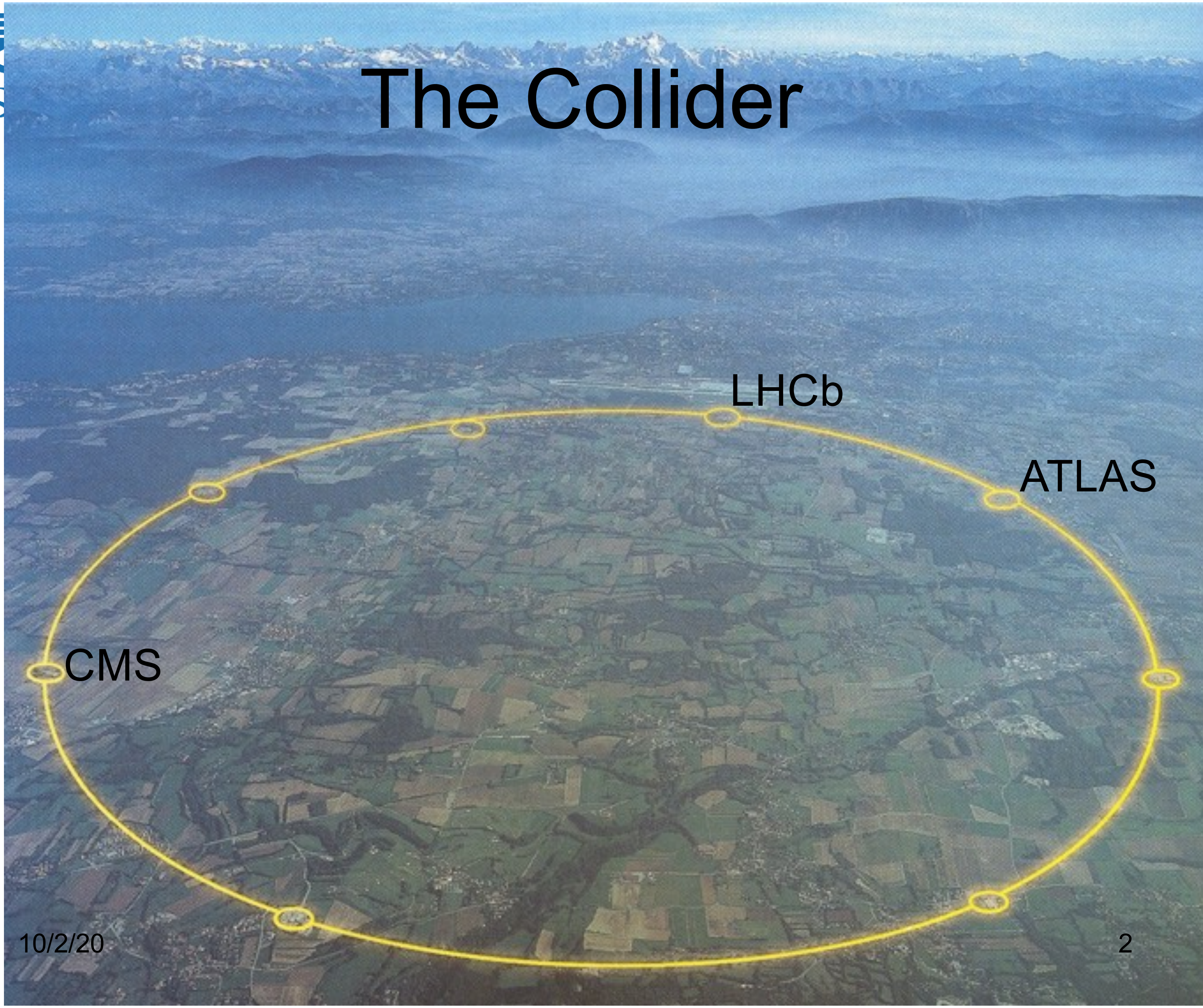
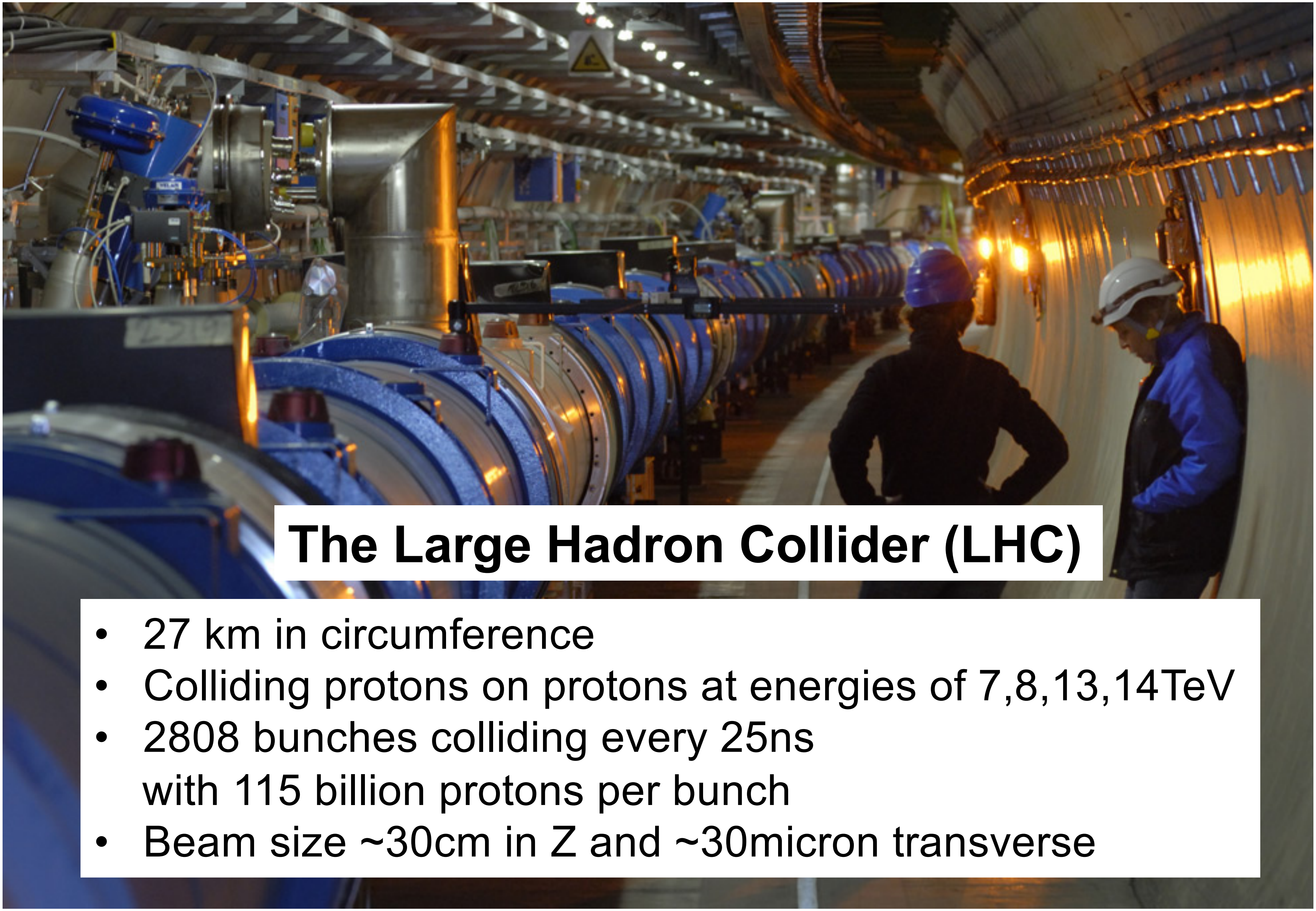
Overview of Experimental Particle Physics at Colliders

Frank Würthwein
SDSC/UCSD

The Collider





The Large Hadron Collider (LHC)

- 27 km in circumference
- Colliding protons on protons at energies of 7,8,13,14TeV
- 2808 bunches colliding every 25ns
  with 115 billion protons per bunch
- Beam size ~30cm in Z and ~30micron transverse

"Big bang" in the laboratory

- We gain insight by *colliding protons at the highest energies* possible to measure:
  - Production rates
  - Masses & lifetimes
  - Decay rates
- From this we *derive the "spectroscopy" as well as the "dynamics" of elementary particles*.
- Progress is made by going to higher energies and more proton proton collisions per beam crossing.
  - More collisions => increased sensitivity to rare events
  - More energy => probing higher masses, smaller distances & earlier times

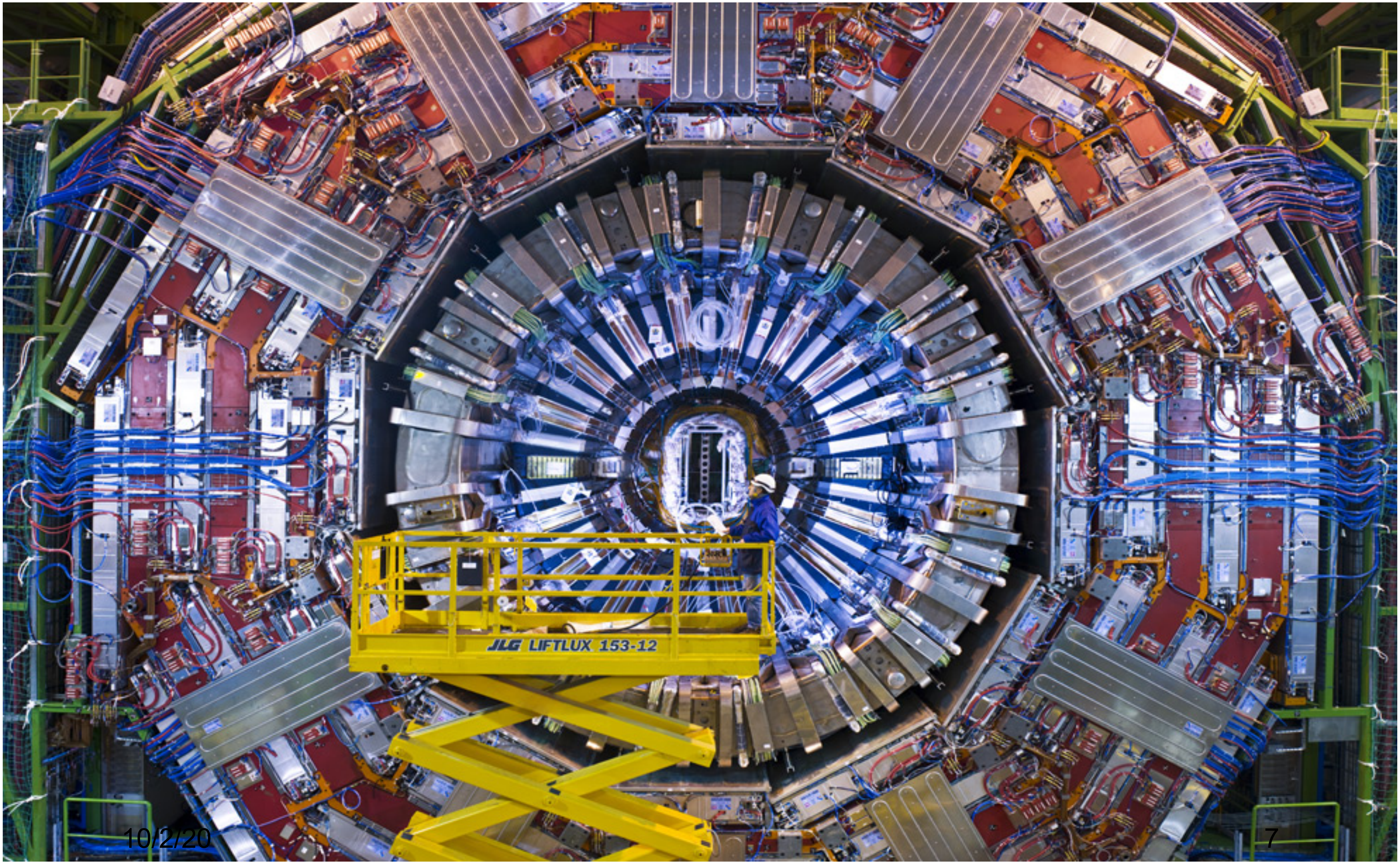
Spectroscopy and Dynamics

- Spectroscopy:
  - What are the particles that exist ?
  - What are their properties ?
- Dynamics:
  - What are the forces ?
  - How do the particles couple to the forces ?
  - How do these depend on energy and angular momentum ?

Data Science in a nutsshell

- We select proton-proton collisions based on some characteristics of the measurement we want to make.
  - 10,000,000,000,000,000 collisions per year out of which we sometimes attempt to select only a few to make a measurement.
- We count them
  - Accounting for background effects
  - Accounting for potential selection biases
- We derive meaning from counting through statistics.

The CMS Experiment (R-Φ view)



The CMS Experiment

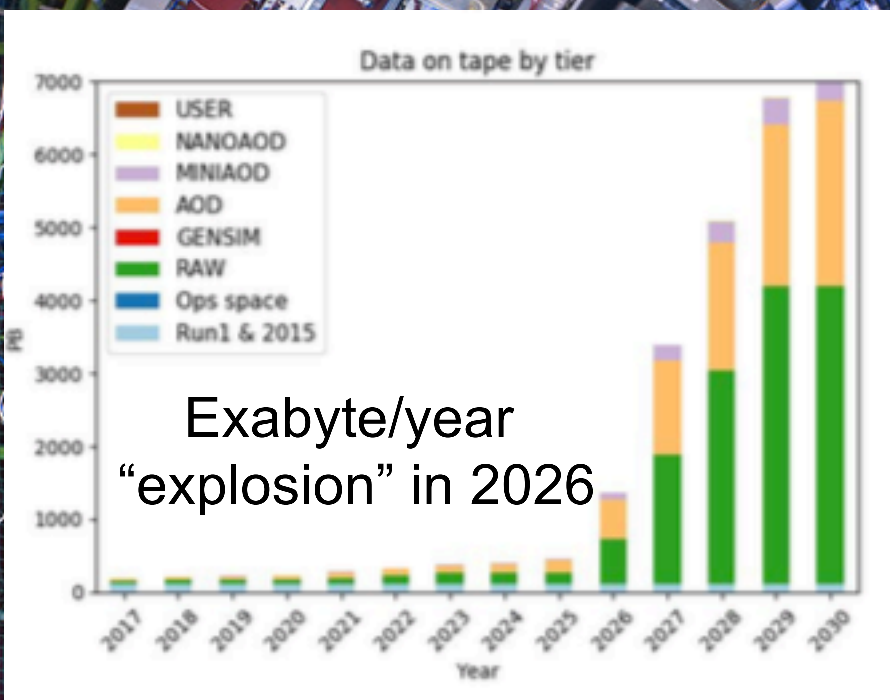
- **80 Million electronic channels**
  x 4 bytes
  x 40MHz

  ~ **10 Petabytes/sec** of information
  x 1/1000 zero-suppression
  x 1/100,000 online event filtering

  ~ 1000 Megabytes/sec raw data to tape
  ~10 Petabytes of raw data per year written to tape, not counting simulations.

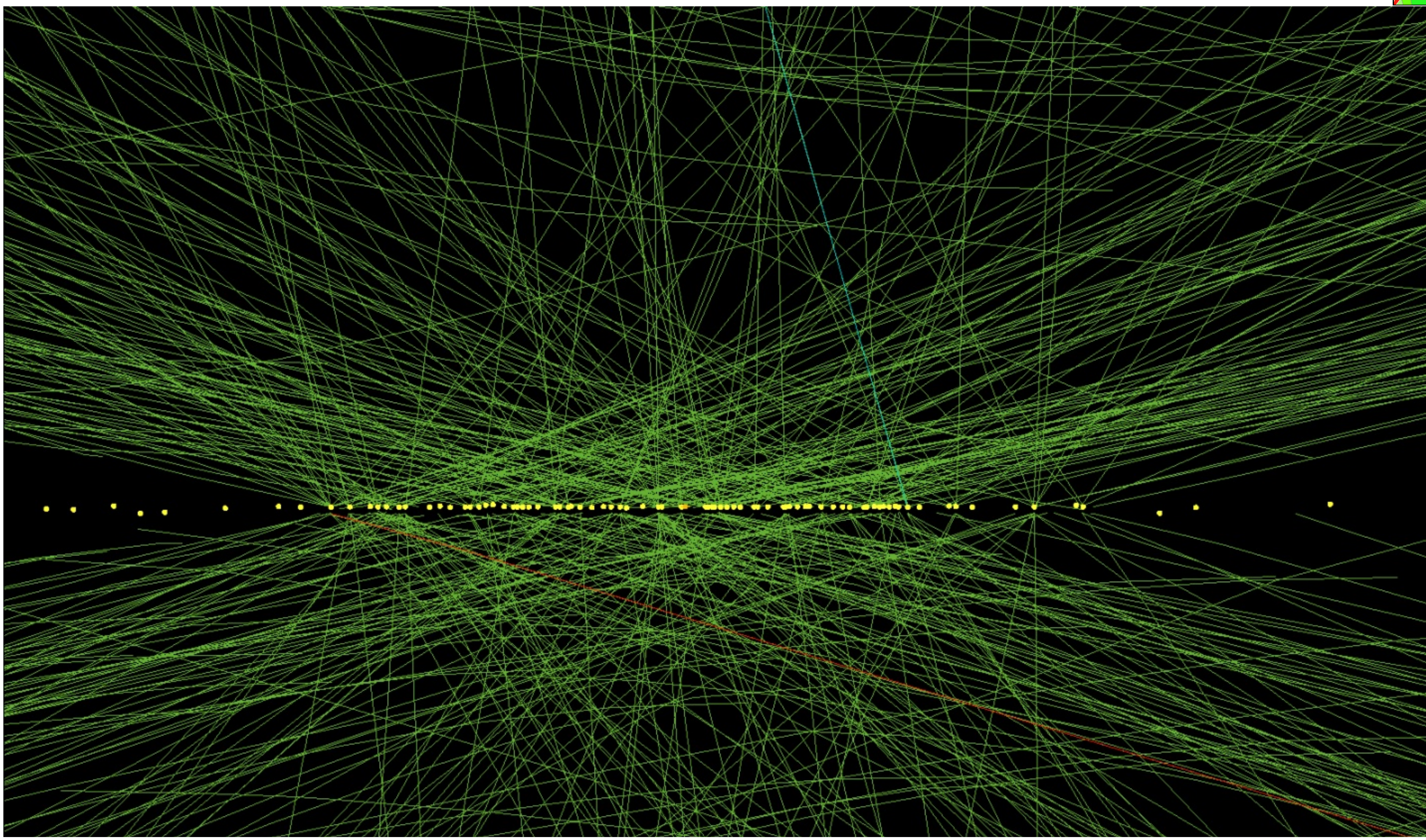
- 2000 Scientists (1200 Ph.D. in physics)
  - ~ 180 Institutions
  - ~ 40 countries

- 12,500 tons, 21m long, 16m diameter



Zoomed in R-Z view of busy event

beam axis



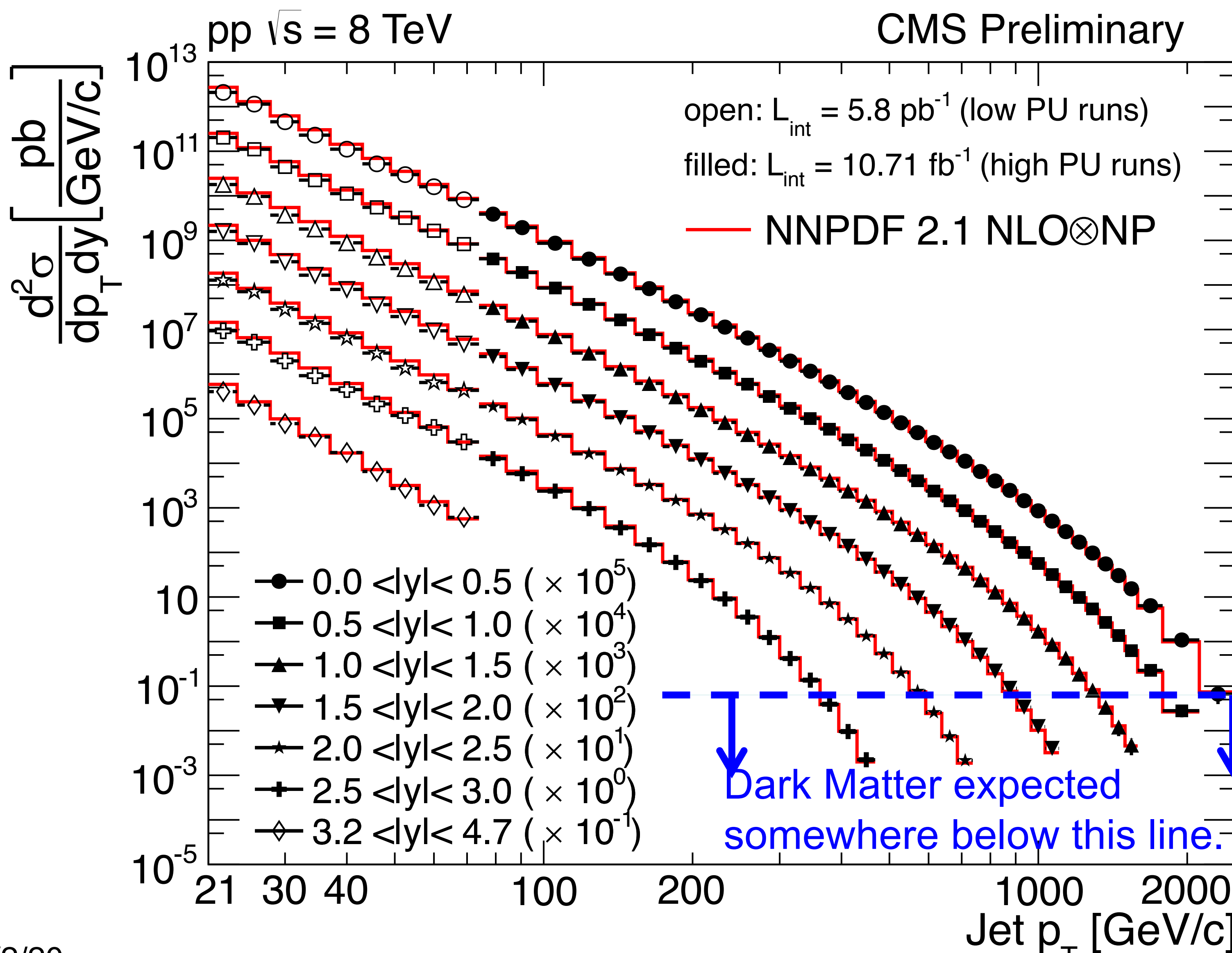
beam

Green lines indicate reconstructed charged particle trajectories.

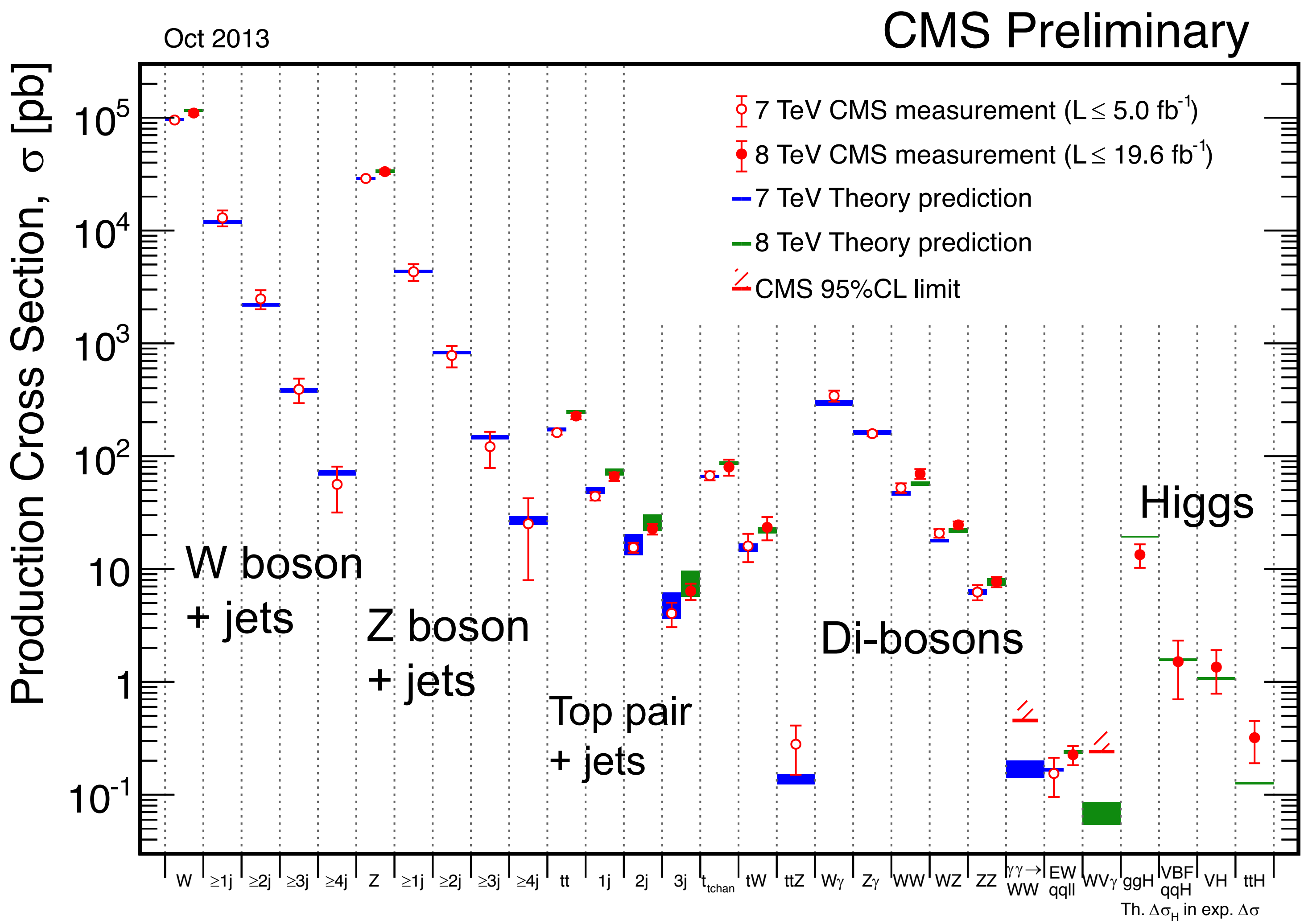
Yellow dots indicate reconstructed collisions, all during the same beam crossing.

Explore Nature over ~13 Orders of magnitude

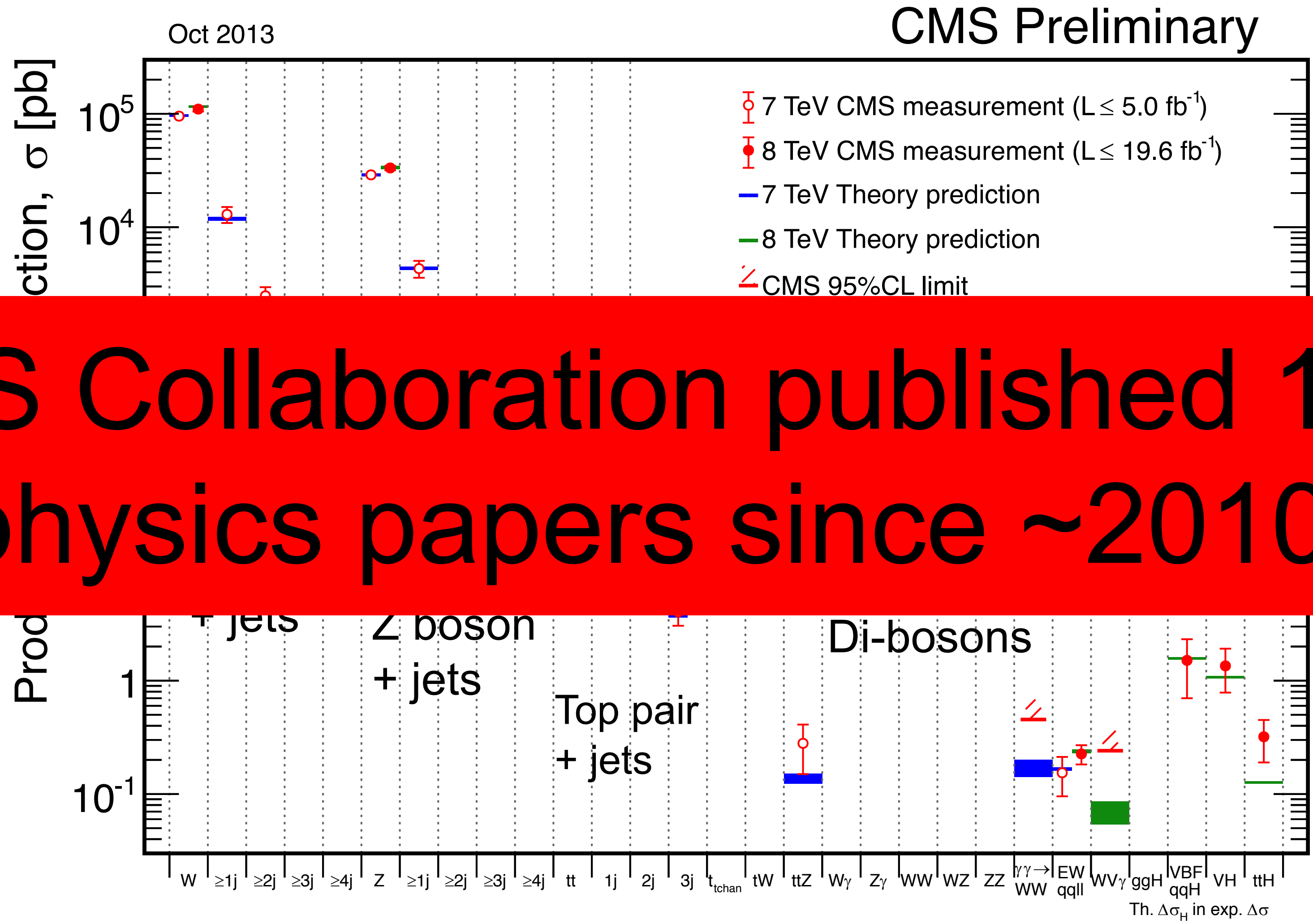
Excellent agreement between Theory & Experiment



Wide variety of different processes are being probed.



Wide variety of different processes are being probed.



CMS Collaboration published 1000 physics papers since ~2010.

Basic Data Flow

10PB/sec Raw Info → Trigger & Zero suppression → ~Gbyte/sec → Prompt Reco

~Gbyte/sec → Tape

Prompt Reco → Tape

2-3 times as much simulation as RAW Data → Tape

~200,000 core global processing capacity across ~100 clusters

150 Petabytes Global Data Federation

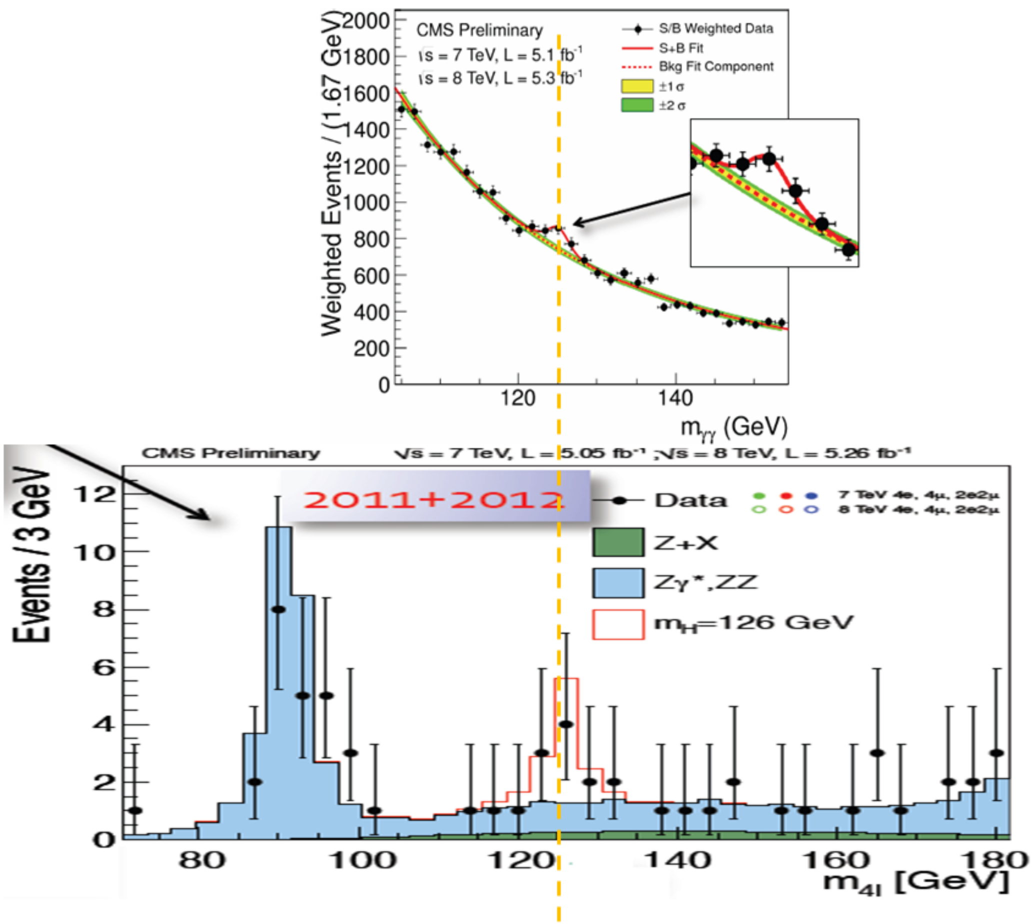
Any data anywhere on disk can be accessed from anywhere with an internet connection.
Something like a Content Delivery Network for Science.

Think of it as YouTube or Netflix for Science

How does this relate to your capstone project ?

You get to reproduce a reconstruction tool essential to measuring the dynamics in Higgs boson production

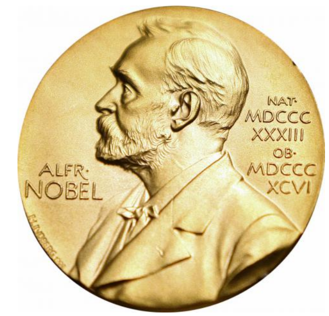
Why the Higgs ?



The Nobel Prize in Physics 2013

Photo: A. Mahmoud
François Englert
Prize share: 1/2

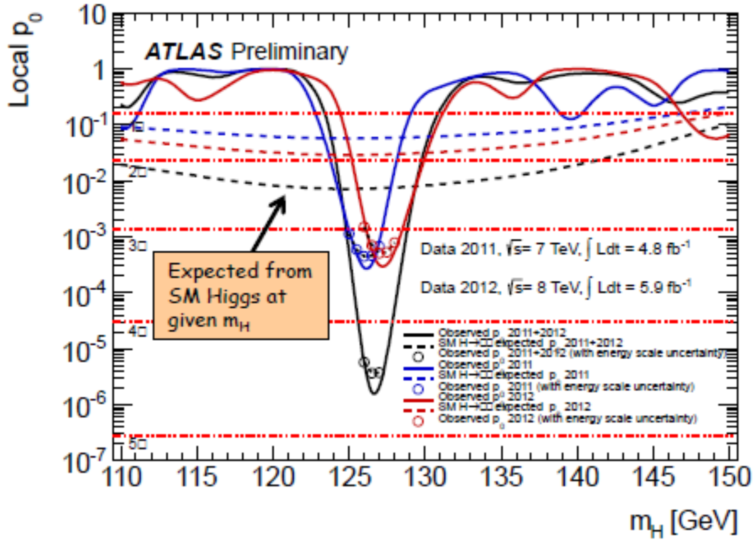
Photo: A. Mahmoud
Peter W. Higgs
Prize share: 1/2

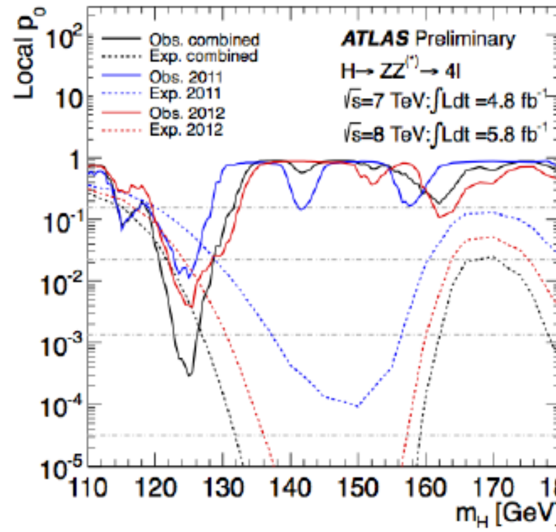


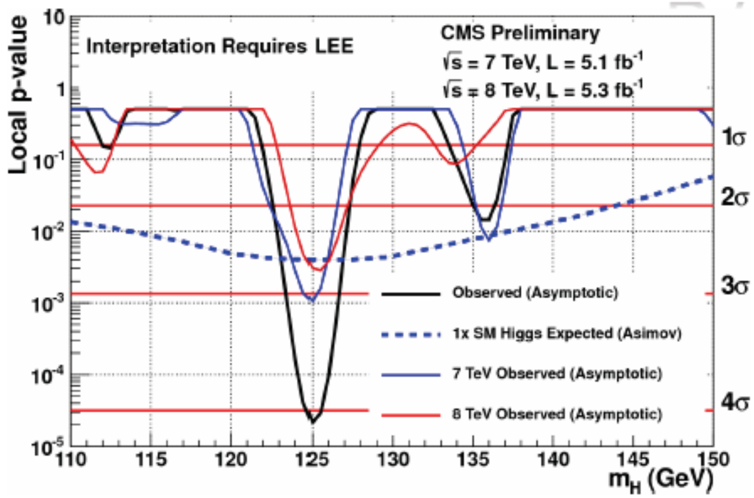
The Nobel Prize in Physics 2013 was awarded jointly to François Englert and Peter W. Higgs *"for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider"*

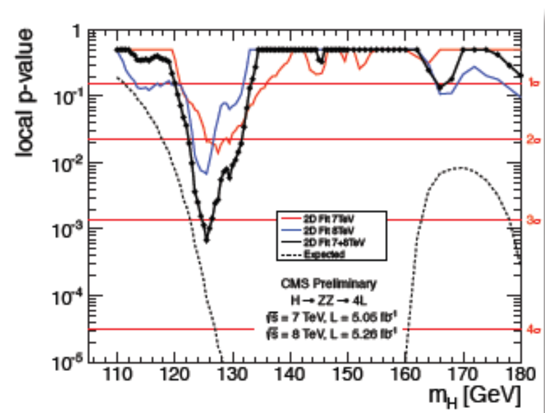
Why not reproduce the measurement that lead to Nobel Prize?

Complexity is staggering









Teams with 100’s of people involved in data analysis

So we picked a specific tool

Tool required **Deep Learning** to succeed.

Tool was used in published physics paper.

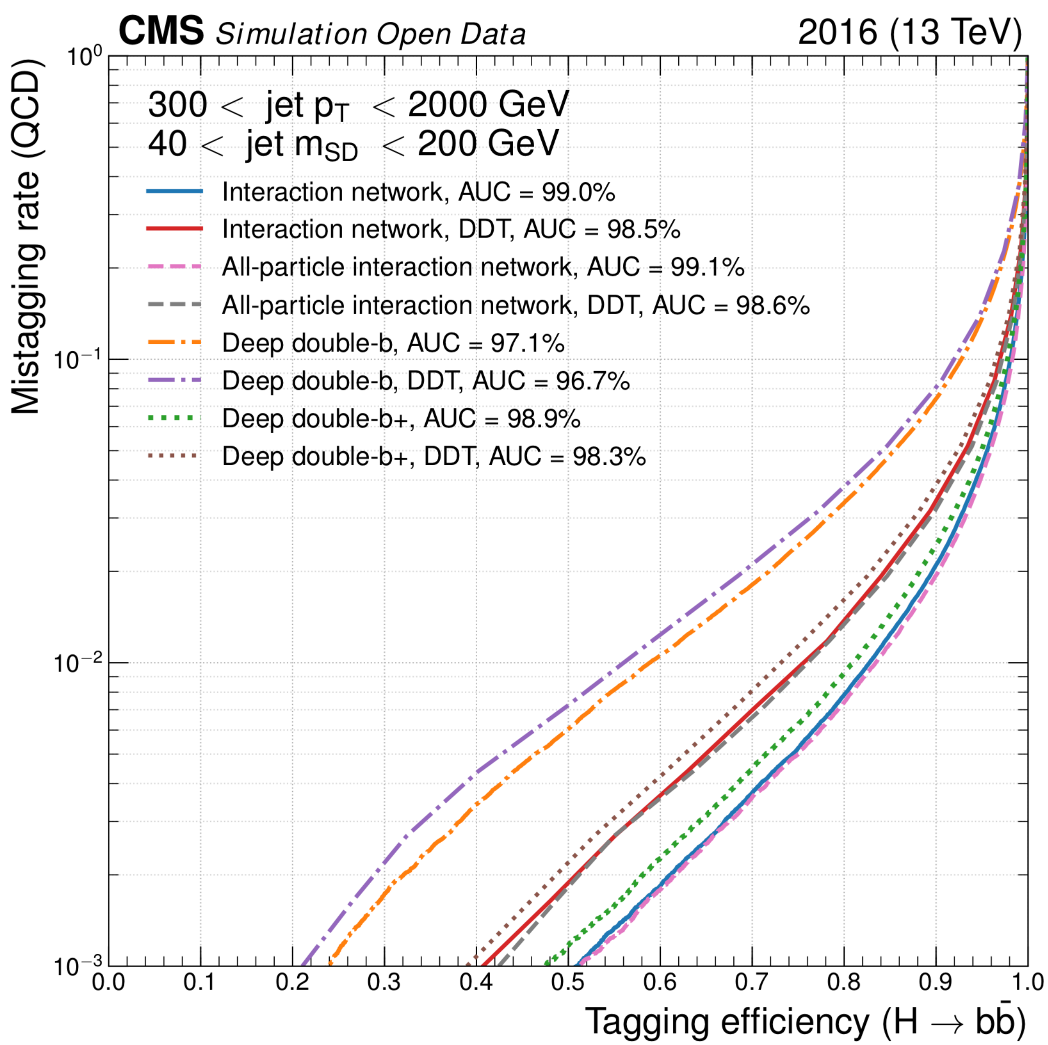
A published technique paper describes the tool.

You **reproduce 2 figures from that paper during Fall Quarter**.

You get to decide among multiple different directions how to **extend that tool in winter quarter**.

Technical Details are next.

The Paper to reproduce



You get to reproduce Figure 4 of this publication

The next few slides explain the context, and figure.

https://arxiv.org/abs/1909.12285

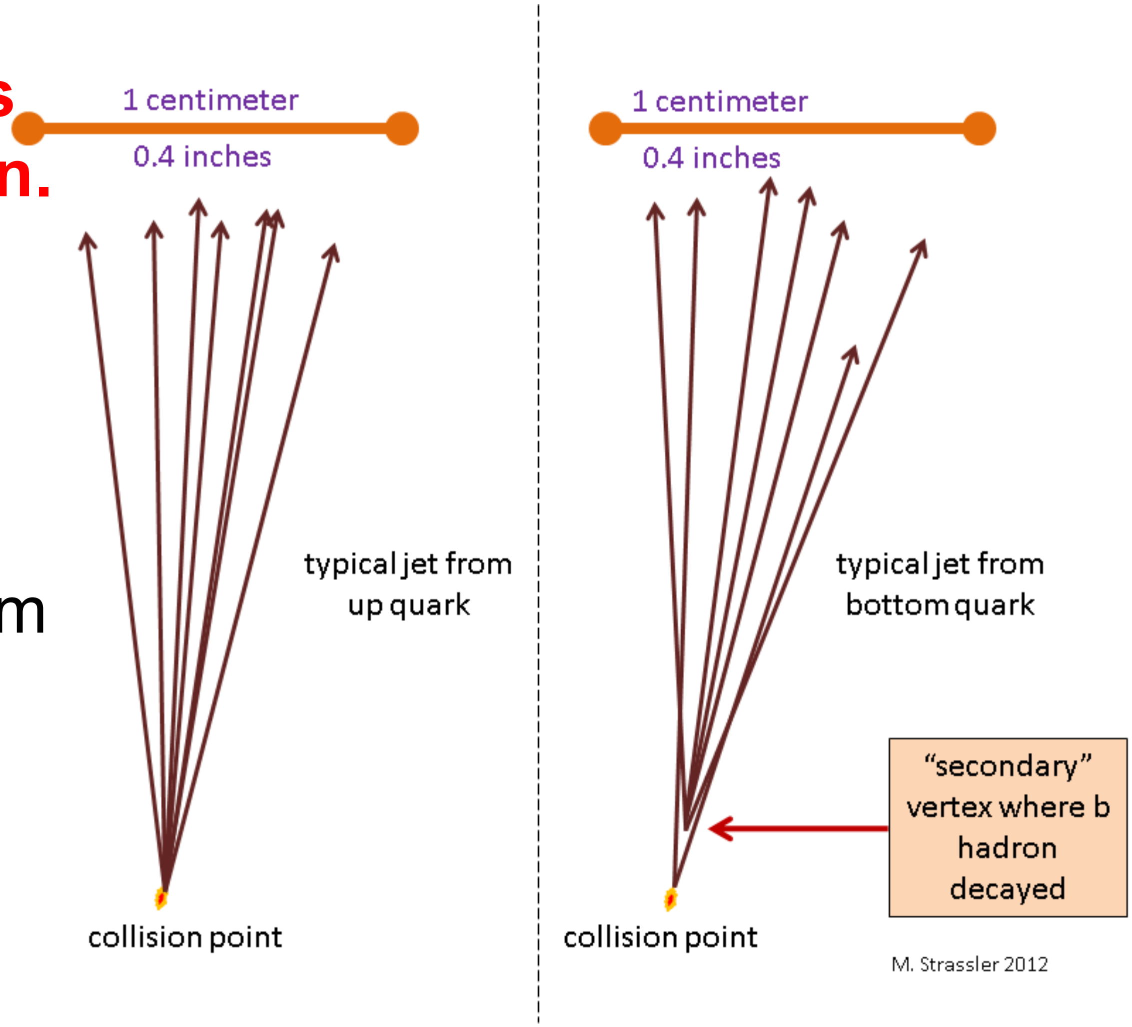
What is a "jet"

https://en.wikipedia.org/wiki/Jet_(particle_physics)

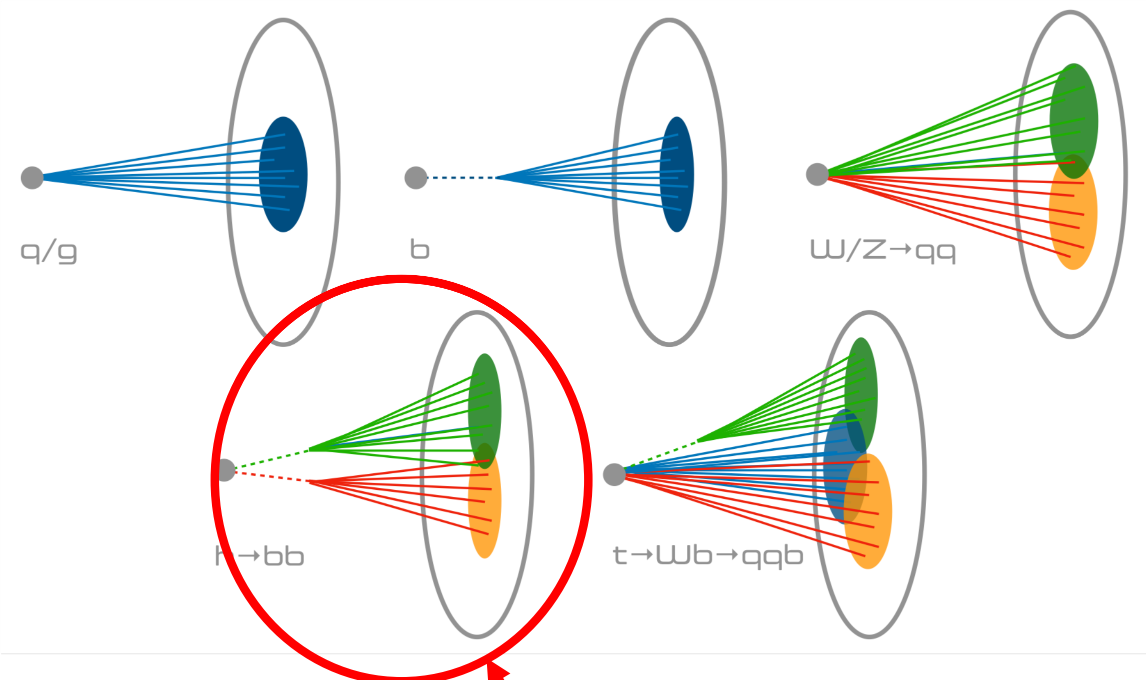
Unfortunately, that's not very helpful for our purposes.

A jet is a spray of particles
Going in the same direction.

We detect the particles,
their direction (momenta),
And reconstruct the "jet" via
some sort of "cone" algorithm
that sums the momentum
in that cone



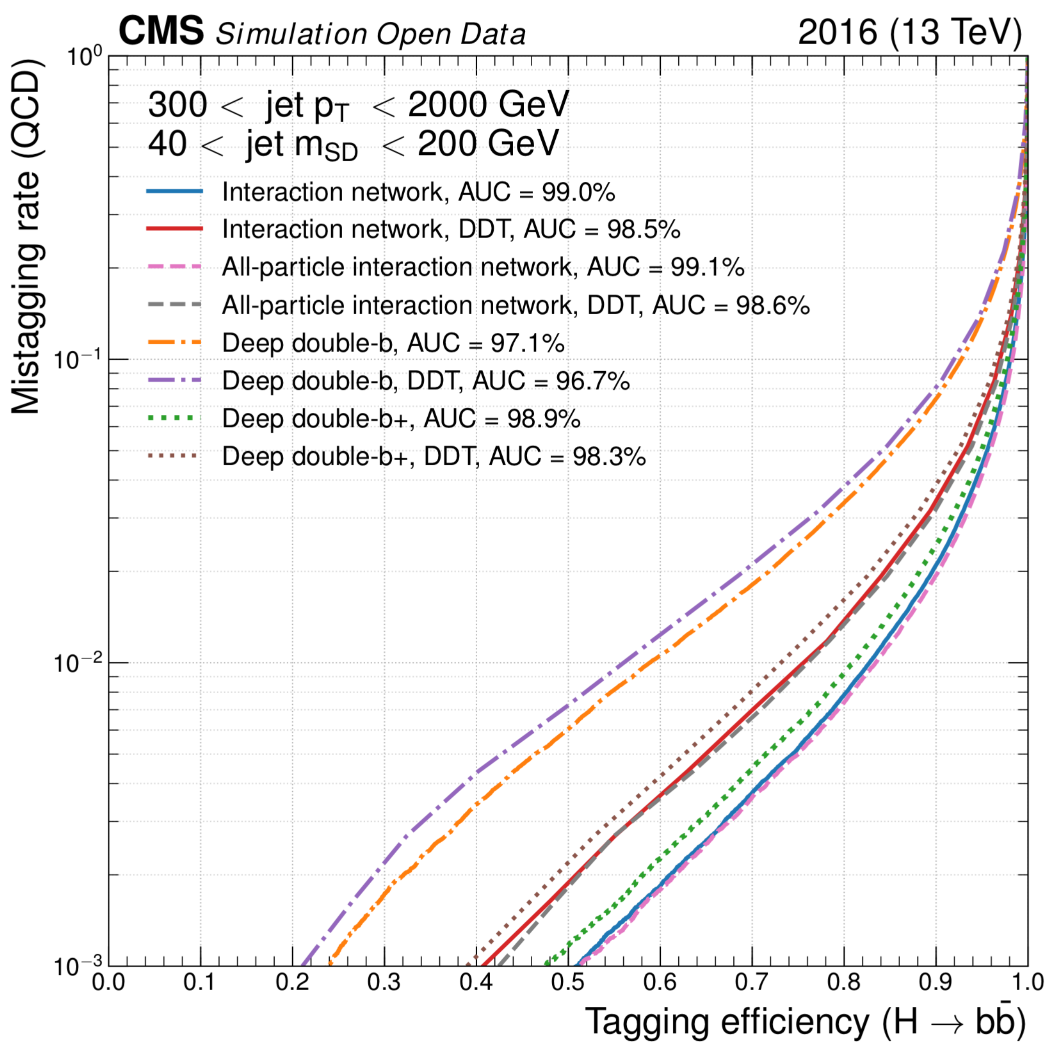
Relevance to “The Paper”



The paper describes how ML was used to design an algorithm that distinguishes a H->bb jet from other types of jets.

We want to distinguish this from all the others !!!

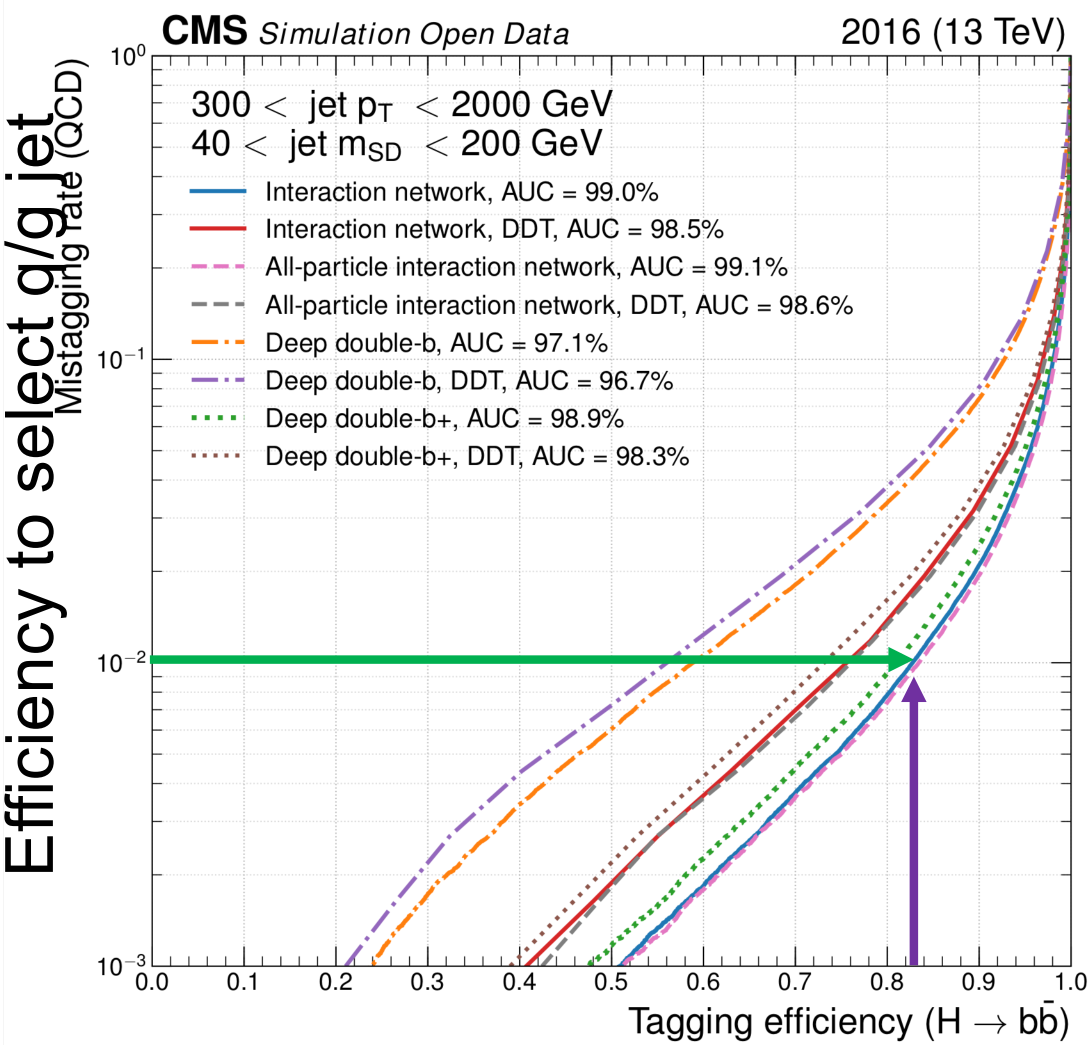
The Figure to reproduce



Efficiency to select q/g jet

Efficiency to select H->bb

How to read this Figure



Efficiency to select H->bb

For any given cut on the Interaction network output, you get an efficiency to select each type of jet.

E.g.:
0.01 for q/g jets
and
0.83 for H->bb jets

Any Questions ?

Some YouTube videos for your amusement.

- https://www.youtube.com/watch?v=FMH3T05G_to
- https://www.youtube.com/watch?v=_wxFYcEuLYo
- https://www.youtube.com/watch?v=328pw5Taeg0
- https://www.youtube.com/watch?v=LraNu_78sCw