Quantum machine learning on CMS open data ~~for new physics search~~

Hwidong Yoo

Yonsei University
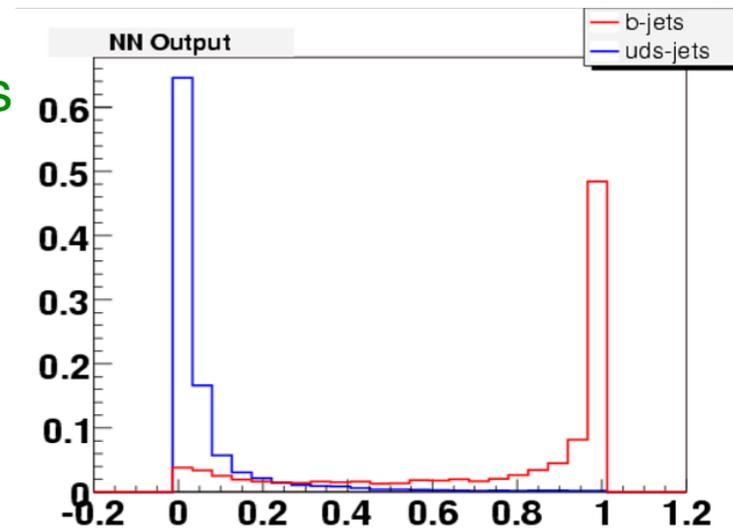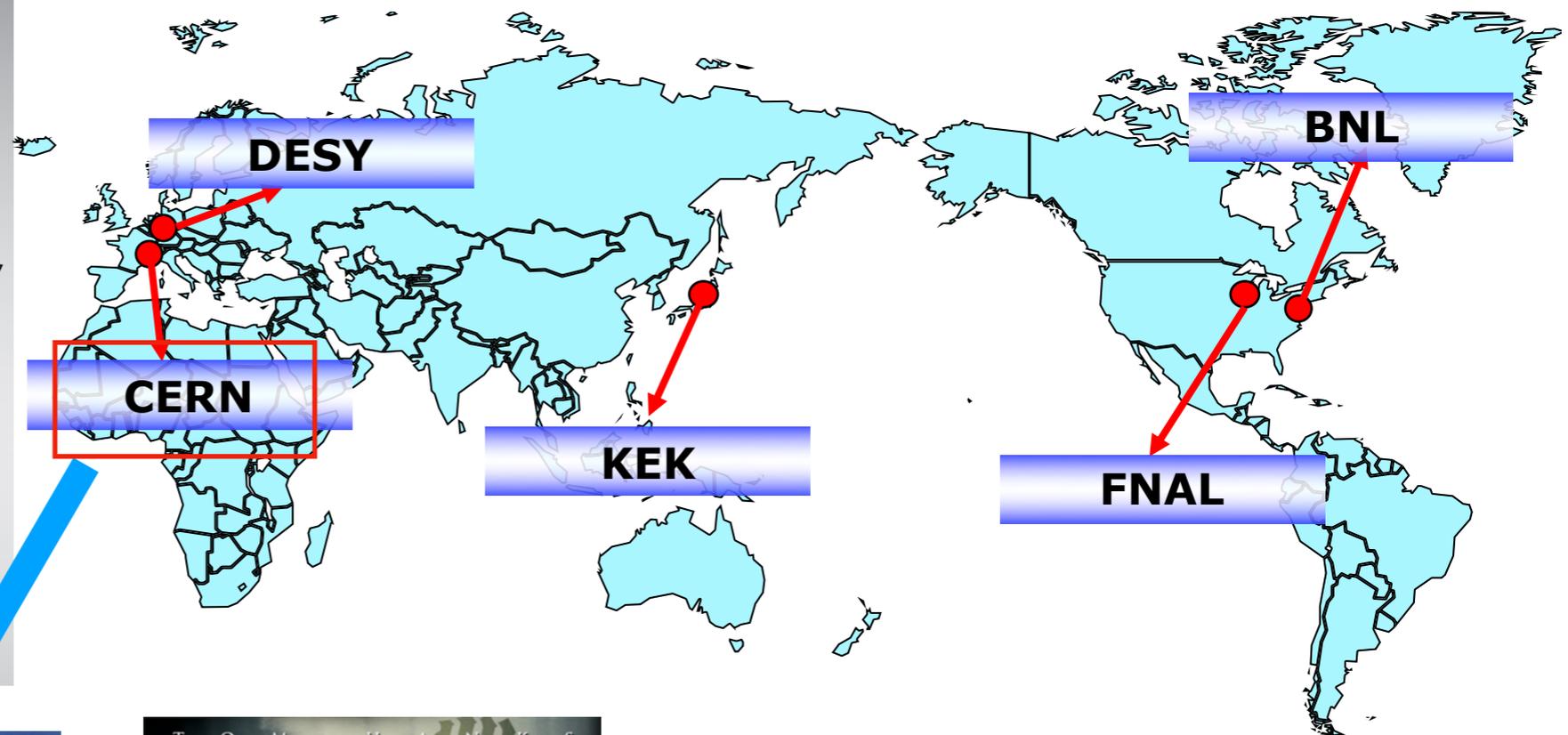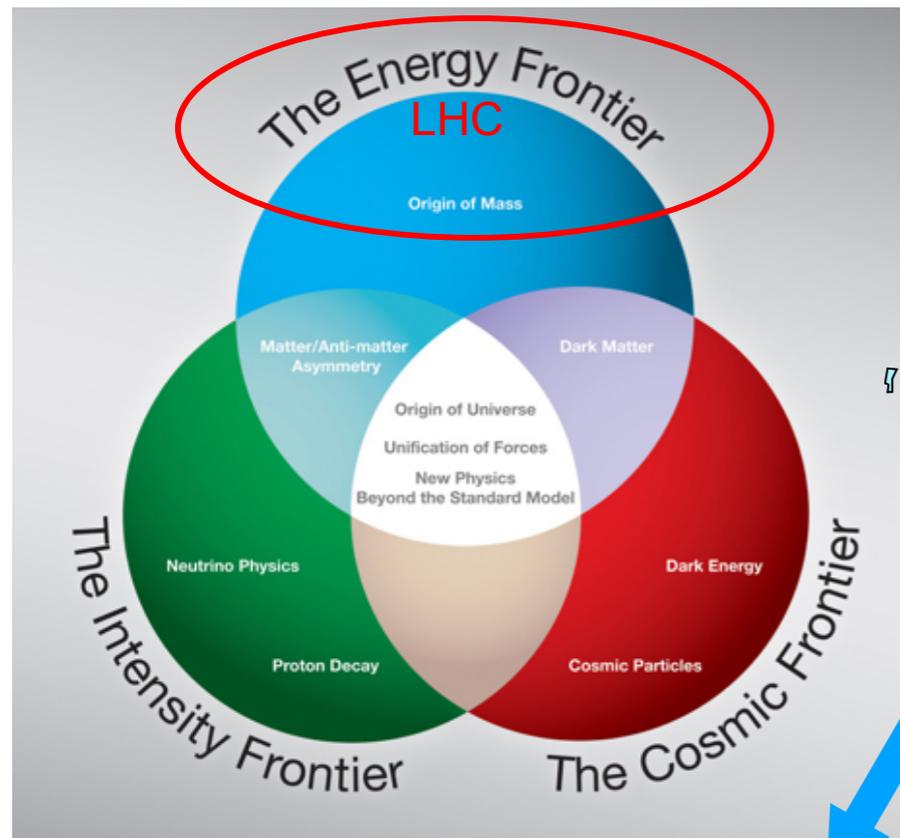
Who I AM ...

- I am experimental particle physicist on hadron collider

  - Since 2008, I have worked in CMS experiment

  - Since 2019, I am strongly involving the Future Collider projects

- I have various experiences on the data analysis

  - 14 CMS papers as a primary author during last 5 years

- I have various experiences on the ML application on the data analysis

  - But my technical knowledge is old-fashioned

- I am super beginner for quantum machine learning

  - I have recently learned several examples using IBM qiskit

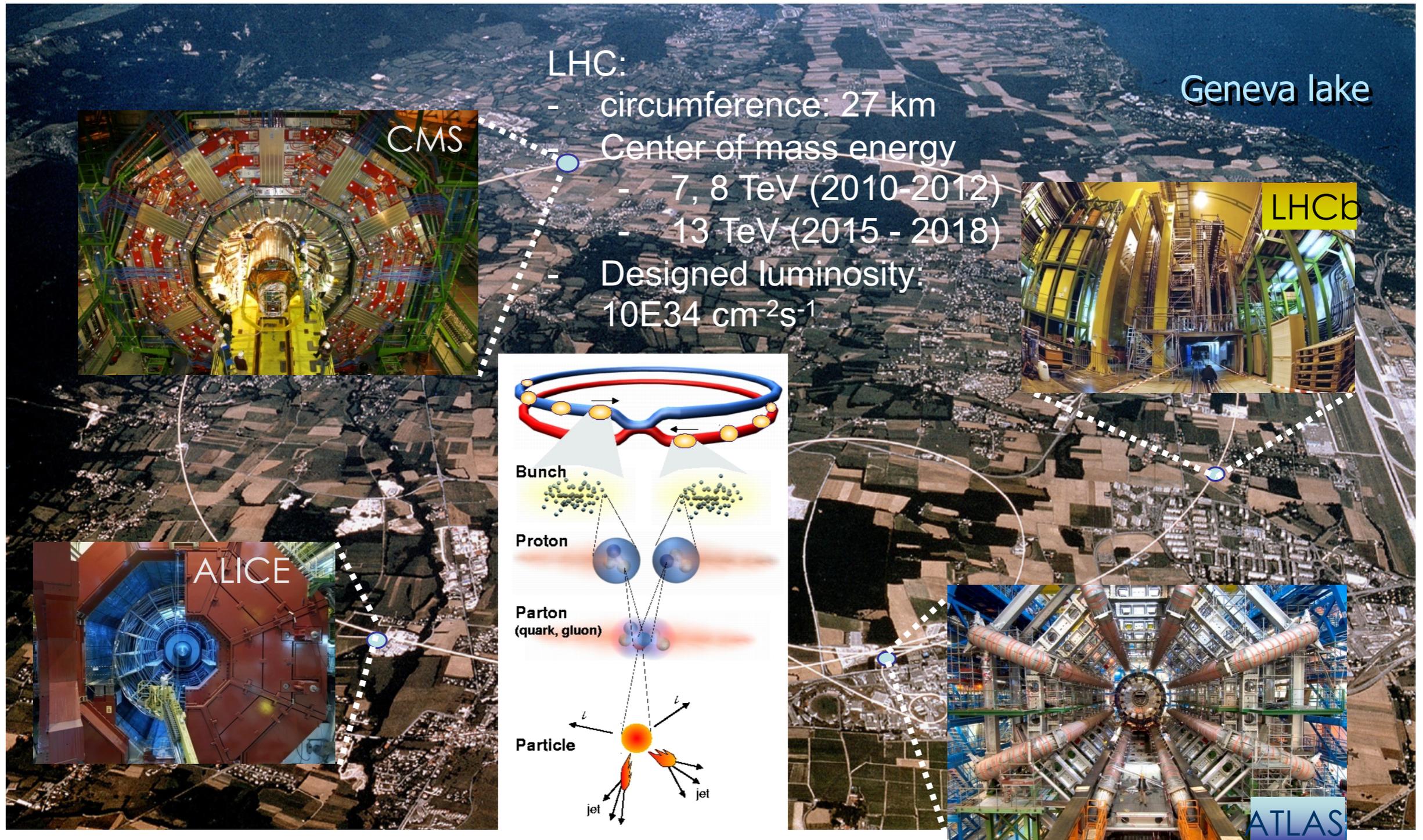 - Do NOT underestimate my poor knowledge/experience on the QML

From my Ph.D. thesis (2008)

Particle Physics in the World

The Energy Frontier — LHC — Origin of Mass
The Intensity Frontier
The Cosmic Frontier

Matter/Anti-matter Asymmetry — Dark Matter
Origin of Universe
Unification of Forces
New Physics Beyond the Standard Model
Neutrino Physics — Proton Decay — Cosmic Particles — Dark Energy

DESY
CERN
KEK
BNL
FNAL

Large Hadron Collider at CERN

CMS

LHC:
- circumference: 27 km
- Center of mass energy
  - 7, 8 TeV (2010-2012)
  - 13 TeV (2015 - 2018)
- Designed luminosity: 10E34 cm^{-2}s^{-1}

Geneva lake

LHCb

ALICE

Bunch

Proton

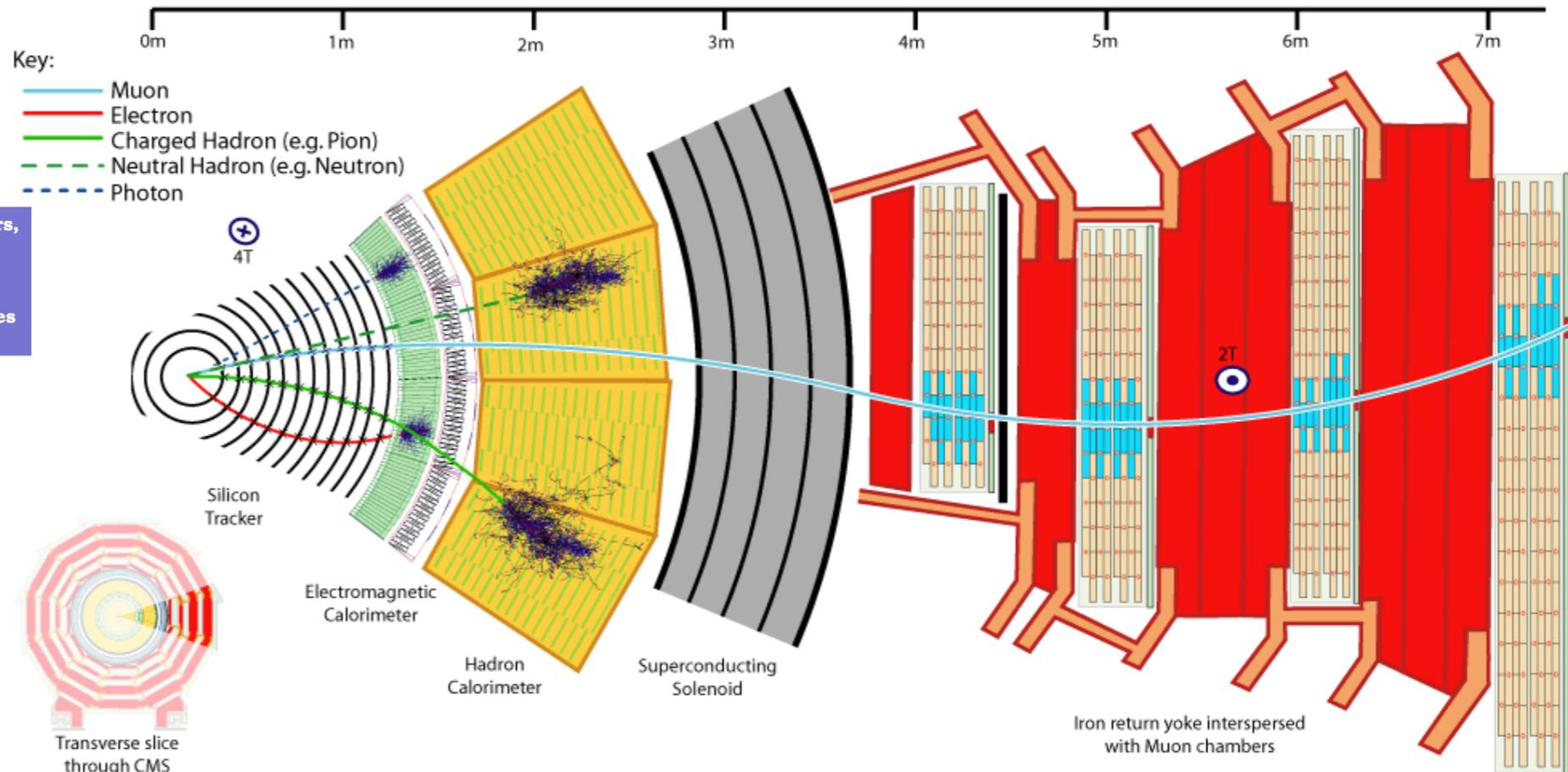Parton (quark, gluon)

Particle
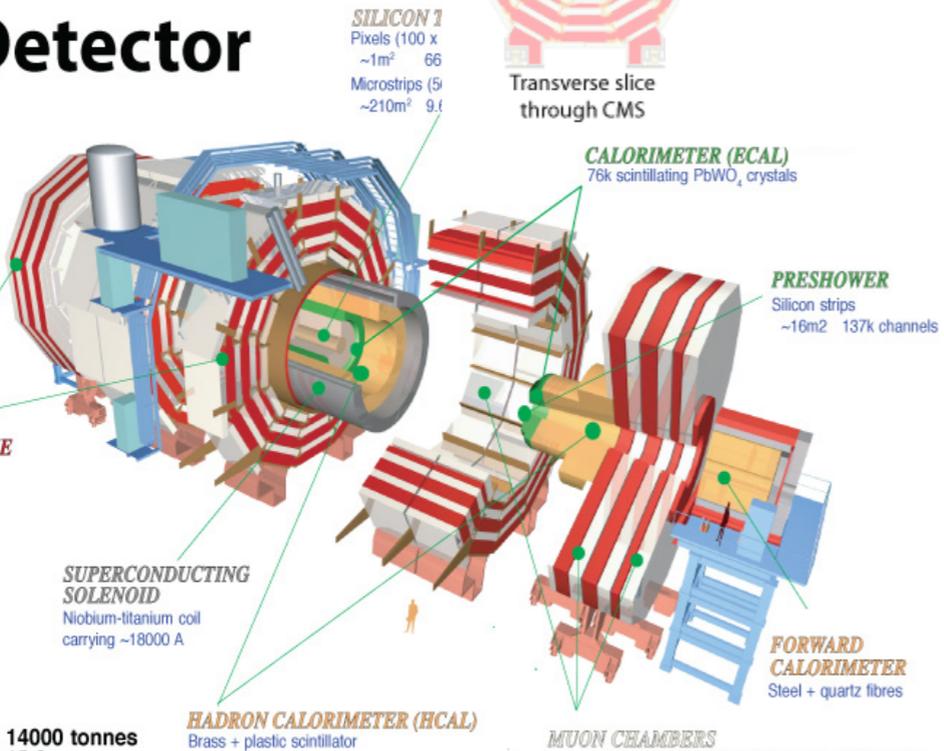
jet

jet

ATLAS

CMS Experiment

CMS Collaboration

Many: 5000 members, 1900 physicists, 1800 students, 950 eng./ techn.
Global: 200 institutes from 43 countries

Key:
— Muon
— Electron
— Charged Hadron (e.g. Pion)
-- Neutral Hadron (e.g. Neutron)
···· Photon

Silicon Tracker

Electromagnetic Calorimeter

Hadron Calorimeter

Superconducting Solenoid

Transverse slice through CMS

Iron return yoke interspersed with Muon chambers

D. Barney, CERN, February 2004

CMS Detector

Pixels
Tracker
ECAL
HCAL
Solenoid
Steel Yoke
Muons

SILICON T
Pixels (100 x
~1m² 66
Microstrips (5
~210m² 9.6

CALORIMETER (ECAL)
76k scintillating PbWO₄ crystals

PRESHOWER
Silicon strips
~16m2 137k channels

STEEL RETURN YOKE
~13000 tonnes

SUPERCONDUCTING SOLENOID
Niobium-titanium coil carrying ~18000 A

FORWARD CALORIMETER
Steel + quartz fibres

HADRON CALORIMETER (HCAL)
Brass + plastic scintillator

MUON CHAMBERS
Barrel: 250 Drift Tube & 500 Resistive Plate Chambers
Endcaps: 450 Cathode Strip & 400 Resistive Plate Chambers

Total weight   : 14000 tonnes
Overall diameter : 15.0 m
Overall length   : 28.7 m
Magnetic field   : 3.8 T

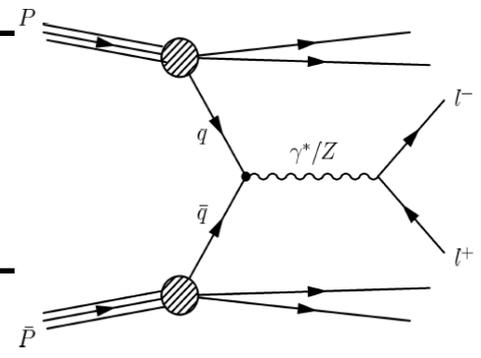| type	tracking	ECAL	HCAL	MUON
γ				
e				
μ				
Jet				
Et miss				

5

Unification of Forces: Z'

- Flagship EXO search in dilepton channel: Z' search

- Probe indirect search for DM: mediator to SM particles

- Also can test the lepton flavor universality between dielectron and dimuon channels
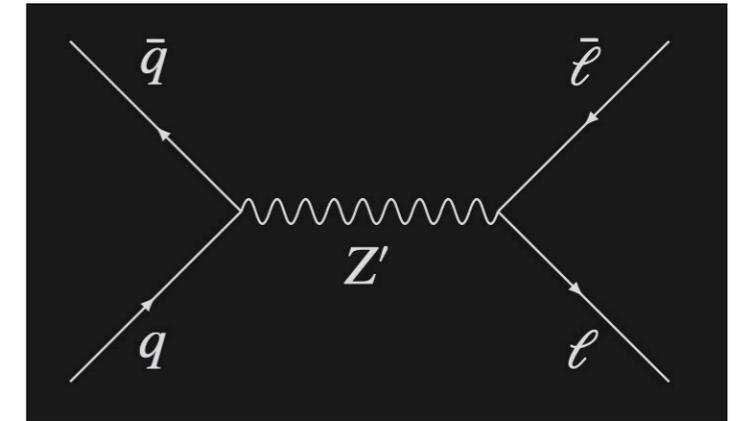
JHEP 07 (2021) 208
JHEP 06 (2018) 120
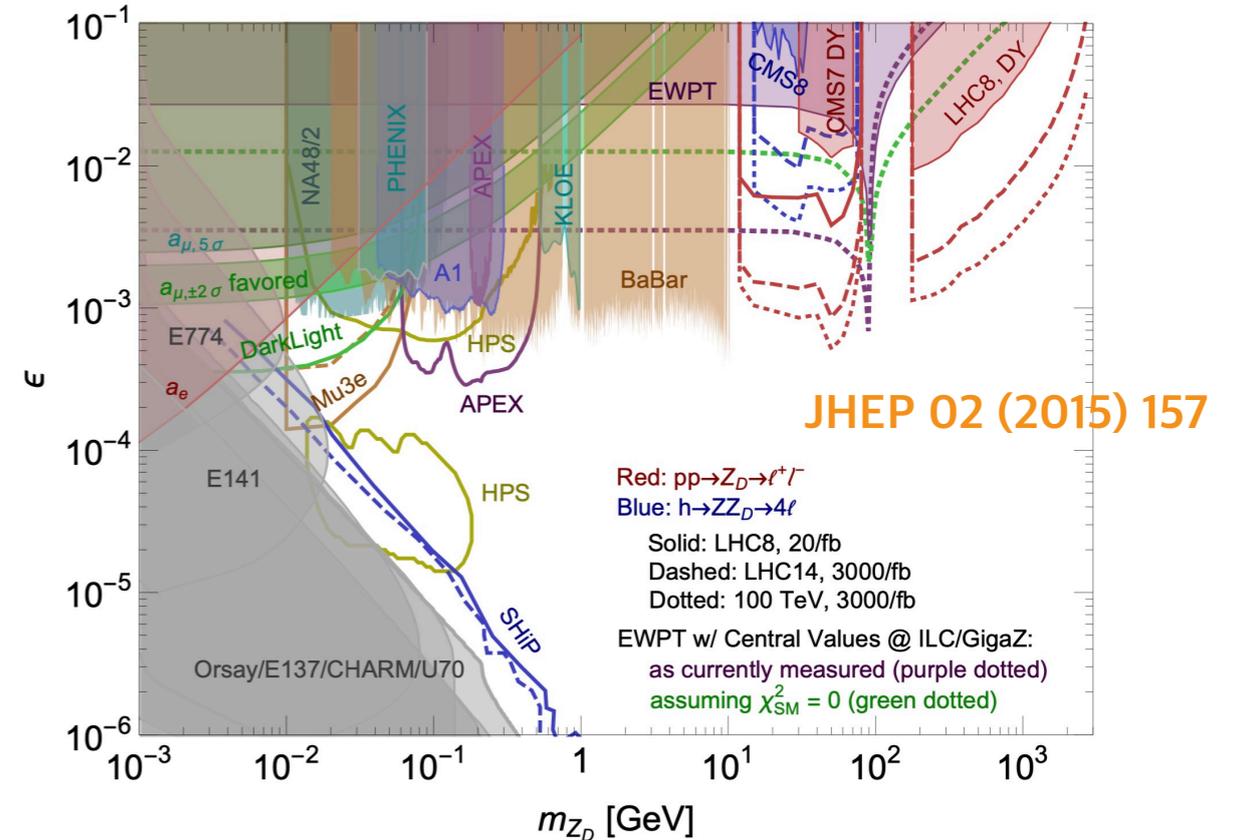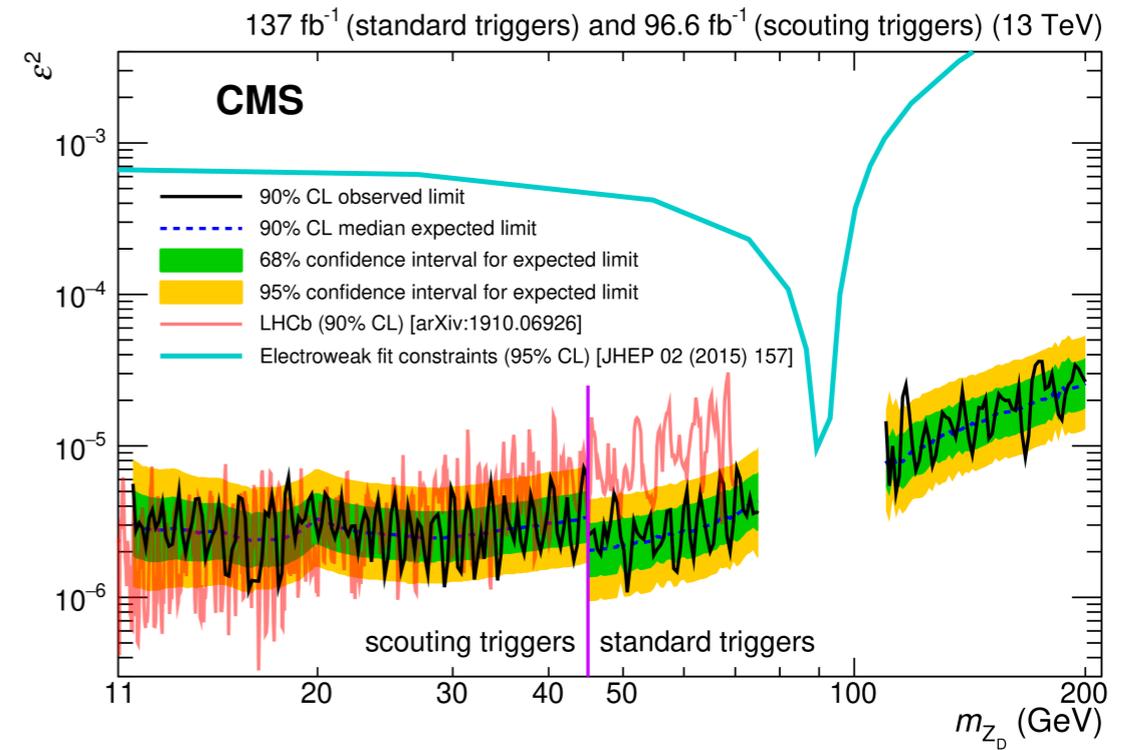**PLB 768 (2017) 57**
JHEP 04 (2015) 025
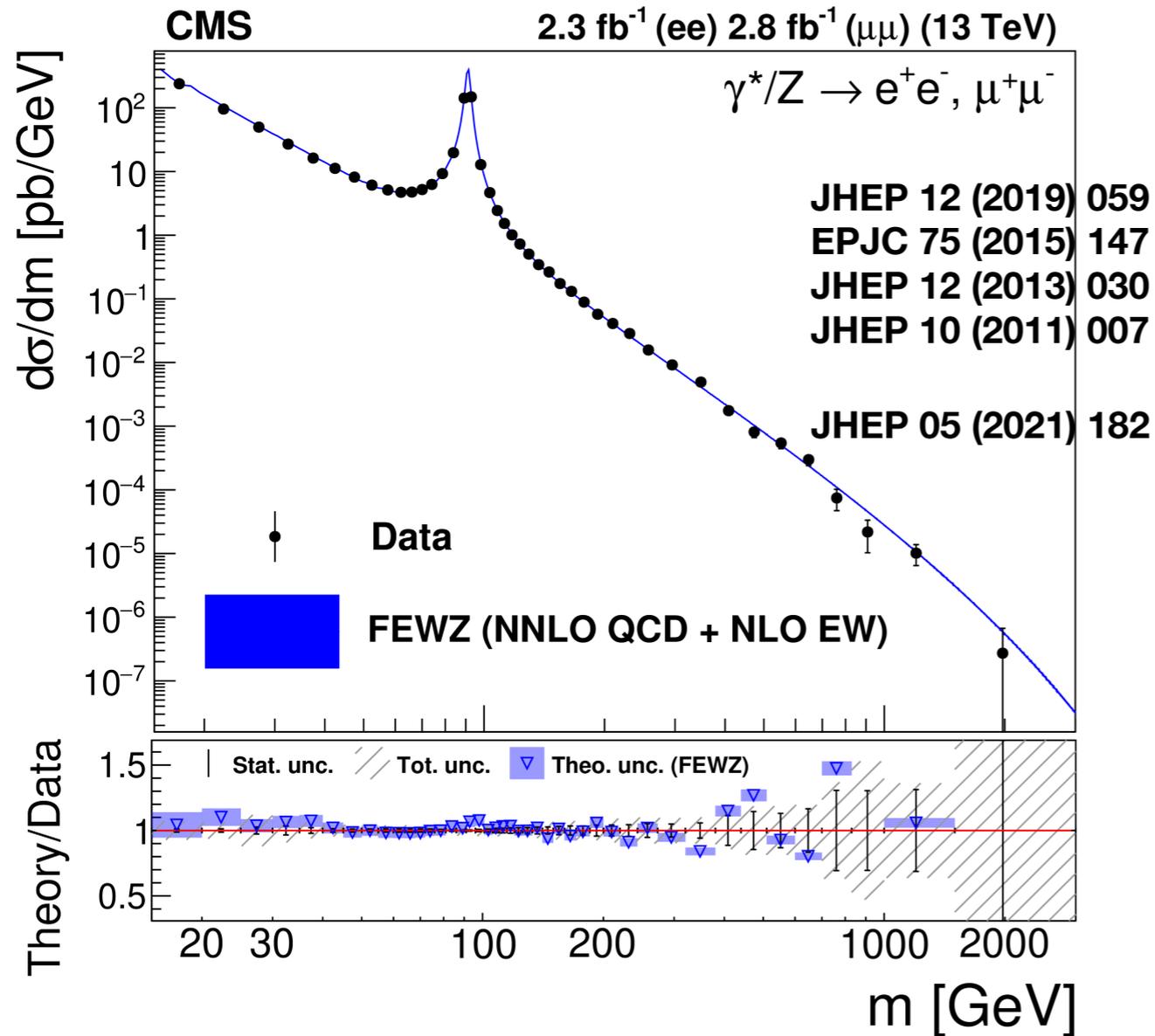
PLB 773 (2017) 563
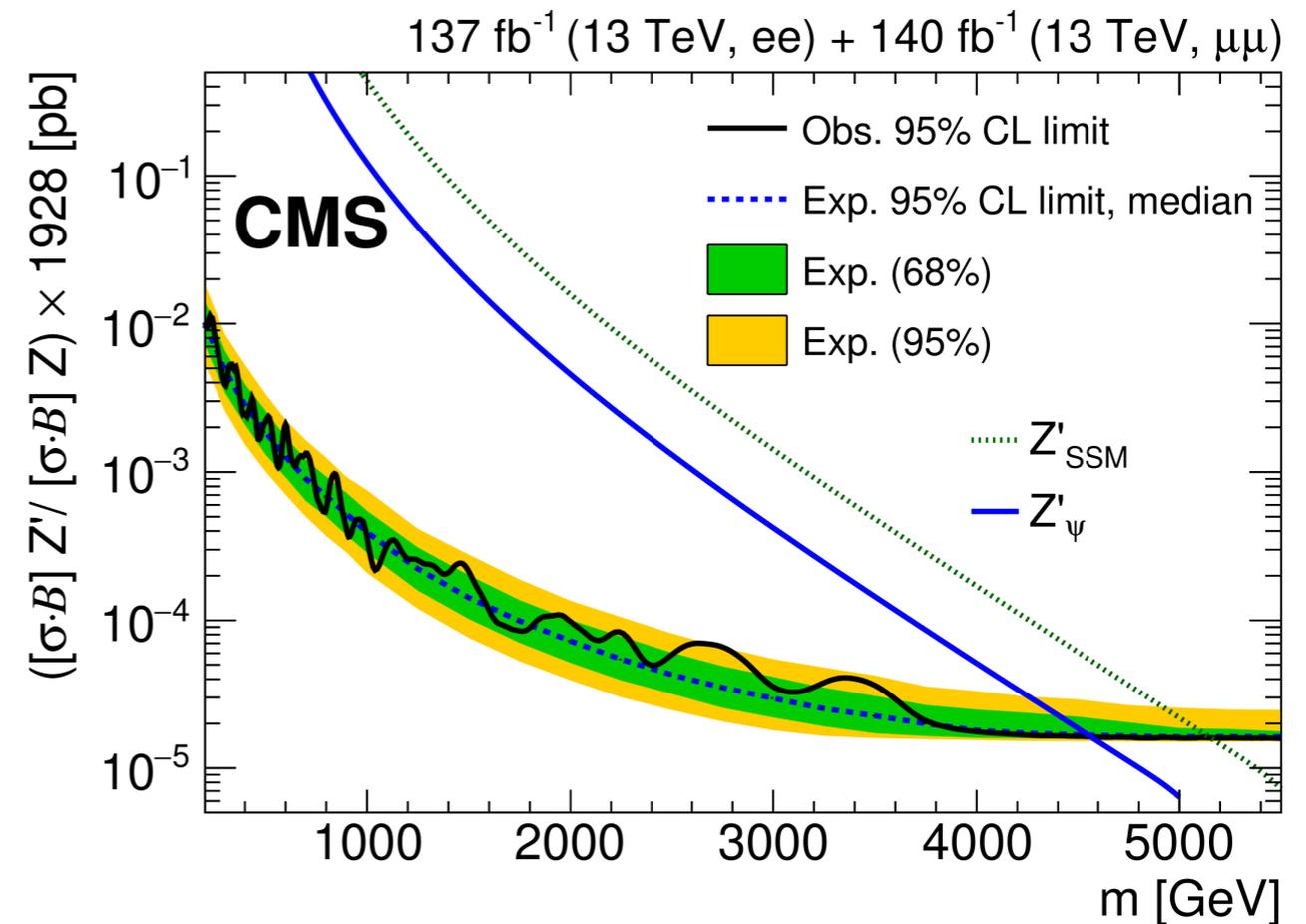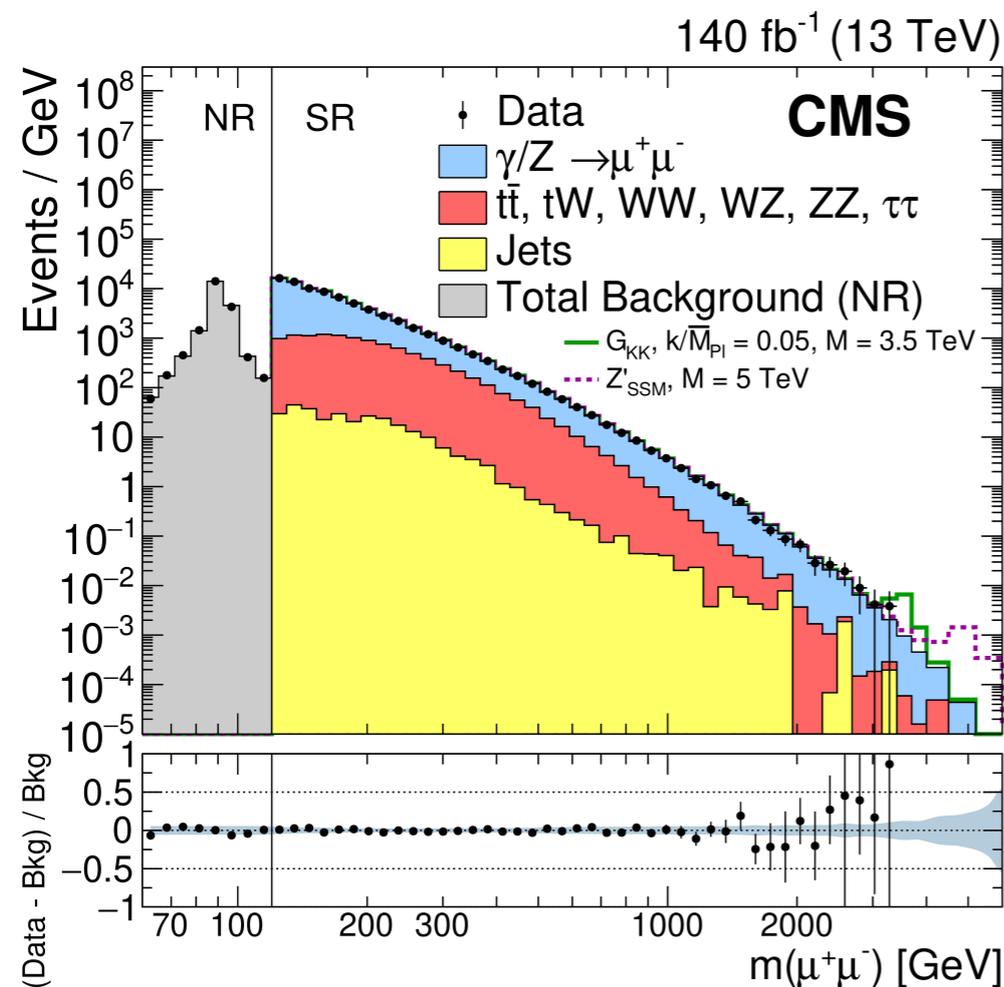
Drell-Yan Process

- Drell-Yan differential cross section measurements in dilepton channel

- Full Run2 data analysis is on-going: low-mass (scouting), high-mass (standard)

  - With various multi-dimensional measurements

137 fb^{-1} (standard triggers) and 96.6 fb^{-1} (scouting triggers) (13 TeV)

CMS

- 90% CL observed limit
- 90% CL median expected limit
- 68% confidence interval for expected limit
- 95% confidence interval for expected limit
- LHCb (90% CL) [arXiv:1910.06926]
- Electroweak fit constraints (95% CL) [JHEP 02 (2015) 157]

scouting triggers | standard triggers

ε^2

m_{Z_D} (GeV)

CMS   2.3 fb^{-1} (ee) 2.8 fb^{-1} (μμ) (13 TeV)

$\gamma^*/Z \rightarrow e^+e^-,\ \mu^+\mu^-$

JHEP 12 (2019) 059
EPJC 75 (2015) 147
JHEP 12 (2013) 030
JHEP 10 (2011) 007

JHEP 05 (2021) 182

Data

FEWZ (NNLO QCD + NLO EW)

dσ/dm [pb/GeV]

Theory/Data

Stat. unc.   Tot. unc.   Theo. unc. (FEWZ)

m [GeV]

JHEP 02 (2015) 157

Red: $pp \rightarrow Z_D \rightarrow \ell^+\ell^-$
Blue: $h \rightarrow ZZ_D \rightarrow 4\ell$

Solid: LHC8, 20/fb
Dashed: LHC14, 3000/fb
Dotted: 100 TeV, 3000/fb

EWPT w/ Central Values @ ILC/GigaZ:
as currently measured (purple dotted)
assuming $\chi^2_{SM} = 0$ (green dotted)

ε

m_{Z_D} [GeV]

Why Quantum Machine Learning?

- Technical problems due to lack of computing resources

  - CPU, GPU, storage etc.

    About one year, I have prepared for this kind of study with undergraduate intern students : Qiskit, CMS open data

- Few examples we face (with my concern)

My long term interesting

  - Large MC samples needed => detector simulation (GEANT4 based) takes a lot of time

    - This problem will be much more serious in Future Collider projects

My expertise

  - Track (vertex) reconstruction in high pileup => can do better offline but challenging online (triggering)

My near term interesting

  - Optimization for wider categories => can speed up new observation

- QC technology is being grown fast! => just curiosity for new technology

  No evidence yet for QML > ML, but worth while trying

QC (QML) in HEP

QC in HEP

- (re)cast combinatorial problems in a quantum formulation
  - ➡ Pattern recognition, tracking : arxiv:1902.08324, arxiv:1902.00498, arxiv:1908.04475, arxiv:2104.11583, …
  - ➡ Vertexing : arxiv:1903.08879, …
  - ➡ Jet clustering : arxiv:1908.08949, arxiv:2012.14514, arxiv:2101.05618, …
- Using the probabilistic character of quantum device
  - ➡ Simulation of quantum effects : arxiv:1811.10085, arxiv:1904.03196, arxiv:1904.10559, arxiv:2010.00046, arxiv:2010.06458, arxiv:2011.13934, arxiv:2105.04765, arxiv:2105.08703, …
  - ➡ Unfolding : arxiv:1908.08519, …

The multi-jet QUBO objective function is

$$O_{\text{QUBO}}(\{x_{ij}\}) = \sum_{r=1}^{M} \sum_{i,j=1}^{N} \left(\frac{\vec{p}_i \cdot \vec{p}_j - E_i E_j \cos R}{1 - \cos R} \right) x_{ir} x_{jr} + \Lambda^2 \sum_{i=1}^{N} \left(1 - \sum_{r=0}^{M} x_{ir} \right)^2.$$

1908.08949

qPDF Workflow 2011.13934
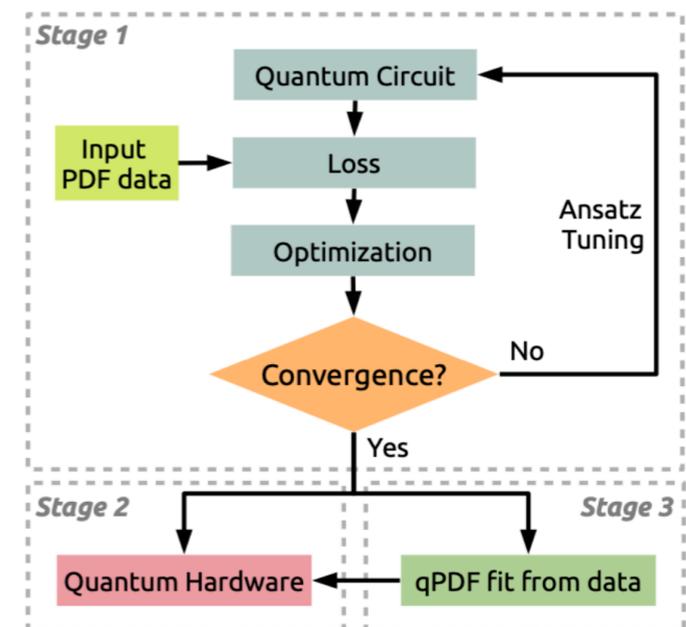
QML Examples in IBM Qiskit


Hybrid quantum-classical Neural Networks with PyTorch and Qiskit

Machine learning (ML) has established itself as a successful interdisciplinary field which seeks to mathematically extract generalizable information from data. Throwing in quantum computing gives rise to interesting areas of research which seek to leverage the principles of quantum mechanics to augment machine learning or vice-versa. Whether you're aiming to enhance classical ML algorithms by outsourcing difficult calculations to a quantum computer or optimise quantum algorithms using classical ML architectures - both fall under the diverse umbrella of quantum machine learning (QML).

In this chapter, we explore how a classical neural network can be partially quantized to create a hybrid quantum-classical neural network. We will code up a simple example that integrates **Qiskit** with a state-of-the-art open-source software package - PyTorch. The purpose of this example is to demonstrate the ease of integrating Qiskit with existing ML tools and to encourage ML practitioners to explore what is possible with quantum computing.
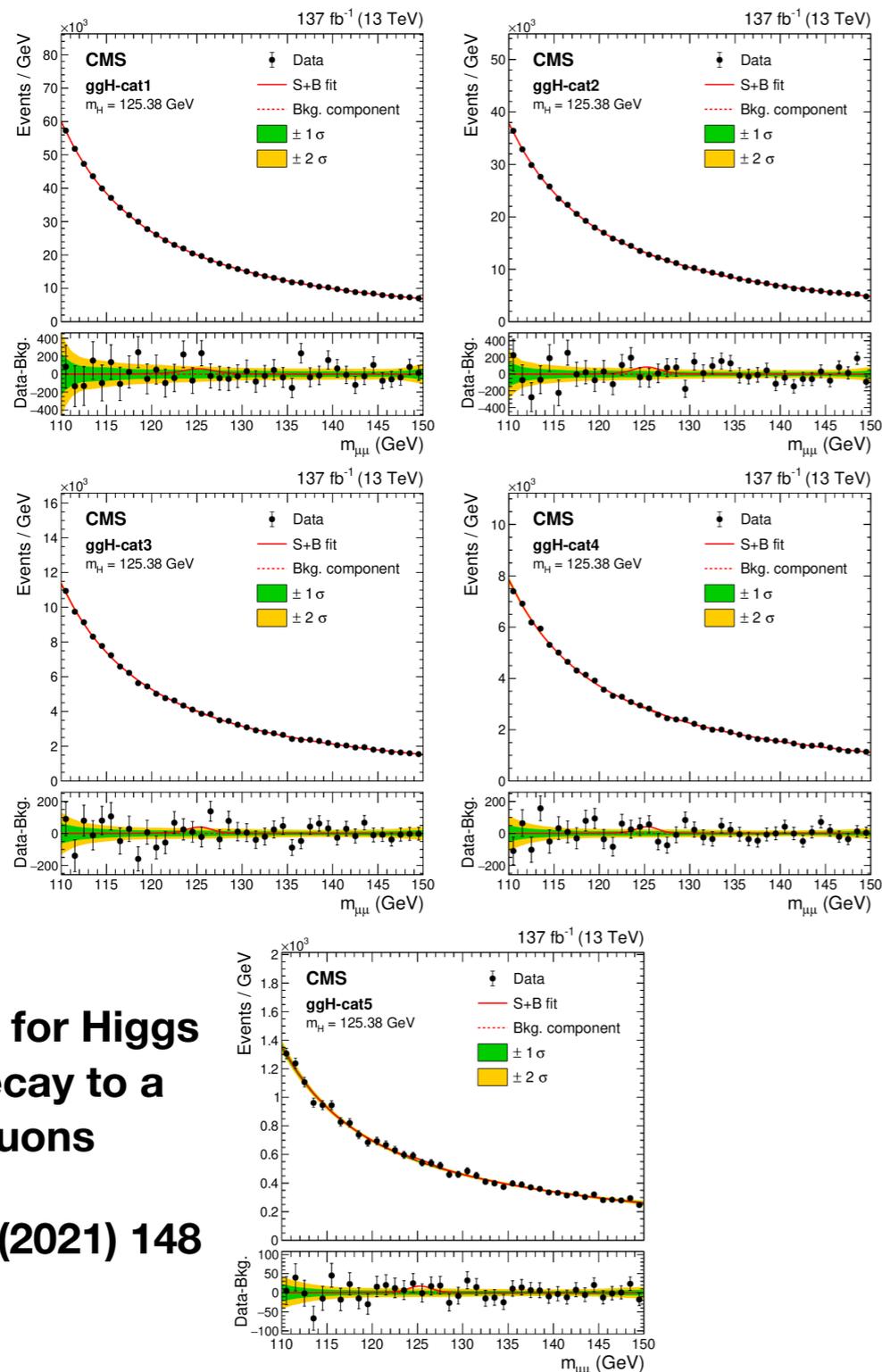
1. How does it work?
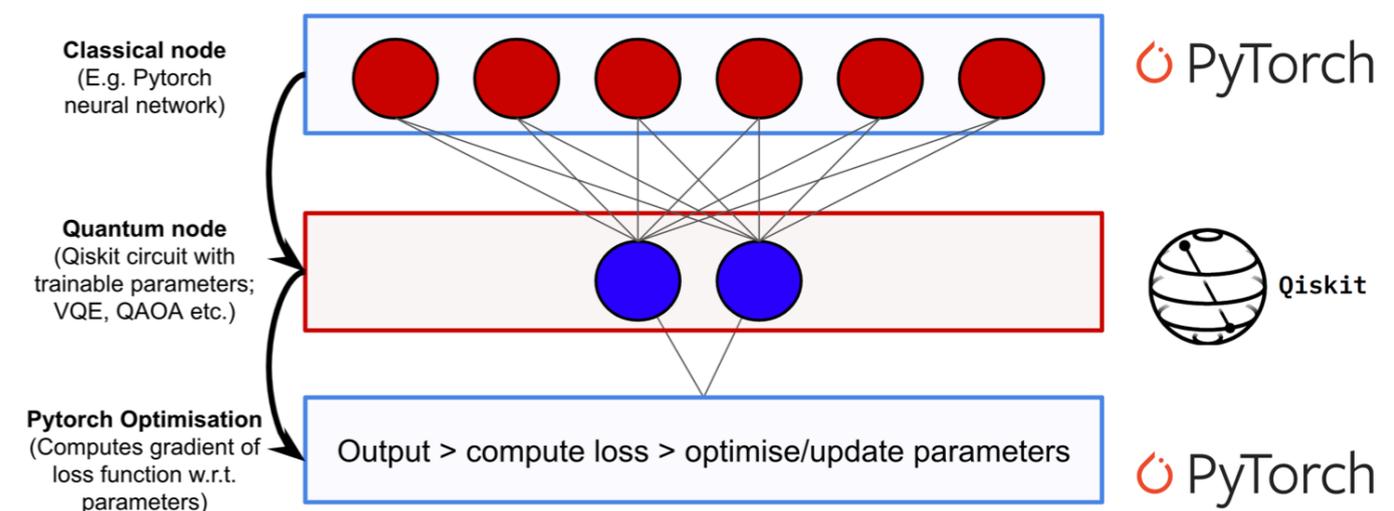
Topic to Start: H->mumu

- Benchmark study is available arxiv: 2012.11560

Sau Lan Wu[1*], Jay Chan[1], Wen Guan[1], Shaojun Sun[1], Alex Wang[1], Chen Zhou[1]
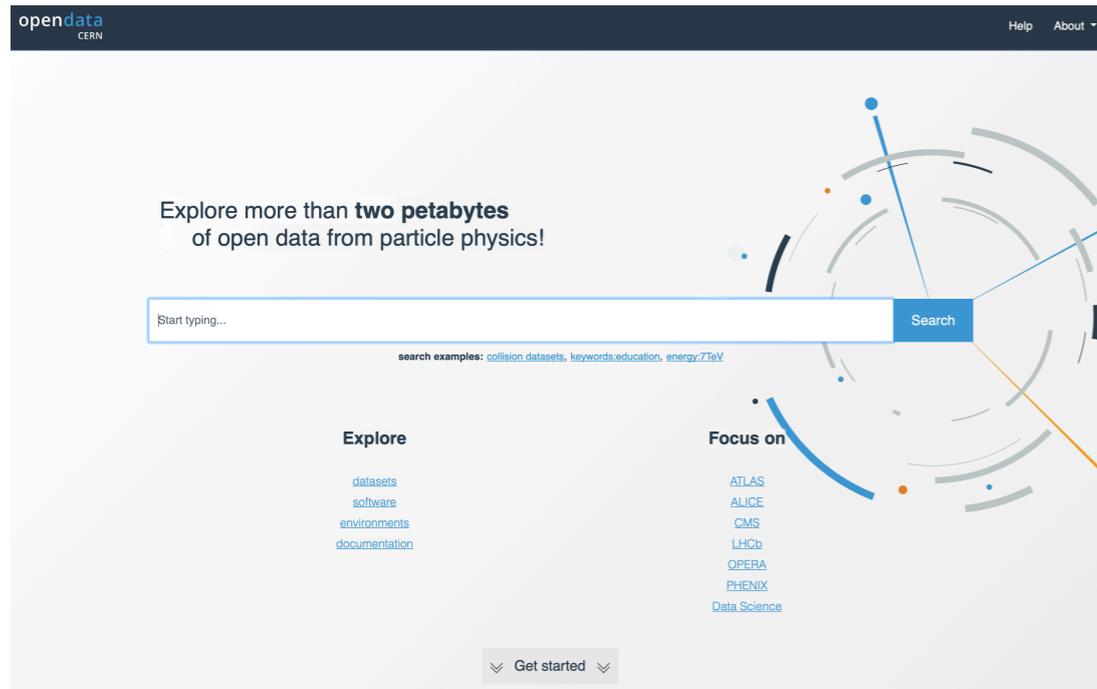Miron Livny[2]
Federico Carminati[3], Alberto Di Meglio[3]

Evidence for Higgs boson decay to a pair of muons

JHEP 01 (2021) 148

11
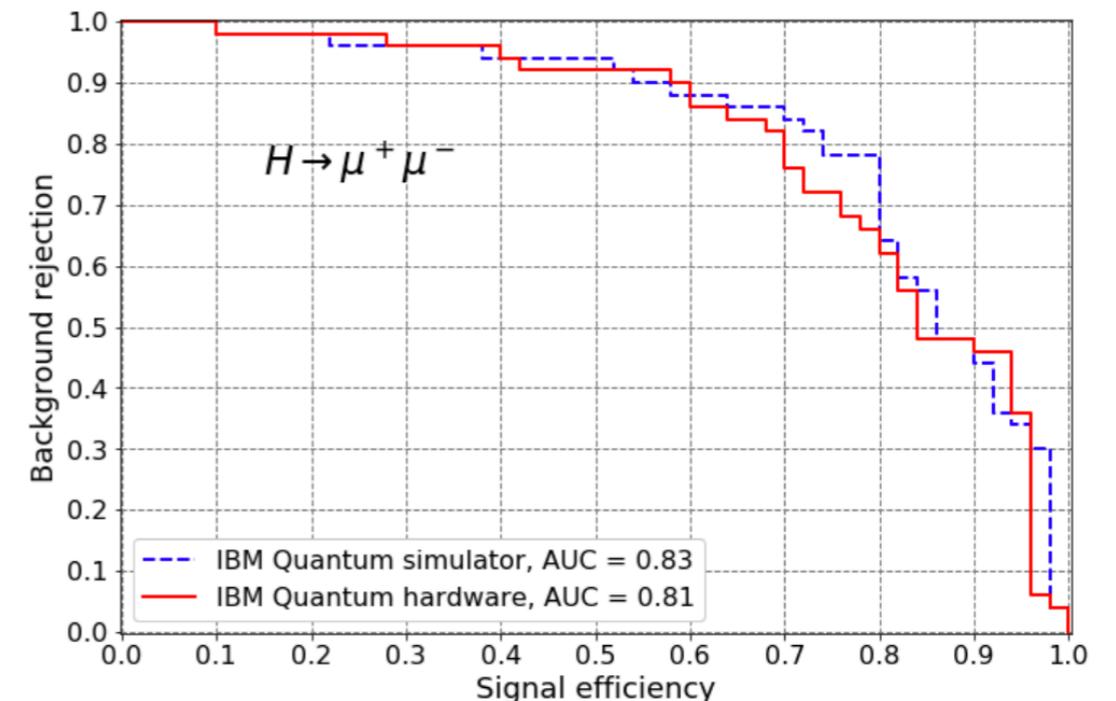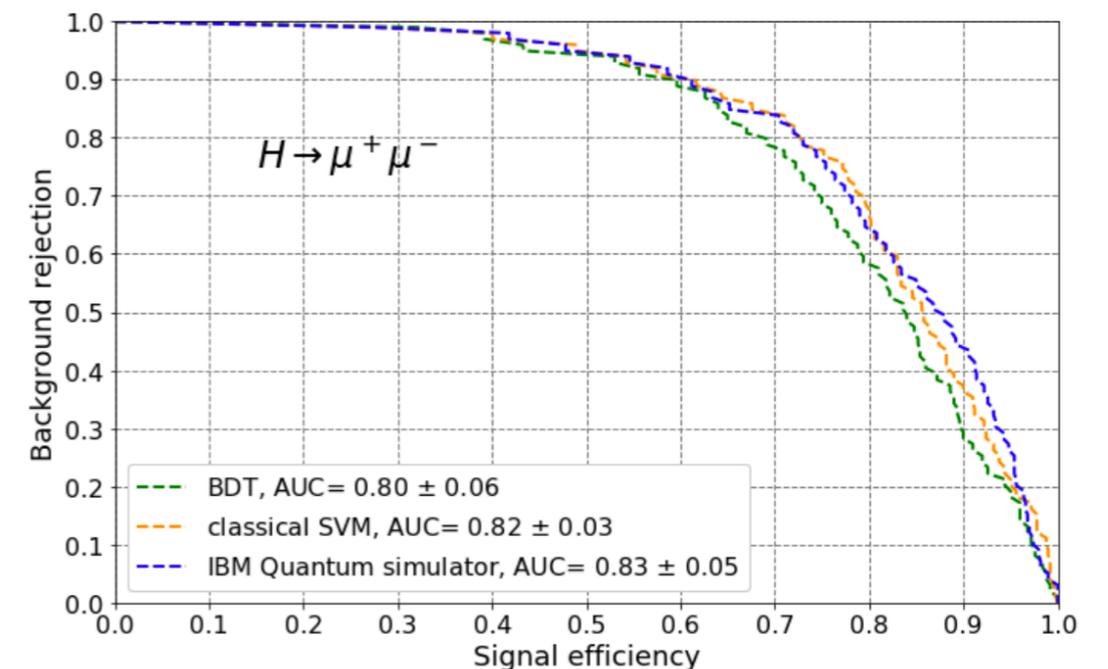
CMS Open Data
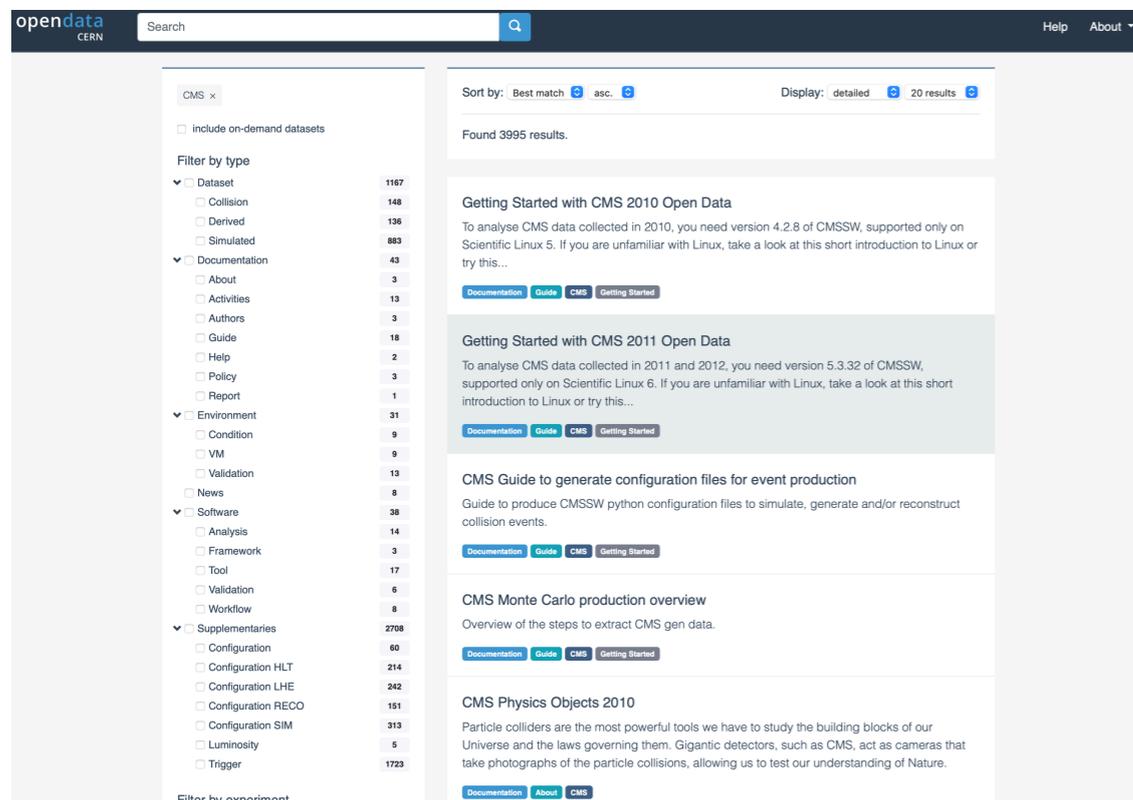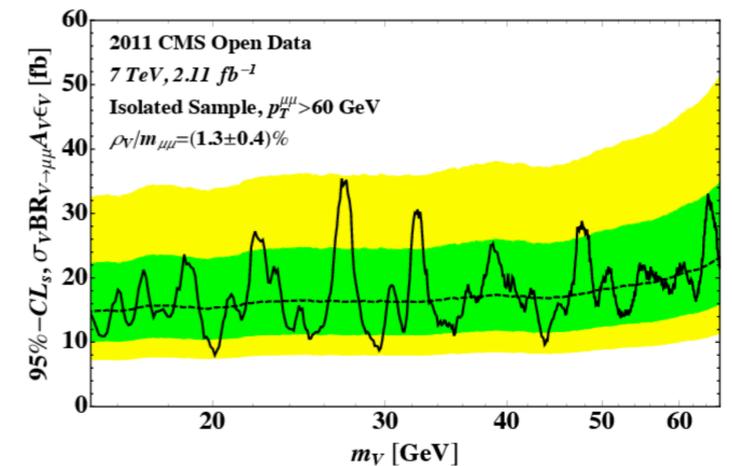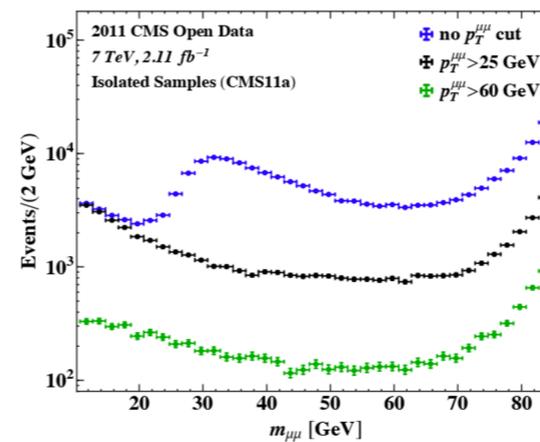

- Open data has been used to perform new searches and measurements

- Searches for non-standard sources of parity violation in jets [arXiv:1904.11195]

- Searches for high-p_T dimuon resonances [arXiv:1902.04222]

Linux(Ubuntu) or Windows 10

Docker

Recipe is available

CMS Software

CMS open data setting on Window 10
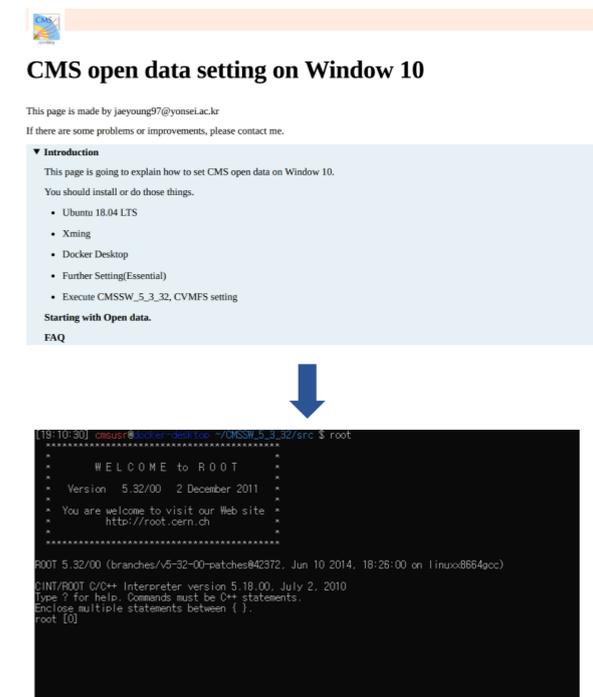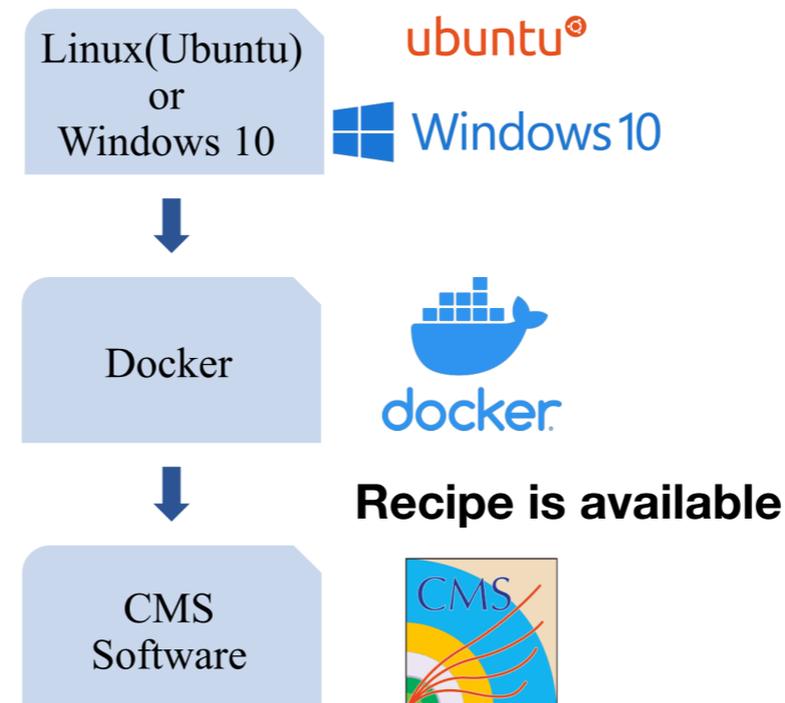
Near Term Goal on QML

- Be familiar to QML technique

  - Accumulate the knowledge/experience on the QML

- Follow up the benchmark study using CMS open data for H->mumu search

  - Extend to other advanced topics afterward

  - Searches in dilepton + X channels or tau channel

- QML vs ML

  - Speed up data sampling, optimization, models

  - Use fewer data samples (benefits for rare decay and new physics search)

  - Quantum generative model

- Essential to follow up the QML for our future

  - Detector design for Future Collider projects should consider the QML

Summary & Plan

- Plan to learn QML and study the benchmark topic

  - Research setup is (almost) ready: Qiskit, QML algorithm, CMS open data

  - Need a help from QML or ML experts who are interested in this project

- Need to understand which kind of benefits will be available from QML for future collider projects

14