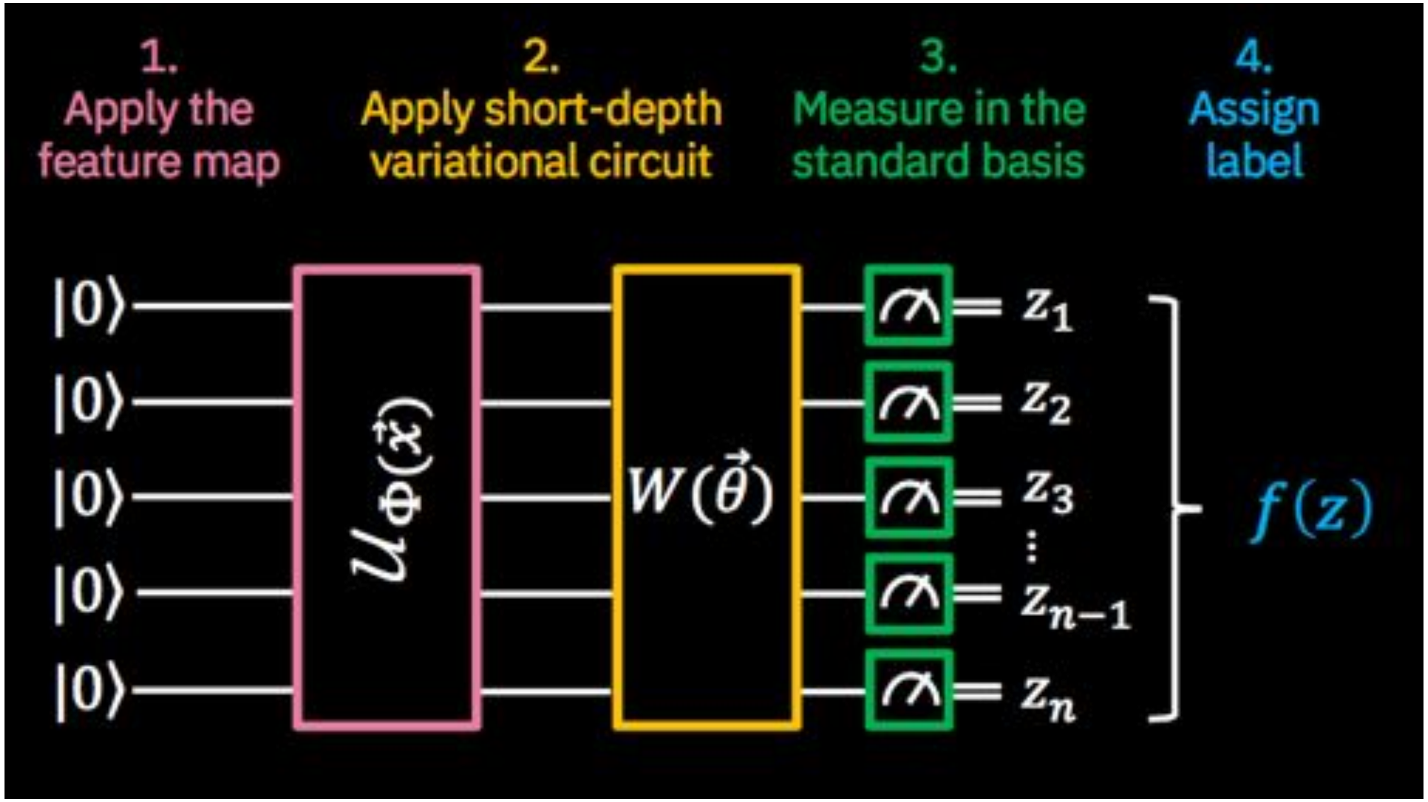
e.g. in ttH (H→γγ) analysis,
number of features (variables) = 23

- **PCA: Principal Component Analysis method is used to convert/combine features into fewer features to match the number of qubits.**
- **FeatureMap: encoding classical variables to quantum states. Currently, number of encoded variables has to equal number of qubits (e.g. 5, 10, 20)**

Quantum Support Vector Machine (QSVM) variational method, for example

Variational Quantum SVM method

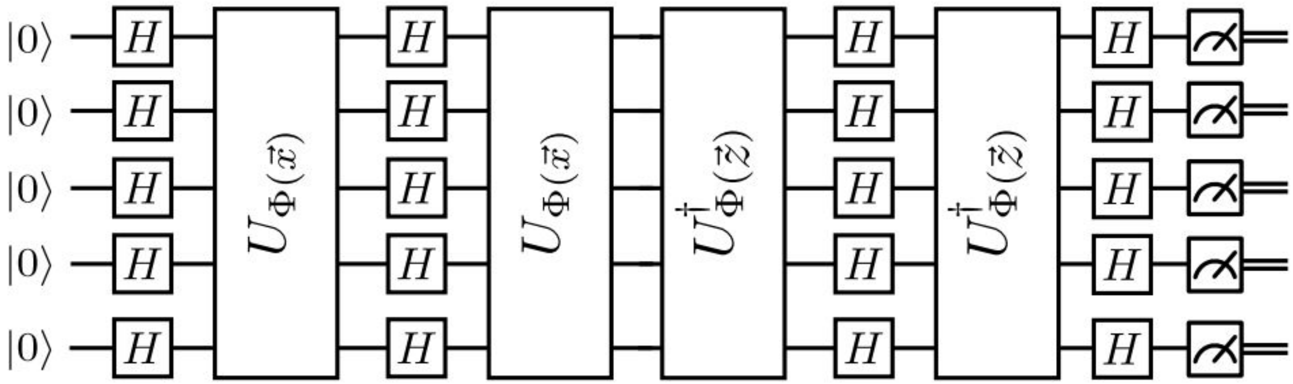
- In 2018, a variational Quantum SVM method was introduced by IBM, published in Nature 567 (2019) 209. The variational Quantum SVM method can be summarized in four steps:



- During the training phase, a set of events are used to train the circuit W(θ) to reproduce correct classification

Kernel Quantum SVM method

- ***Kernel Trick for Classical SVM**: mapping the non-linear separable data into a higher dimensional feature space using a kernel function that measures the similarity between two data points; then using the kernel to find a separating hyperplane.*
- ***Quantum Kernel Estimation** (introduced by IBM, published in Nature 567 (2019) 209): mapping classical data $\vec{x}$ non-linearly to a quantum state using Quantum Feature Map function; calculating the kernel matrix $K(\vec{x}, \vec{z})=|\langle\Phi(\vec{x})|\Phi(\vec{z})\rangle|^2$ using a quantum computer; then training the quantum SVM in the same way as a classical SVM.*

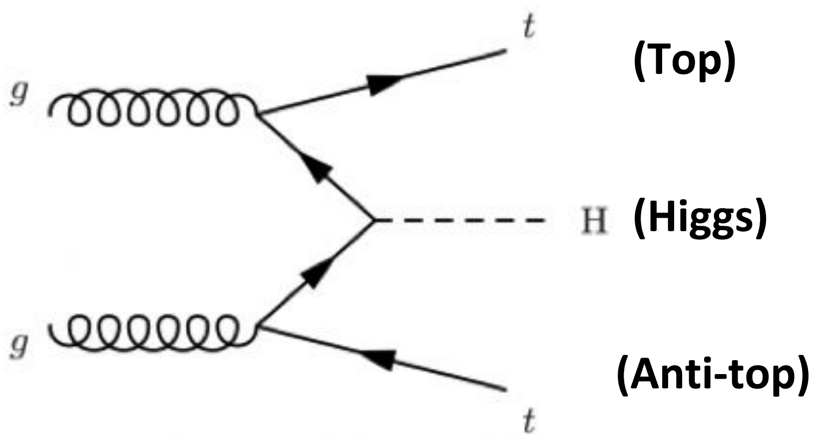


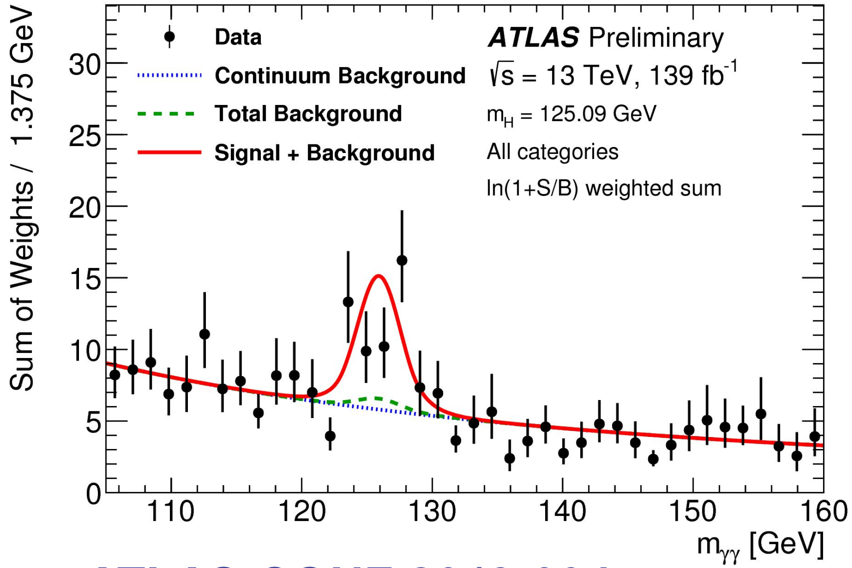
Nature 567 (2019) 209

Employing QSVM Variational to ttH ($H \rightarrow \gamma\gamma$) and $H \rightarrow \mu\mu$ analysis

ttH (H → γγ) analysis by the ATLAS Collaboration

The observation of ttH production (Higgs boson production in association with a top quark pair) by ATLAS and CMS directly confirmed the interaction between the Higgs boson and the top quark, which is the heaviest known fundamental particle



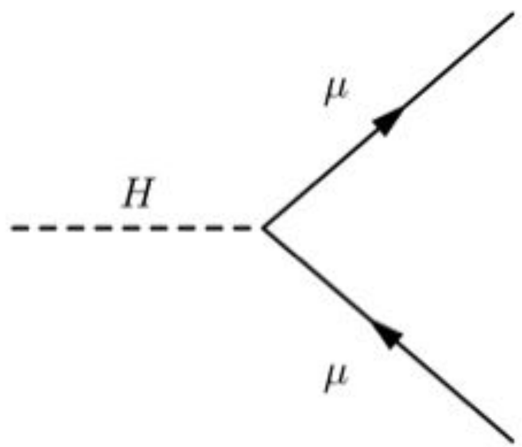


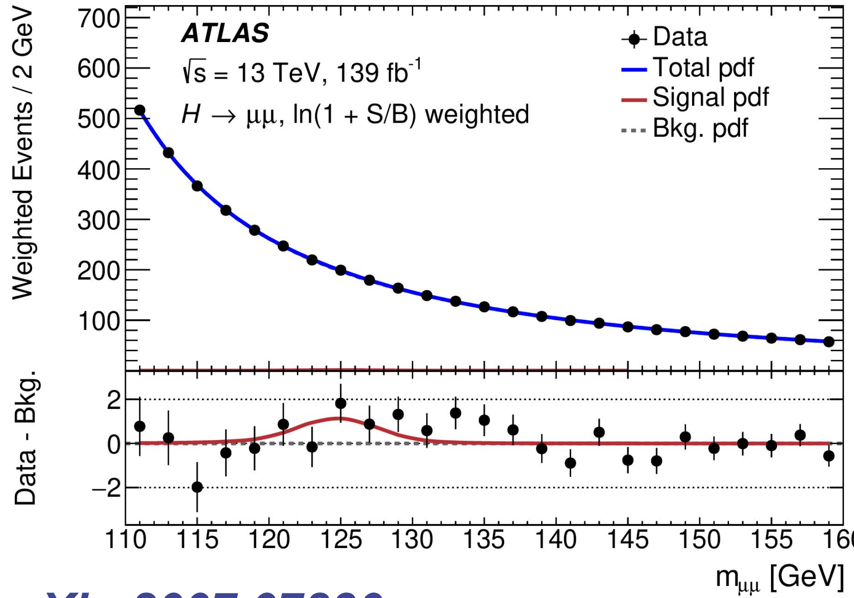
ATLAS-CONF-2019-004

- **Using Boosted Decision Tree (BDT, a classical machine learning technique) with XGBoost package, the ATLAS Collaboration observes the ttH (H→γγ) process**
- **This talk will perform the machine learning-based event classification of the ttH (H→γγ) analysis (hadronic channel) with delphes simulation samples and quantum machine learning**

H → $\mu\mu$ analysis by the ATLAS Collaboration

Although the coupling between the Higgs boson and 3rd-generation fermions has been observed, currently there is no evidence for the coupling between the Higgs boson and 2nd-generation fermions. H→$\mu\mu$ is the most promising process to observe such a coupling at the LHC)





arXiv:2007.07830 *(ATLAS submitted it to P.L.B)*

- **Using Boosted Decision Tree (BDT, a classical machine learning technique) with XGBoost package, the ATLAS Collaboration searches for the H→$\mu\mu$ decay**
- **This talk will perform the machine learning-based event classification of the H→$\mu\mu$ analysis (VBF channel) with delphes simulation samples and quantum machine learning**

Employing QSVM Variational with IBM Q simulator, ttH ($H \rightarrow \gamma\gamma$) analysis and $H \rightarrow \mu\mu$ analysis

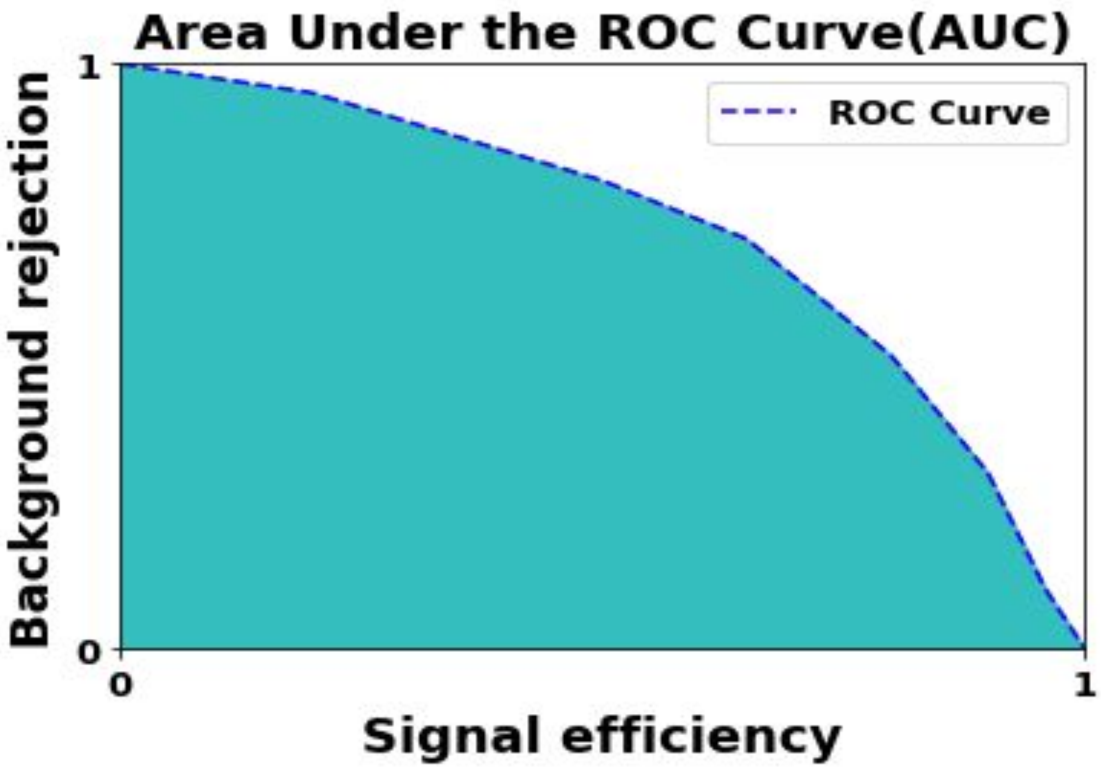
With 10 qubits, we successfully finished training and testing with 100 events with IBM Qiskit qasm simulator (where '100' events means 100 training events and 100 test events).

- **Q simulator (Quantum circuits simulator): here Qiskit Qasm simulator is used. Our simulation incorporates the hardware noise**
- **Quantum circuits are optimized to best fit the constraints imposed by hardware (e.g. qubit connectivity, hardware noise) and the nature of data**
- **SPSA optimizer is used with 1000 iterations***

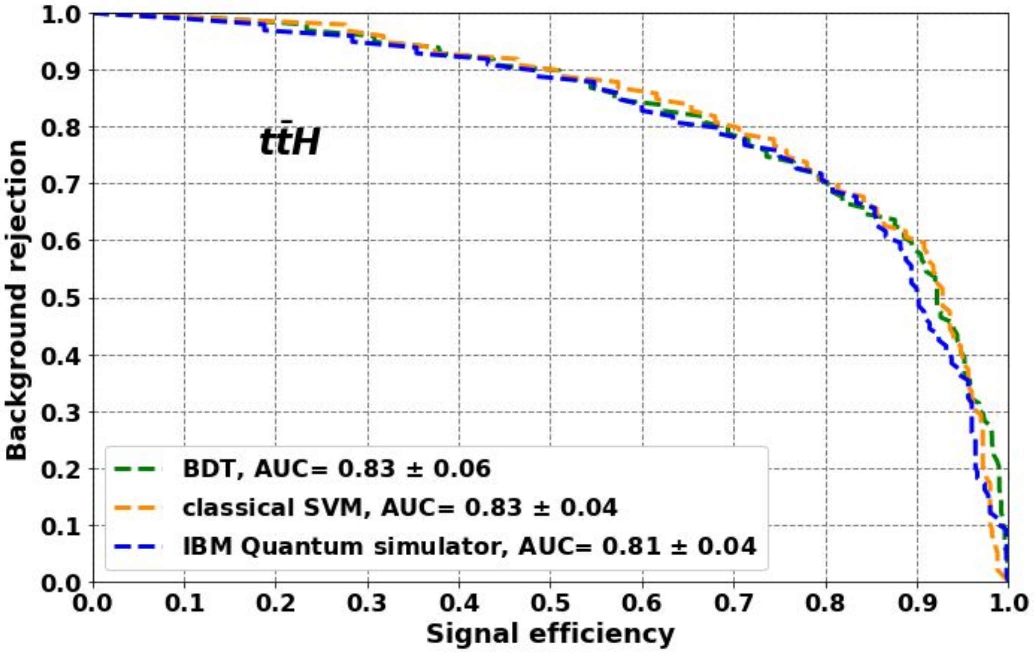
*** "iteration" indicates the number of times the algorithm's parameters are updated in training**

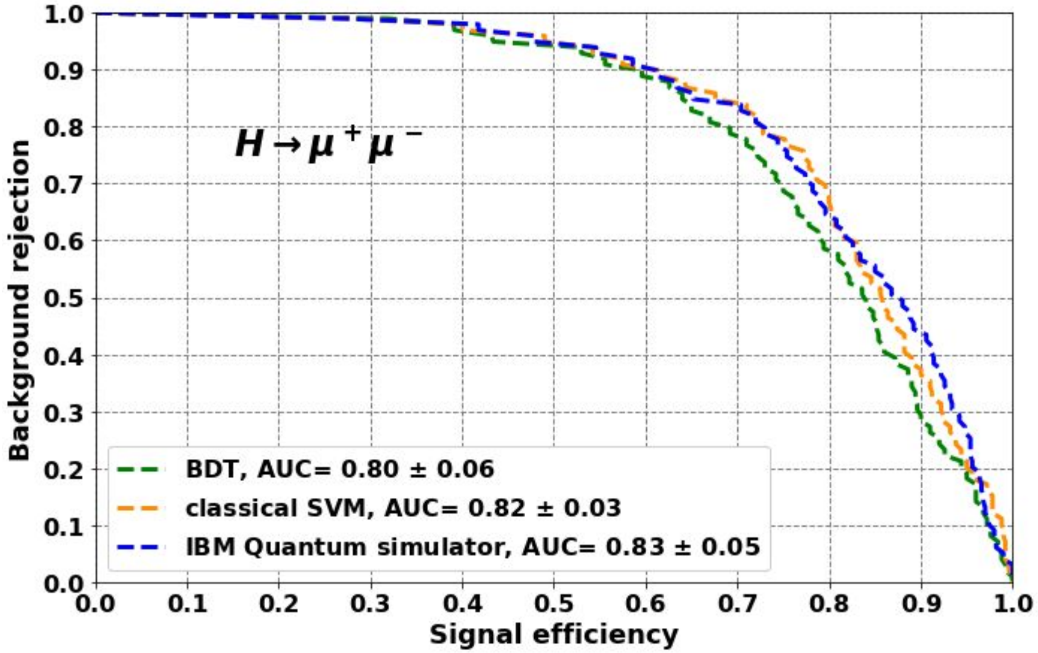
Employing QSVM Variational with IBM Q simulator, ttH ($H \rightarrow \gamma\gamma$) analysis and $H \rightarrow \mu\mu$ analysis

- **Definitions**
  - **ROC Curve**: a graph showing background rejection vs signal efficiency.
  - **AUC**: Area Under the ROC Curve



Employing QSVM Variational with IBM Q simulator, ttH ($H \to \gamma\gamma$) analysis and $H \to \mu\mu$ analysis



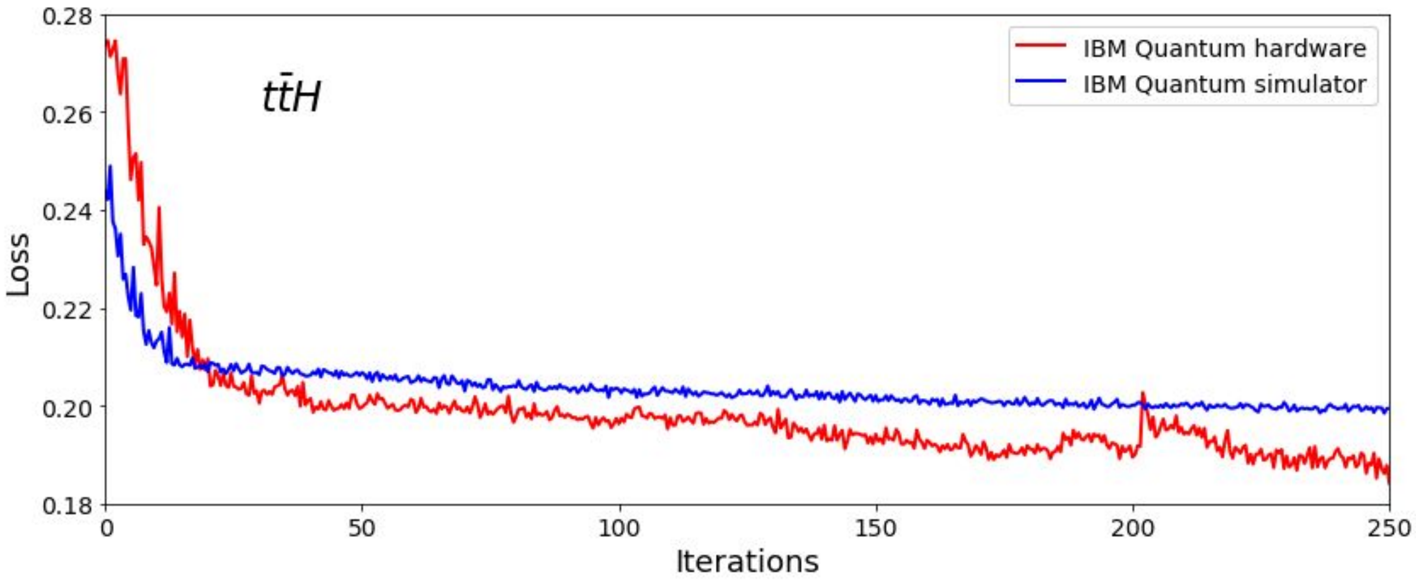


- Using ttH analysis dataset (100 events, 10 variables) and $H \to \mu\mu$ analysis dataset (100 events, 10 variables), QSVM Variational on simulator (blue) performs similarly with classical BDT (green) and classical SVM (yellow).

Employing QSVM Variational with IBM hardware (ibmq_boeblingen, a 20-qubit machine), ttH (H $\rightarrow \gamma\gamma$) analysis (ibmq_paris, a 27-qubit machine), H $\rightarrow \mu\mu$ analysis

- **With the help of IBM Research Zurich, Fermilab and BNL, we have finished multiple jobs on the IBM hardware based on superconducting electronic circuits (ibmq_boeblingen, a 20-qubit machine and ibmq_paris, a 27-qubit machine) with 100 training events and 100 test events using 10 qubits.**
- **For each analysis, due to current limitation of access time, we apply the QSVM variational method to one dataset on quantum hardware (rather than ten datasets on quantum simulator)**

Employing QSVM Variational with IBM hardware (ibmq_boeblingen, a 20-qubit machine), ttH ($H \rightarrow \gamma\gamma$) analysis (ibmq_paris, a 27-qubit machine), $H \rightarrow \mu\mu$ analysis



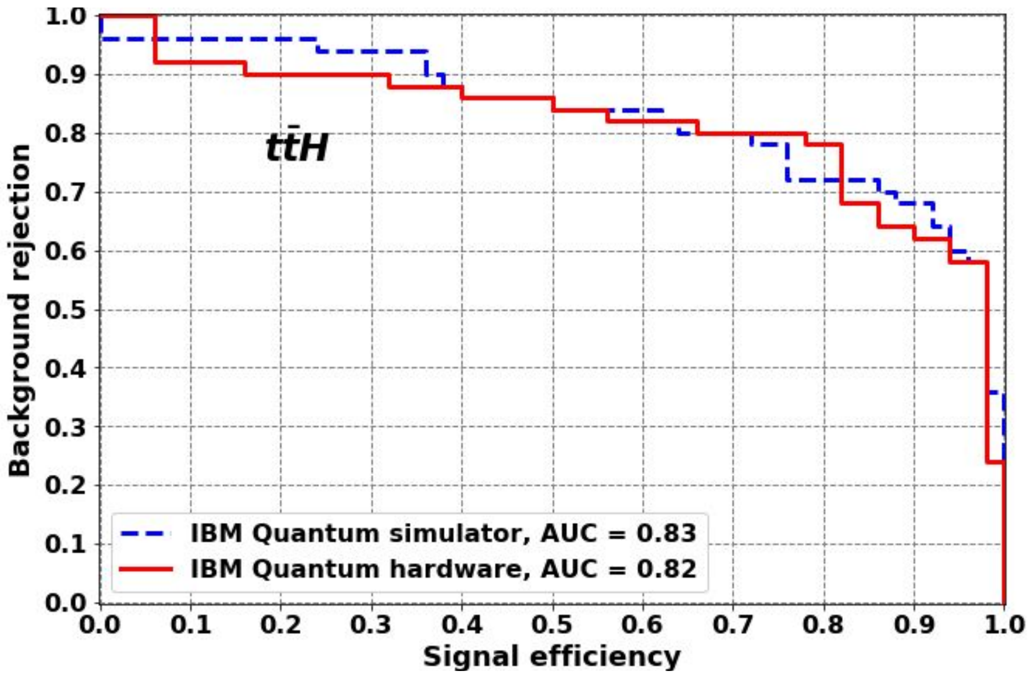
Blue: Quantum Simulation

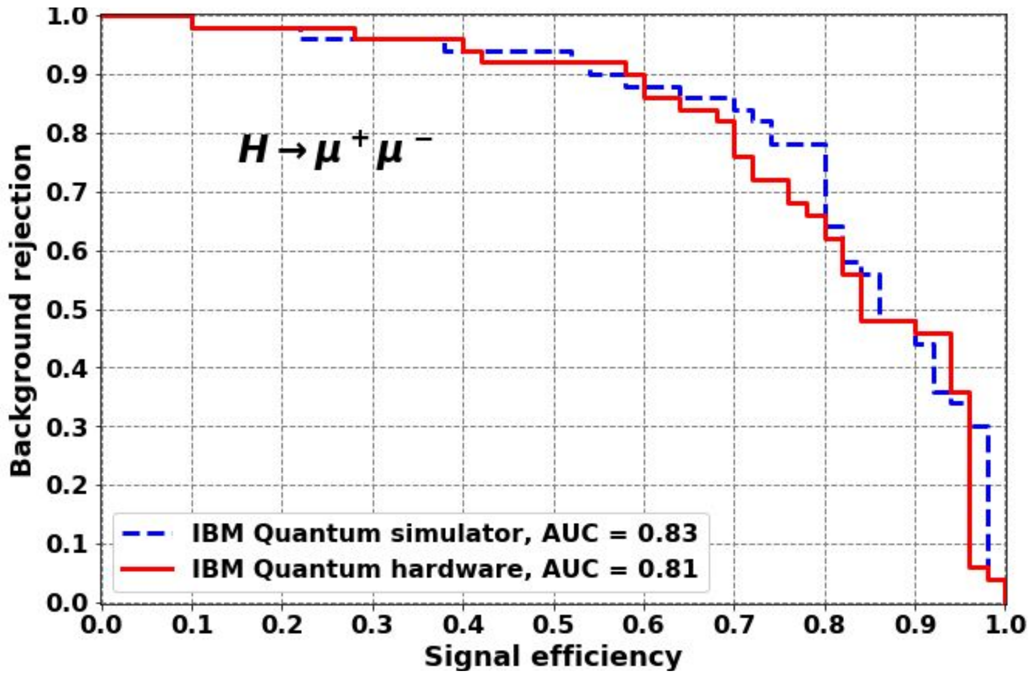
Red: Quantum Hardware

Loss: the mean of the squared differences between the output scores from the quantum algorithm and the ideal scores

- **The hardware loss (red) is decreasing with the increase of number of iterations. This indicates that the Quantum Computer has the ability to learn how to differentiate between the signal and the background for a HEP analysis.**

Employing QSVM Variational with IBM hardware (ibmq_boeblingen, a 20-qubit machine), ttH (H → $\gamma\gamma$) analysis (ibmq_paris, a 27-qubit machine), H → $\mu\mu$ analysis





hardware AUC = 0.82, simulator AUC = 0.83

hardware AUC = 0.81, simulator AUC = 0.83

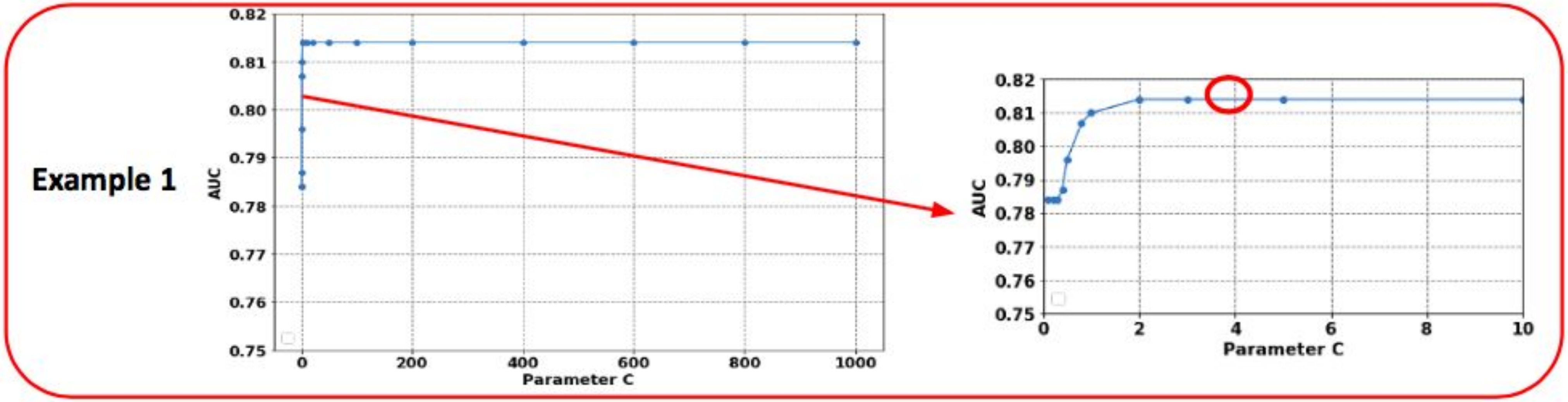
- **Using ttH analysis dataset (100 events, 10 variables) and H → $\mu\mu$ analysis dataset (100 events, 10 variables), with 250 iterations, the QSVM Variational on Quantum Hardware and QSVM Variational on quantum simulator results are in good agreement.**

Employing QSVM Kernel to ttH ($H \to \gamma\gamma$) analysis

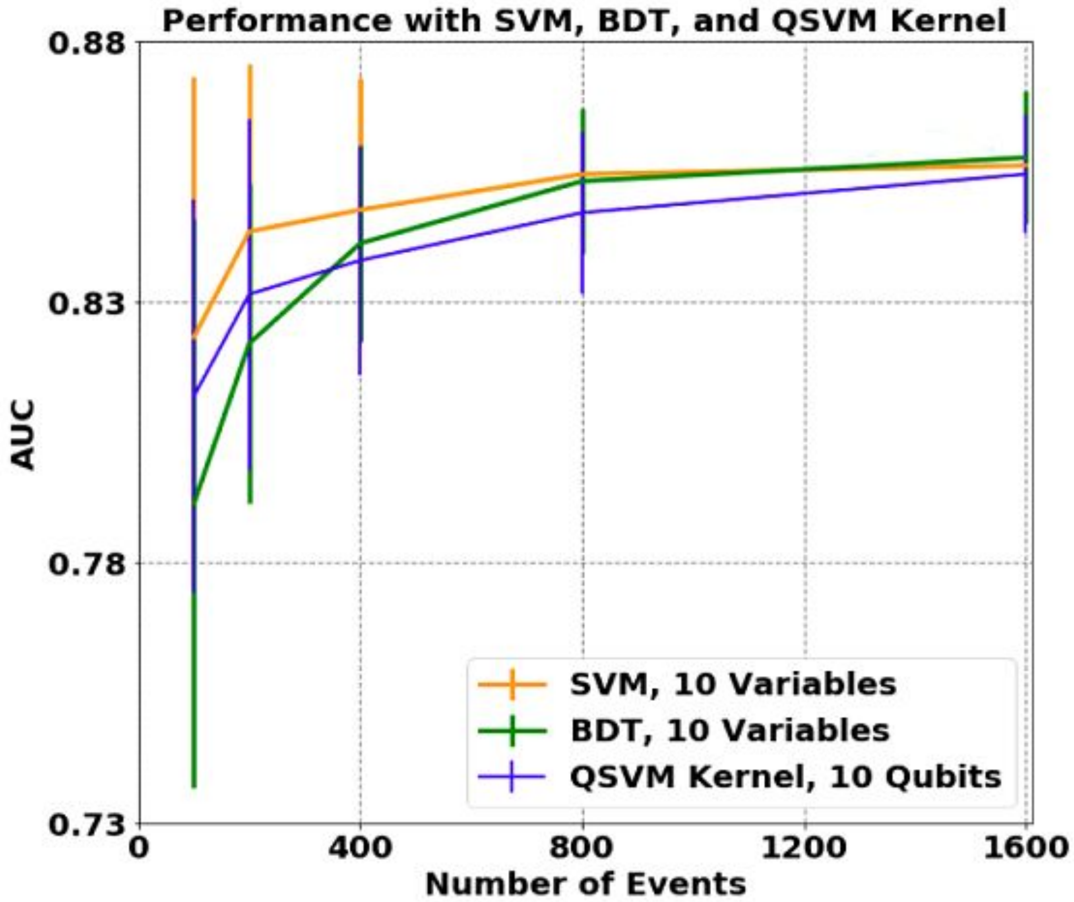
Employing QSVM Kernel with IBM Q simulator, ttH ($H \rightarrow \gamma\gamma$) analysis

We are performing the ttH analysis using QSVM Kernel method with 10 qubits:

- ***A customized FeatureMap is used***
- ***Grid-Search with cross-validation is used to optimize the SVM hyperparameters, for example, C (regularization parameter)***



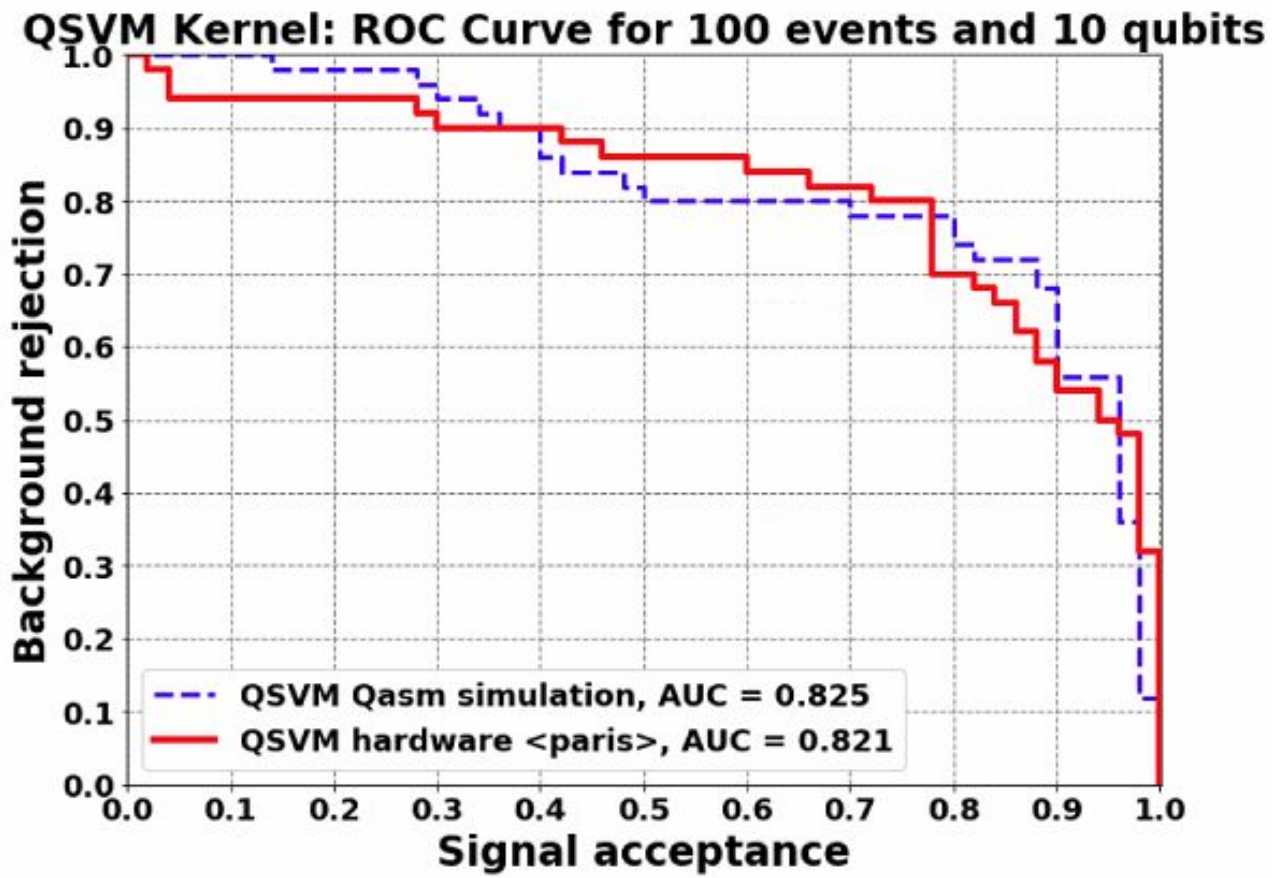
Employing QSVM Kernel with IBM Q simulator, ttH ($H \to \gamma\gamma$) analysis



- At the moment, an advantage of using QSVM Kernel method is that the running for both simulator and hardware are much faster. It enables us to work with a larger number of events. This was not feasible with the QSVM Variational method.

- Using ttH analysis dataset (100-1600 events, 10 variables), QSVM Kernel on simulator (blue) performs similarly with classical BDT (green) and classical SVM (yellow).

Employing QSVM Kernel with IBM hardware (ibmq_paris, a 27-qubit machine), ttH ($H \rightarrow \gamma\gamma$) analysis



hardware AUC = 0.82

simulator AUC = 0.83

- Using ttH analysis dataset (100 events, 10 variables), the discrimination power of the **QSVM Kernel on the Quantum Hardware** is currently similar to that of the **QSVM Kernel on quantum simulator**.

Summary

- Using IBM Quantum Computer simulators and hardware (20-qubit ibmq_boeblingen and 27-qubit ibmq_paris), we have employed Quantum Machine Learning (QSVM Variational and Kernel methods) to two LHC HEP flagship analyses (ttH ($H \rightarrow \gamma\gamma$) and $H \rightarrow \mu\mu$) with Delphes simulation events.
- With 100 events* and 10 qubits, QSVM variational method on Quantum Simulator perform similarly with classical BDT and classical SVM.
- With 100-1600 events* and 10 qubits, QSVM kernel method on Quantum Simulator perform similarly with classical BDT and classical SVM.
- With 100 events* and 10 qubits, for both variational and kernel methods, Quantum Hardware and Quantum Simulator show comparable performance

* '100' events means 100 training events and 100 test events

Summary-2

ttH (H->$\gamma\gamma$) 100 events, 10 qubits	QSVM Variational	QSVM Kernel
Quantum Simulator AUC	0.83	0.83
Quantum Hardware AUC	0.82	0.82

* Note QSVM Variational and QSVM Kernel runs are not using identical datasets

- Our results demonstrate quantum machine learning on the hardware of gate-model quantum computers has the ability to differentiate signal and background in realistic physics datasets
- We will investigate further optimization and hopefully we will see QSVM methods outperform classical machine learning
- Furthermore, future quantum computers might offer speedups in machine learning which could be critical for the HEP community