Victory of Machine Learning in High Energy Physics

Tae Jeong Kim (Hanyang)

Dec. 1 in 2023

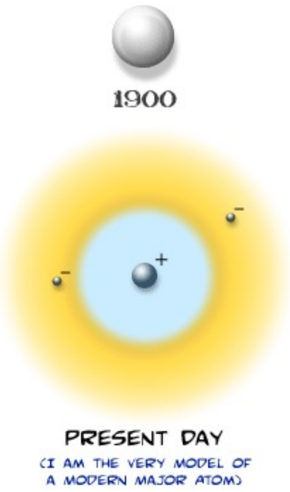
For High Energy Physics and AI workshop

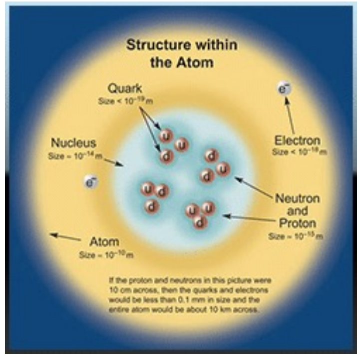
Outline

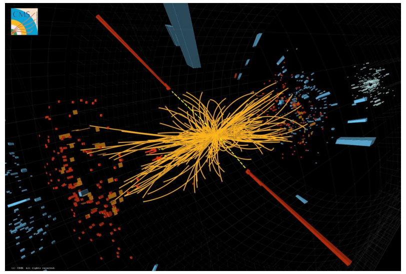
- Introduction to High Energy Physics
- LHC
- Standard model
- Discovery of single top quark, Higgs boson and its coupling
- Challenges

High Energy Physics

Zeptospace: 10^{-21} m



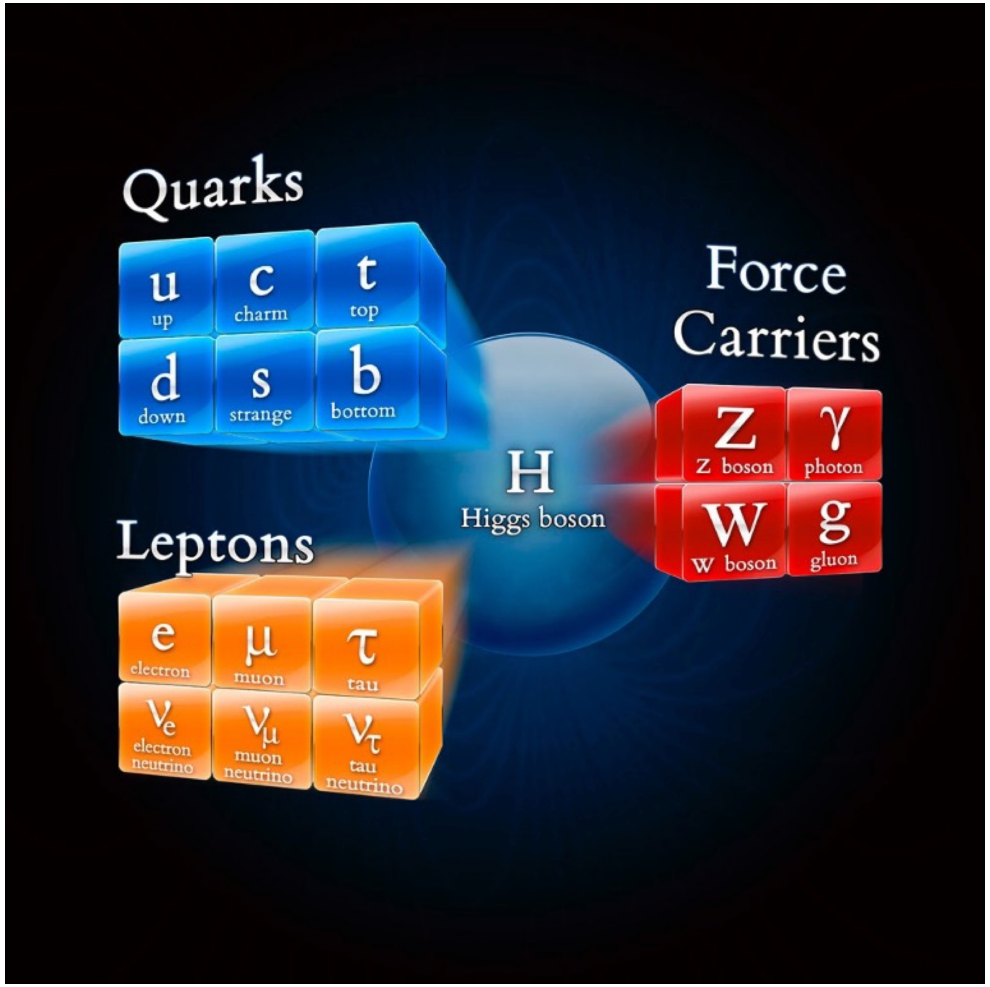




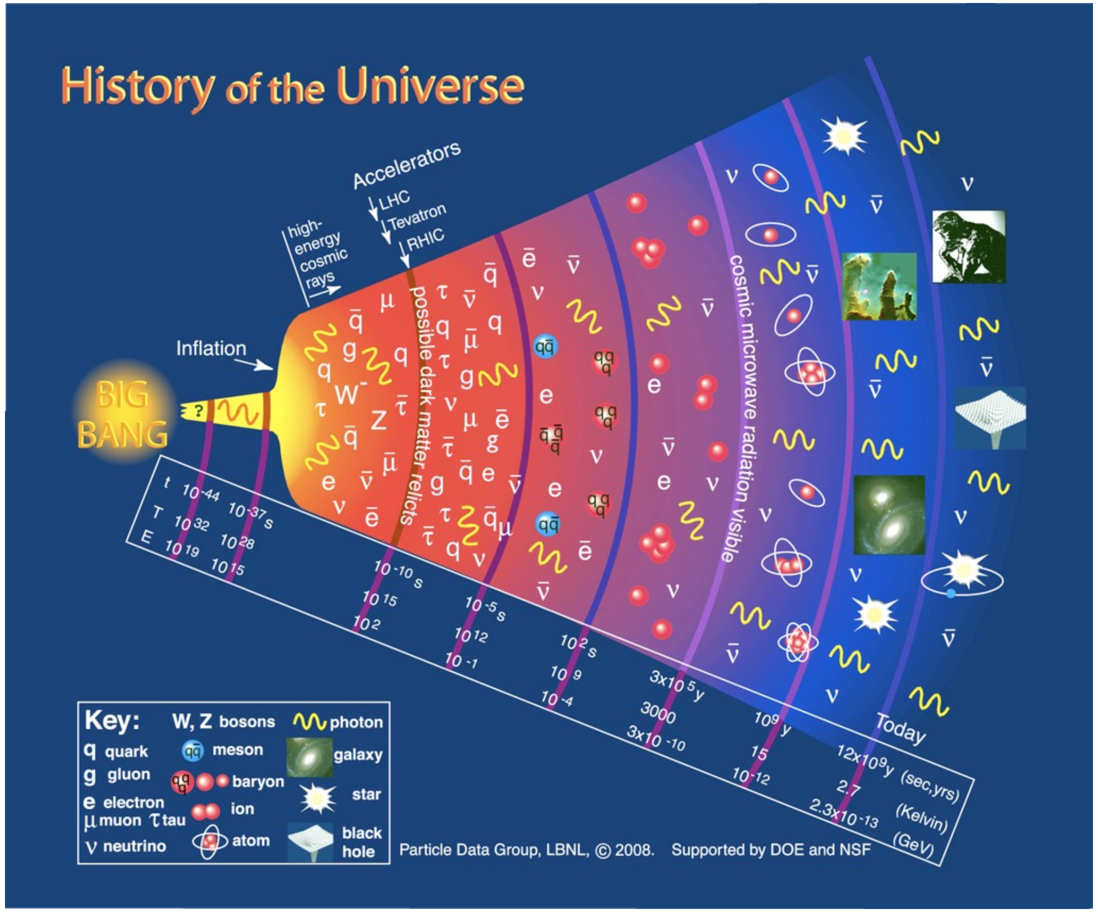
B.C.400 1900 1930 1960 1995

Standard Model

- Electromagnetic interaction (photon)
  - Electric charge
  - $\alpha \sim 1/137$
- Weak interaction (W,Z)
  - Fundamental particle decay
  - $\alpha_W \sim 1/40$
- Strong interaction (gluon)
  - Color charge
  - $\alpha_S \sim 1$

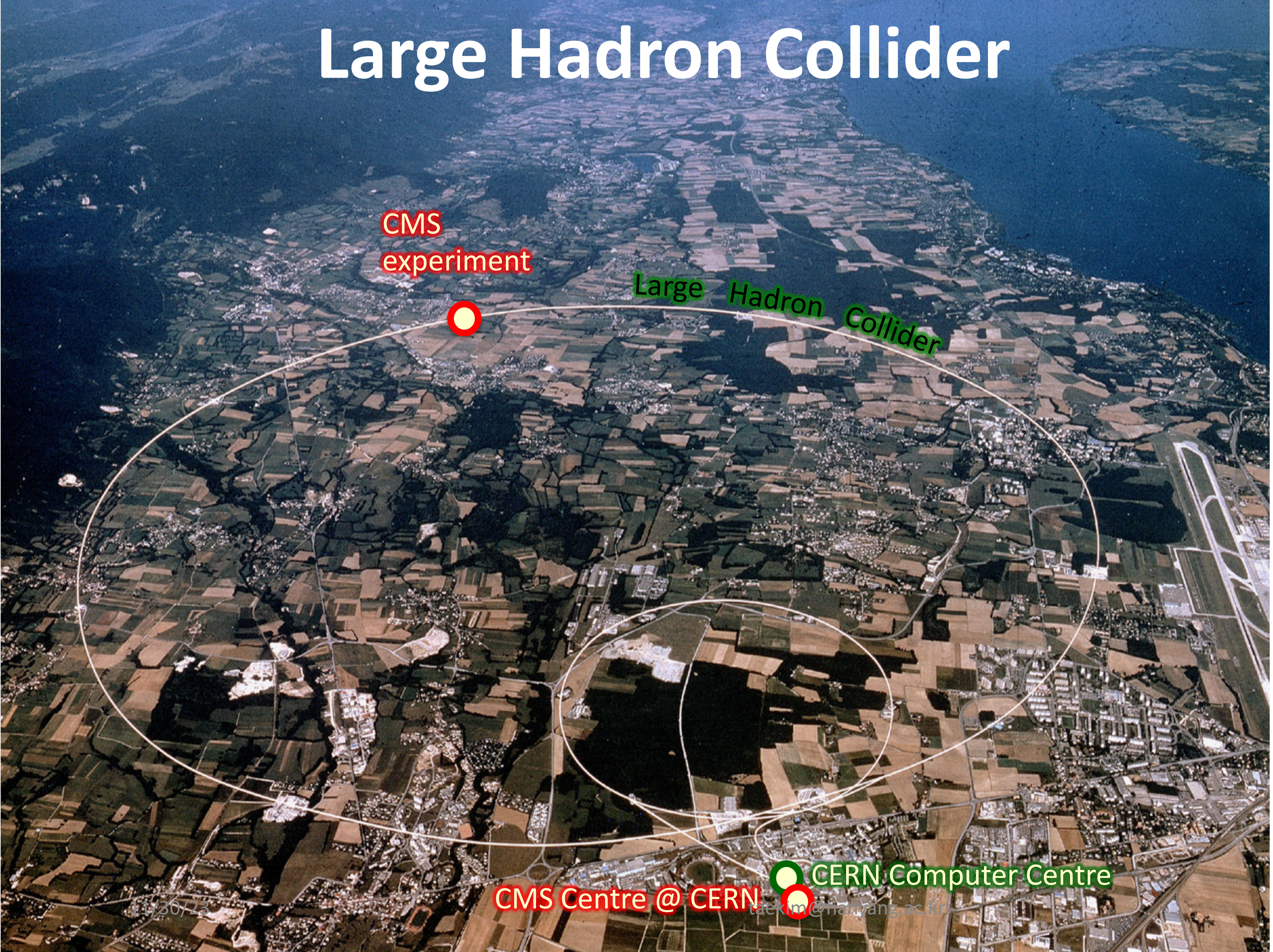


Understand the very first moment of our Universe



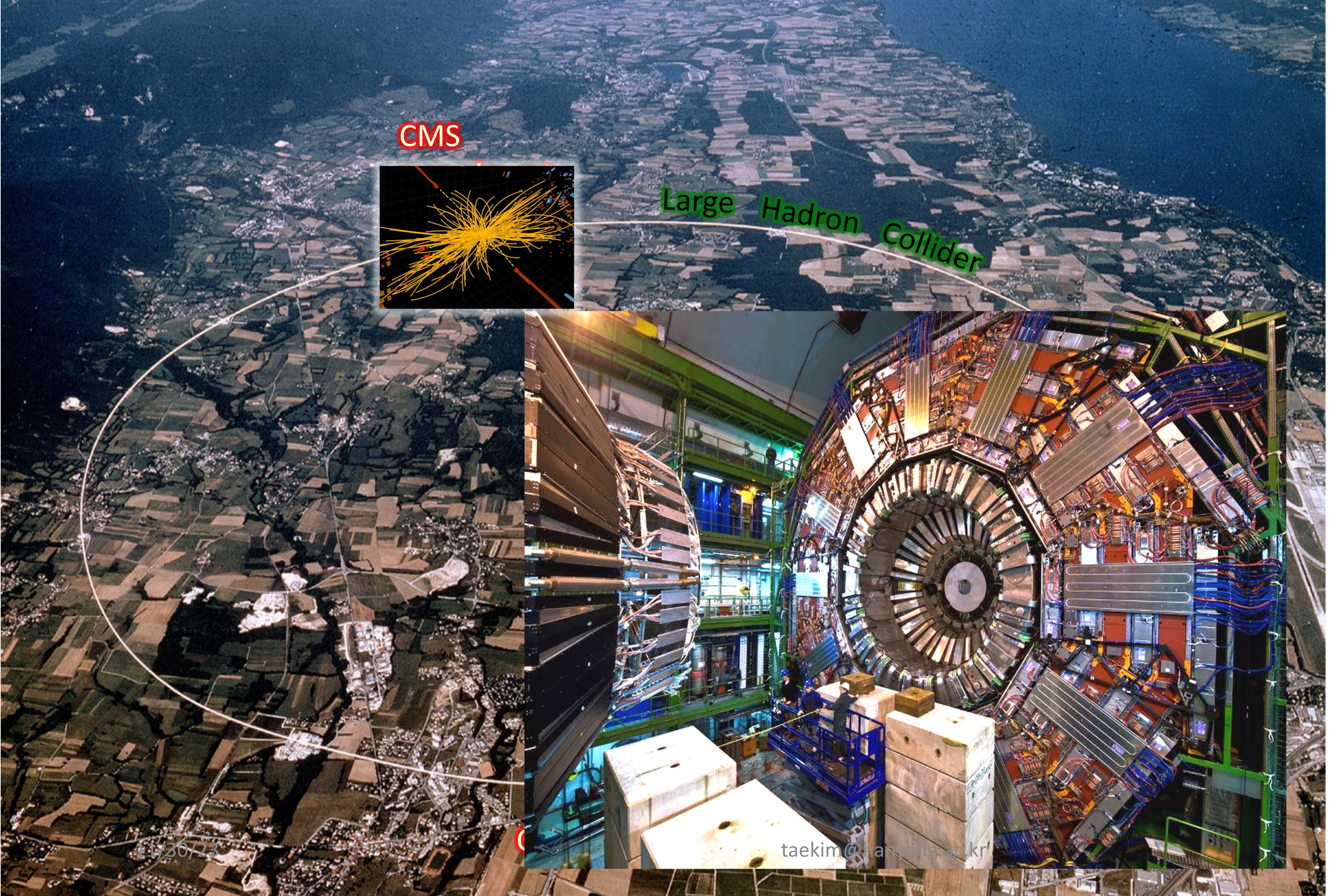
- Using powerful telescope, we can only look back until cosmic microwave radiation visible
- Using LHC, we can recreate conditions shortly after Big Bang and study them
- Need high energy to probe hotter space: $\lambda = \frac{h}{p}$

Large Hadron Collider

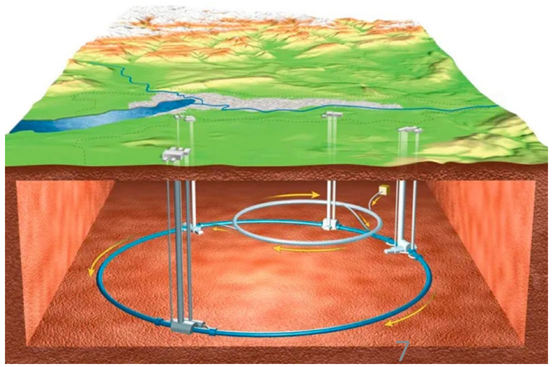


- 27 km circumference
- 9593 superconducting magnets at 1.9 K
- 120 tonne of liquid helium
- Accelerates beams of protons to 99.9999991% the speed of light
- Proton beams circulate 11245 times/sec
- 1 billion of p-p collisions/second
- Collisions with a center of mass energy 13.6 TeV

Compact Muon Solenoid (CMS)

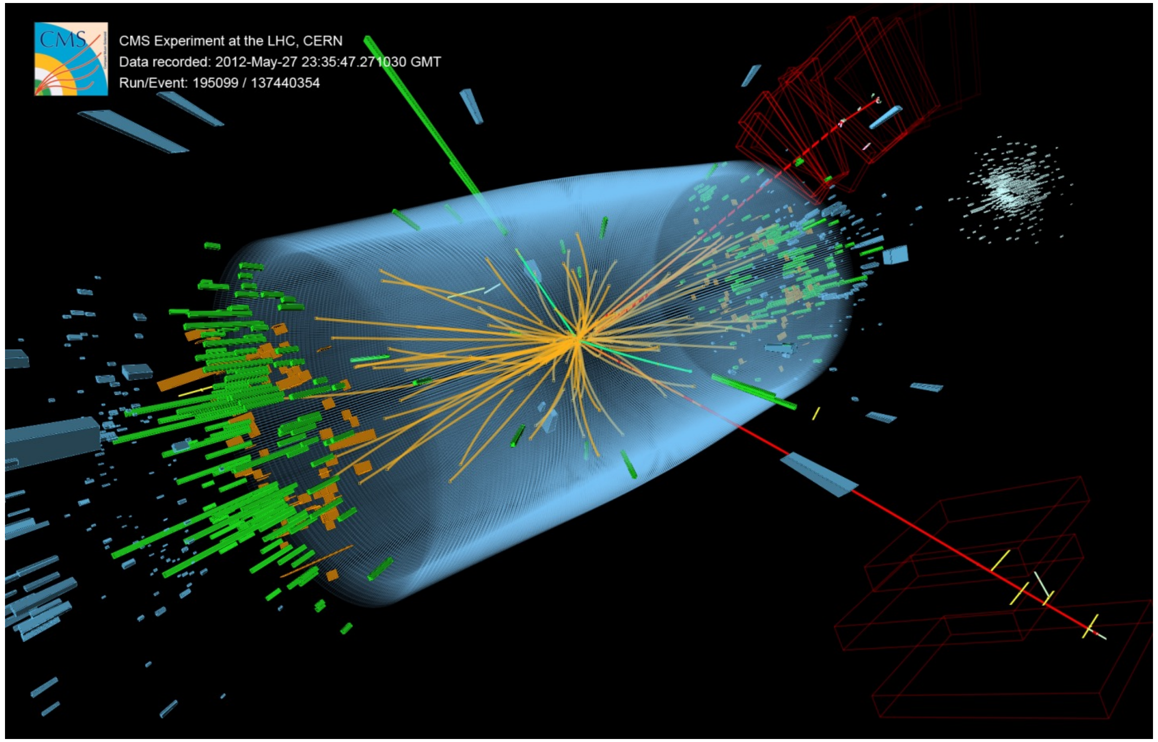


- Detects new particles created in LHC collisions
- 21 m long and 15 m high in a huge cavern 100 m underground
- 12,500 tonne

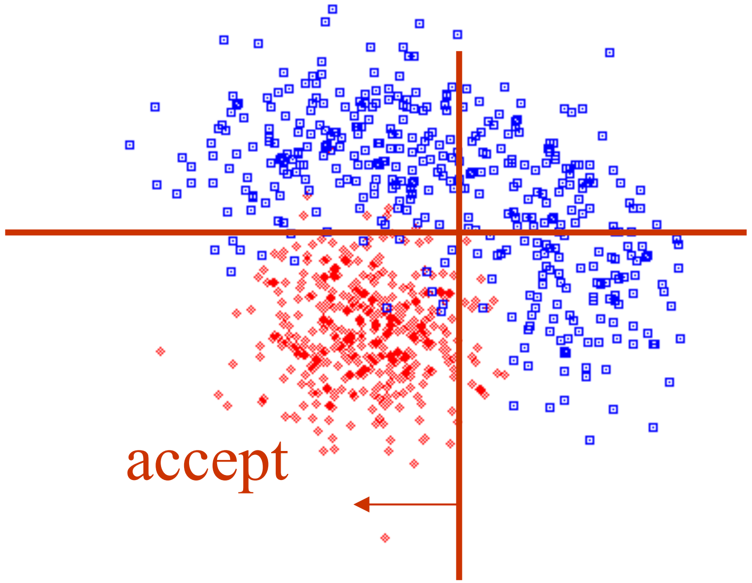


LHC data size

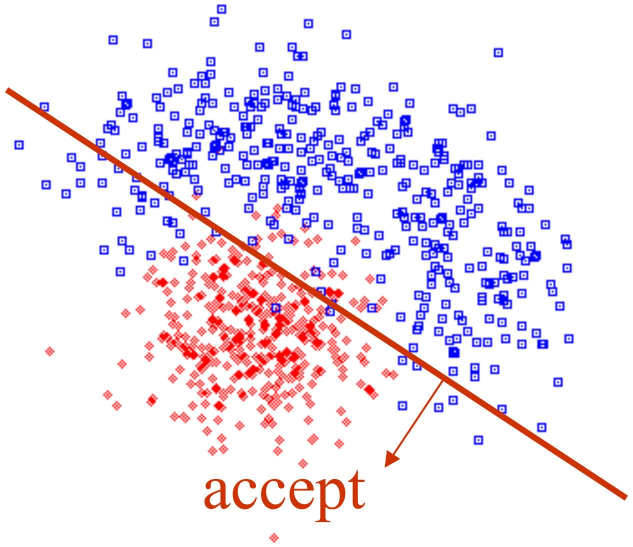
- At the LHC, particles collide 1 billion times per second generating one petabyte per second
- Save only interesting events
  - Level 1 trigger ~ 1 kHz
  - High level trigger ~ 100 Hz
- Store 20 petabyte every year



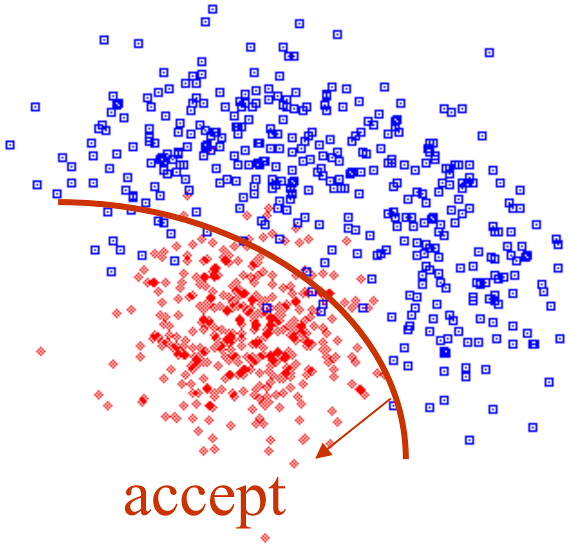
Event selection



traditional way

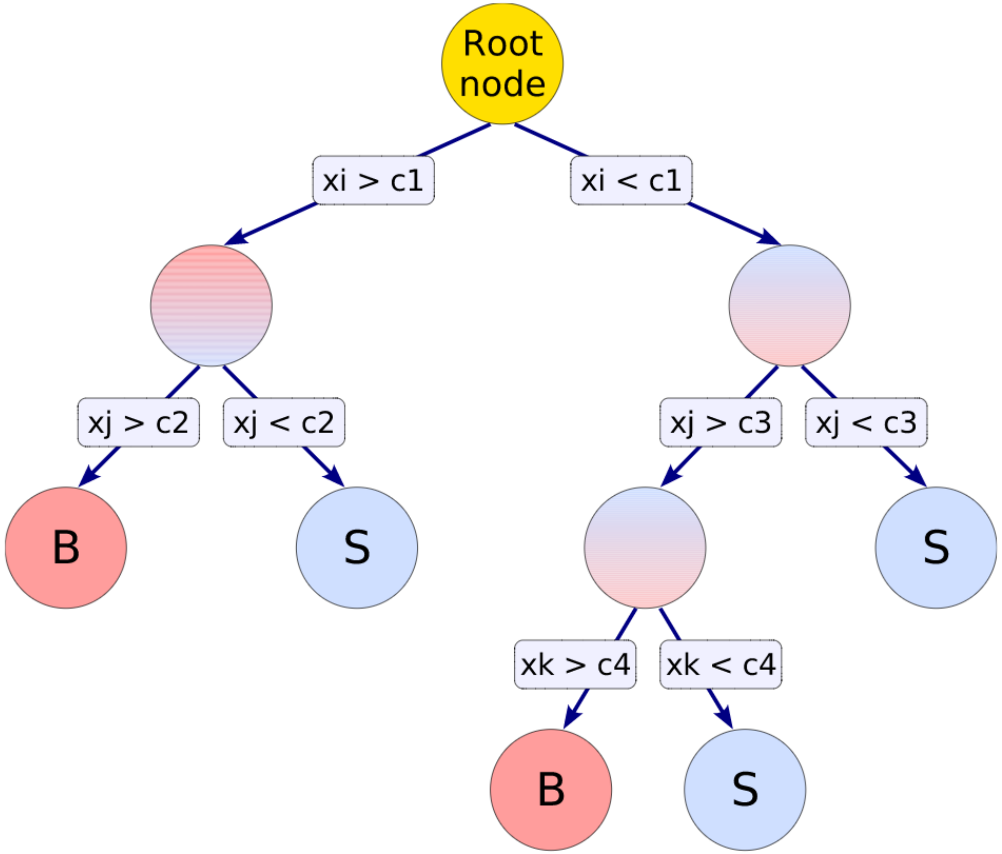


linear

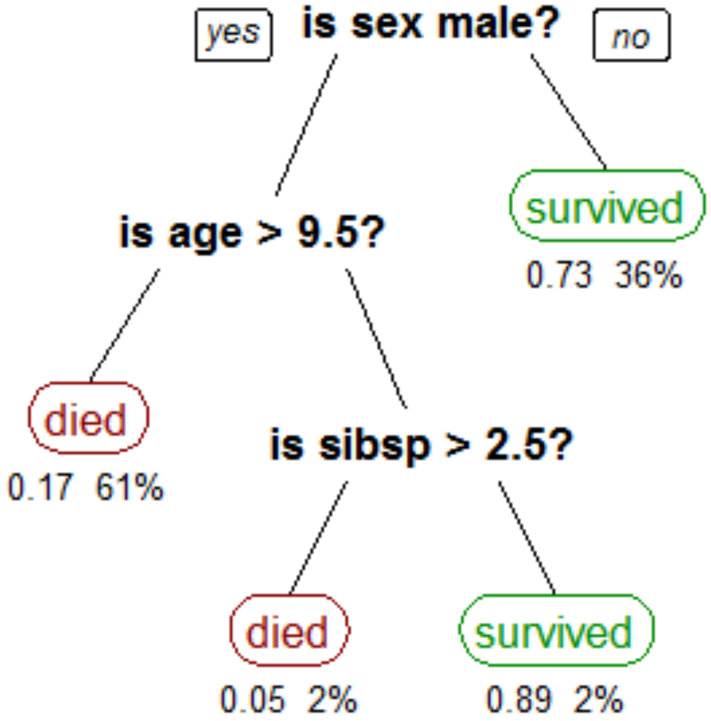


nonlinear

Boosted Decision Tree (BDT)

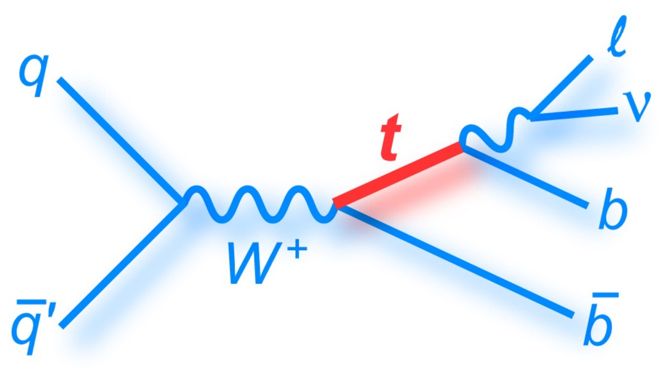


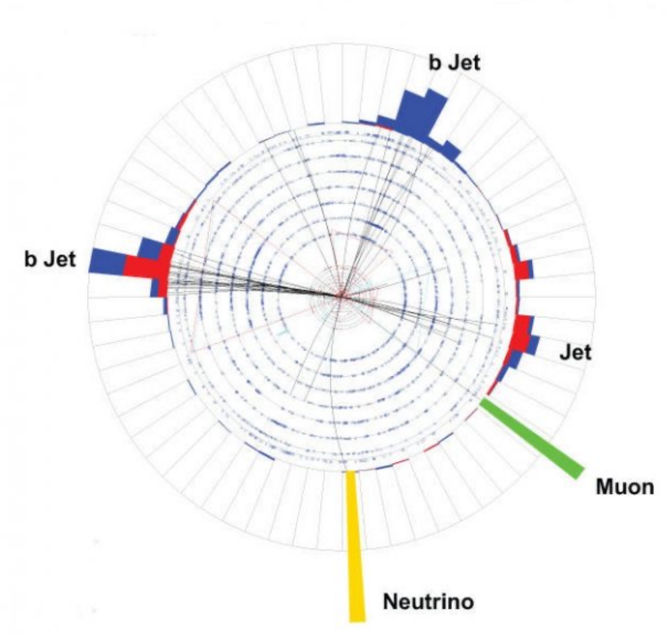
Toolkit for Multivariate Data Analysis with ROOT (TMVA)

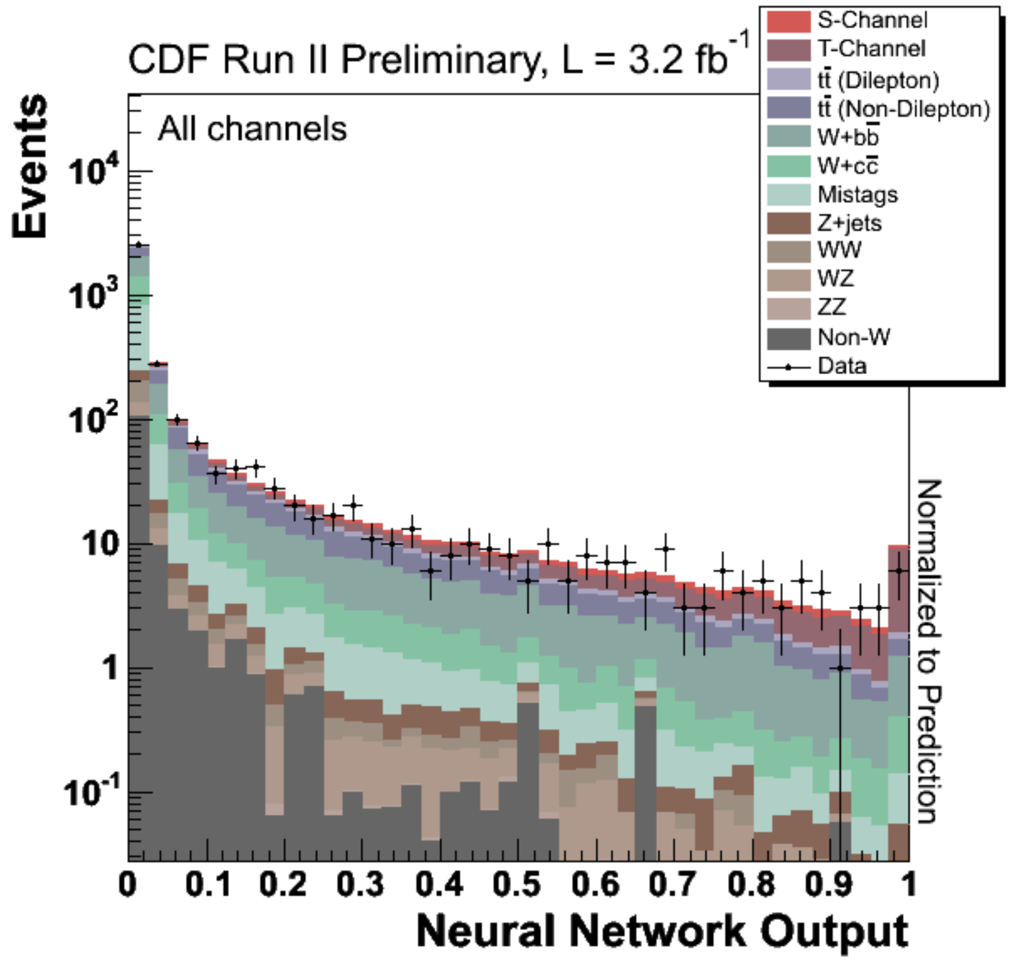


Survival passengers on the Titanic
from WIKIPEDIA

Single top quark discovery with ML

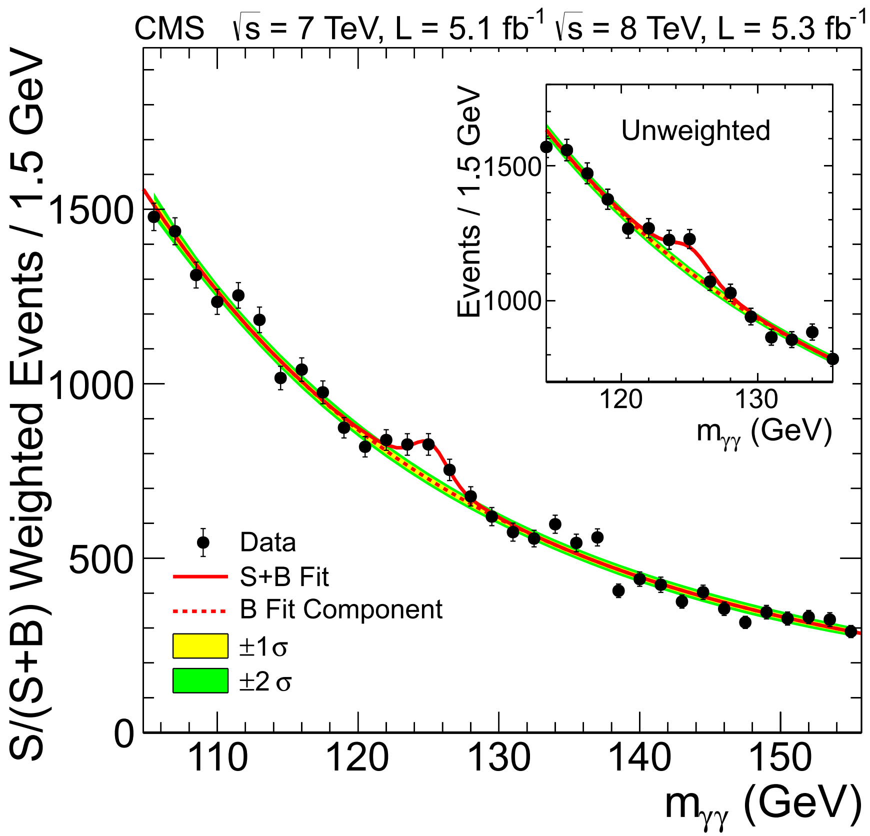






- Small cross section → simple cut and count does not work
- Single top quark was discovered with a help of machine learning technique in 2009

Discovery of Higgs

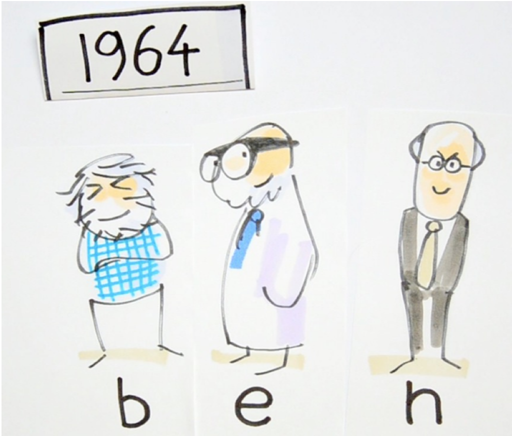


PLB 716 (2012) 30

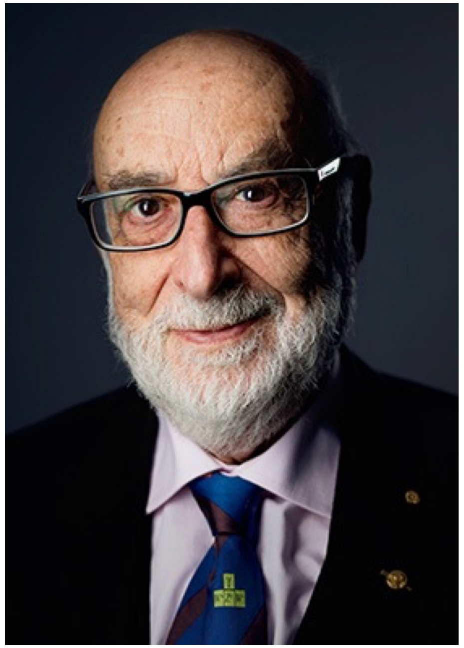
- Higgs predicted in 1964
- Machine learning in 2012
  - Photon energy by regression
  - Photon identification by Boosted Decision Tree (BDT)
  - Multivariate Data Analysis for event classification

Discovery of Higgs

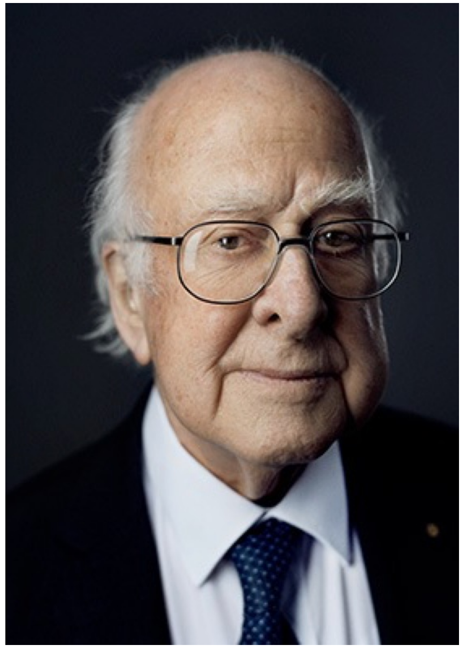
Breakthrough of the year 2012



Nobel prize in physics 2013

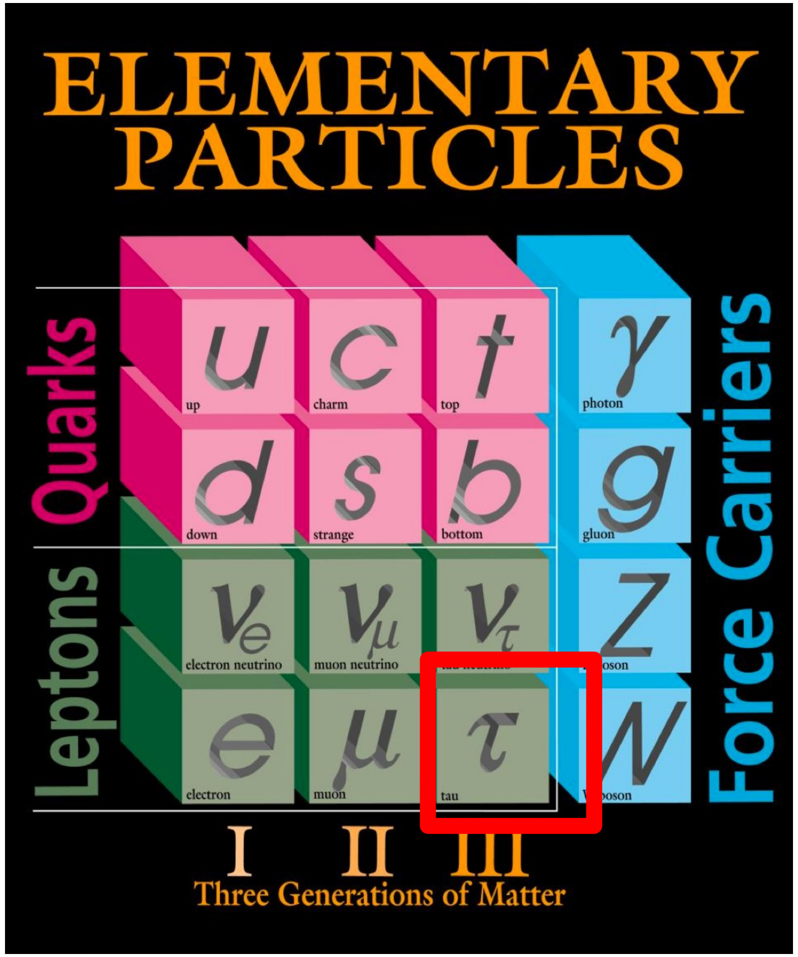


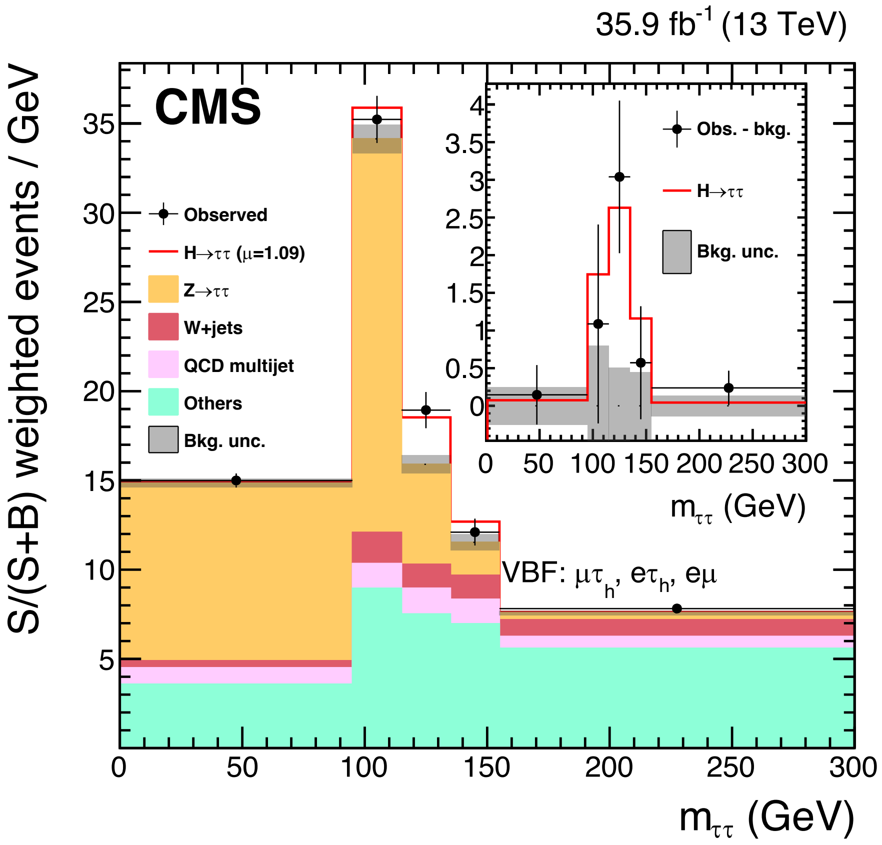
Francois Englert



Peter Higgs

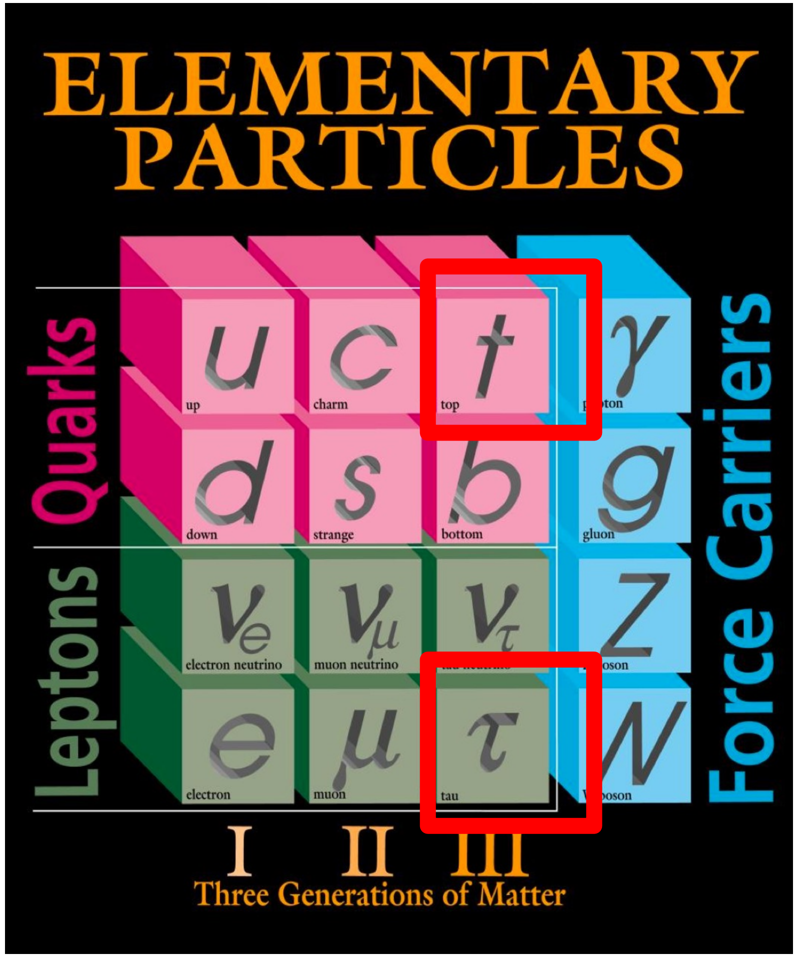
Couling with the third generation particle

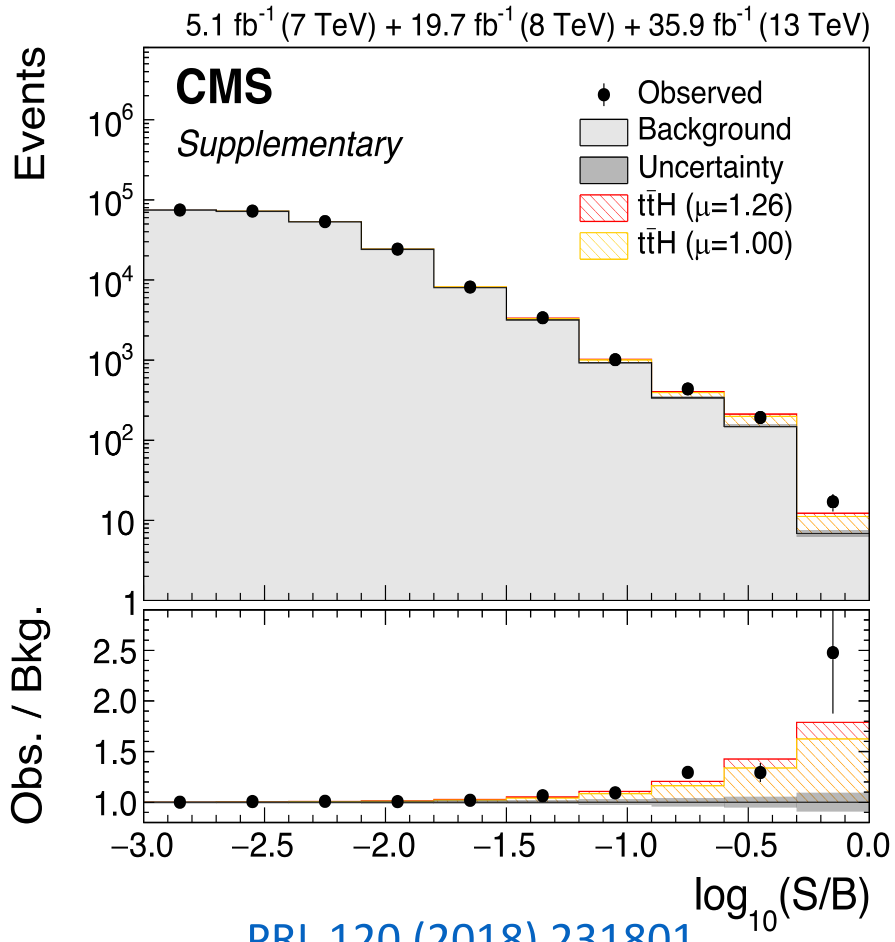




PLB 779 (2018) 283

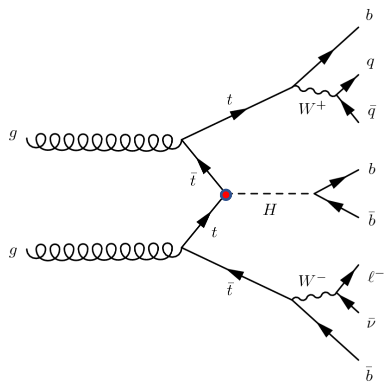
Couling with the third generation particle

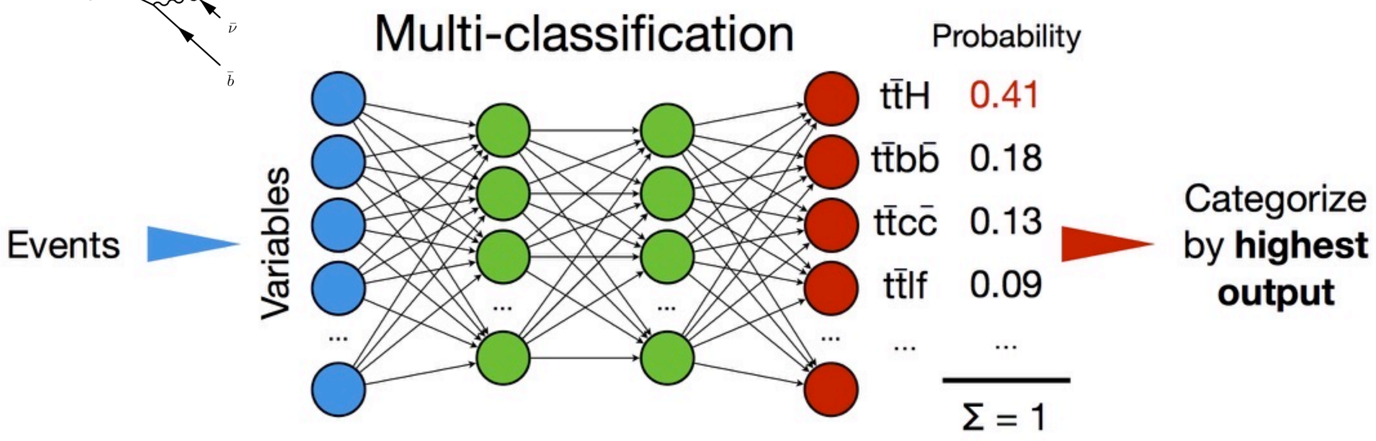


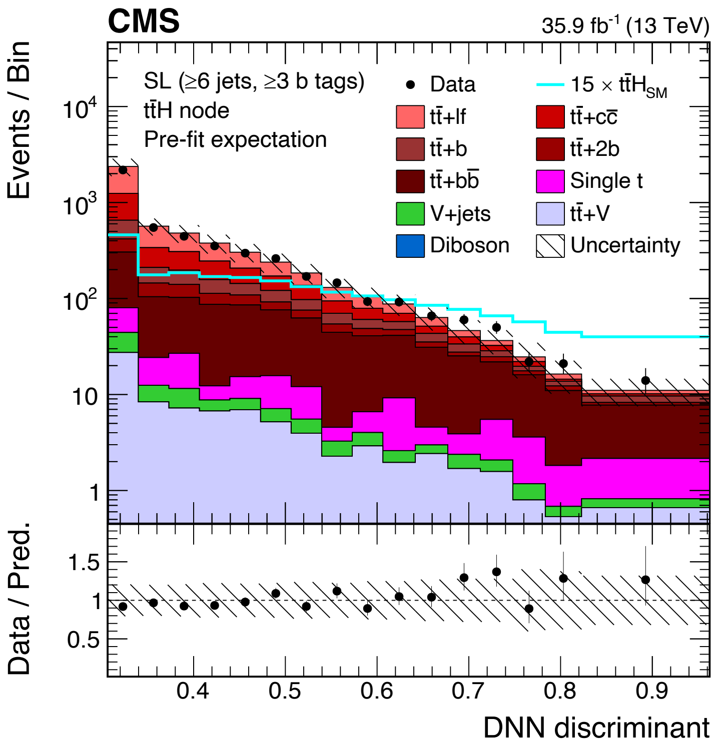


PRL 120 (2018) 231801

$t\bar{t}H(b\bar{b})$ searches for the coupling with top





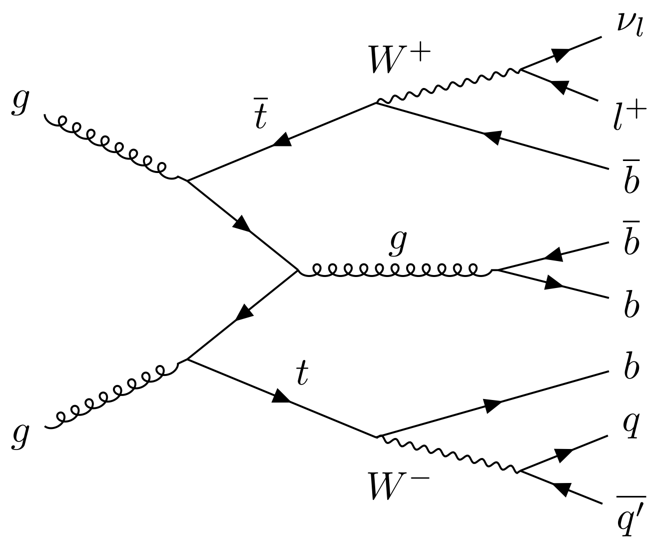


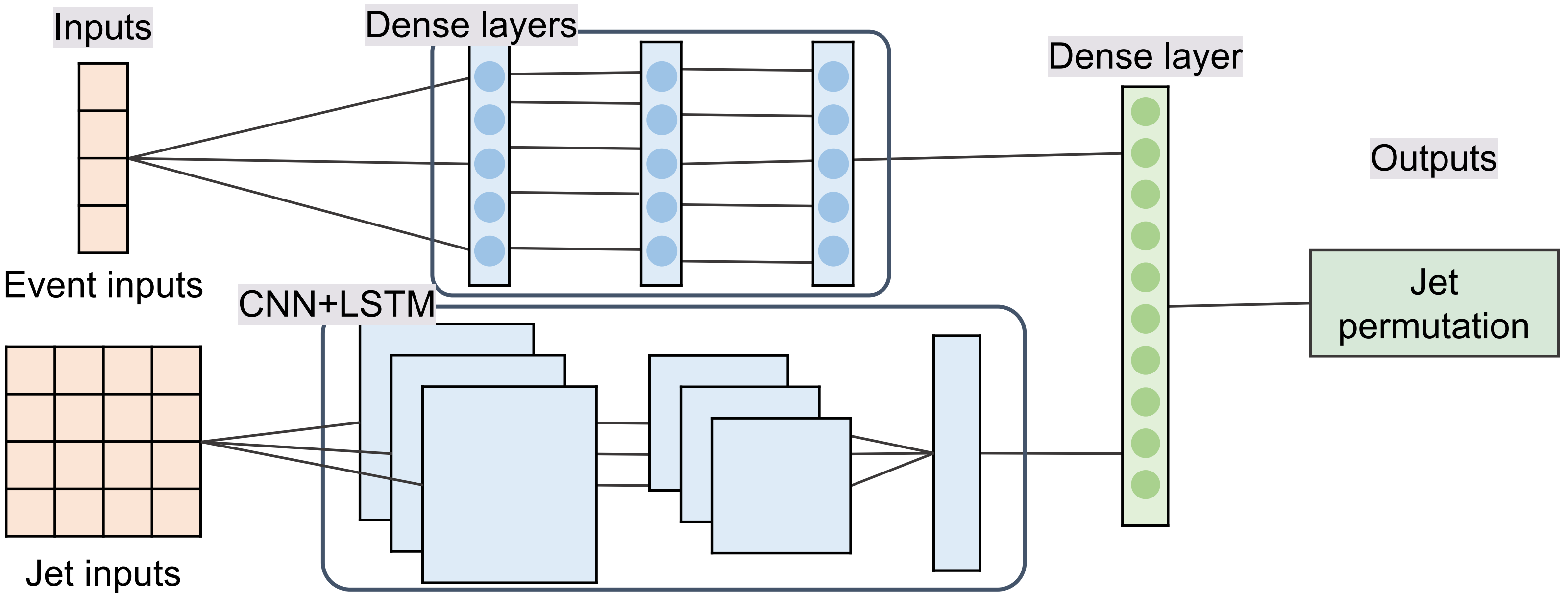
taekim@hanyang.ac.kr

$t\bar{t}b\bar{b}$ cross section measurement

Submitted to JHEP

arXiv:2309.14442





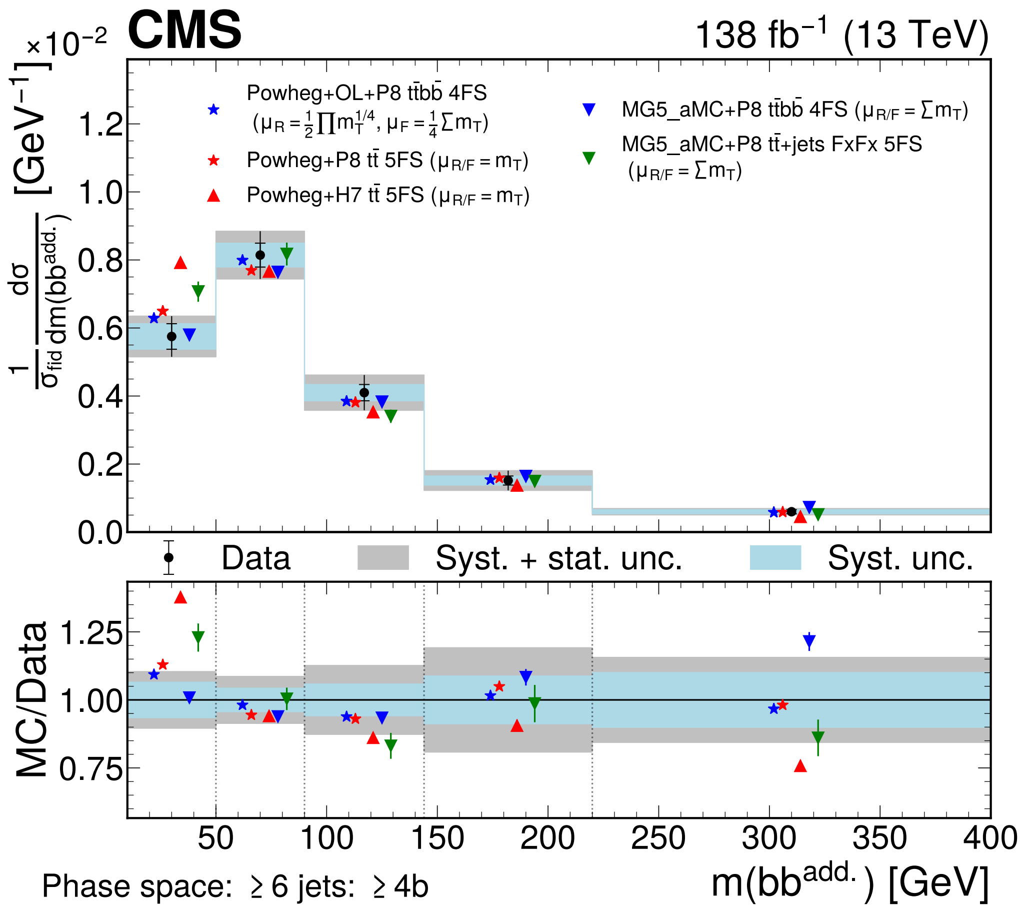
Matching efficiency = 46.9%

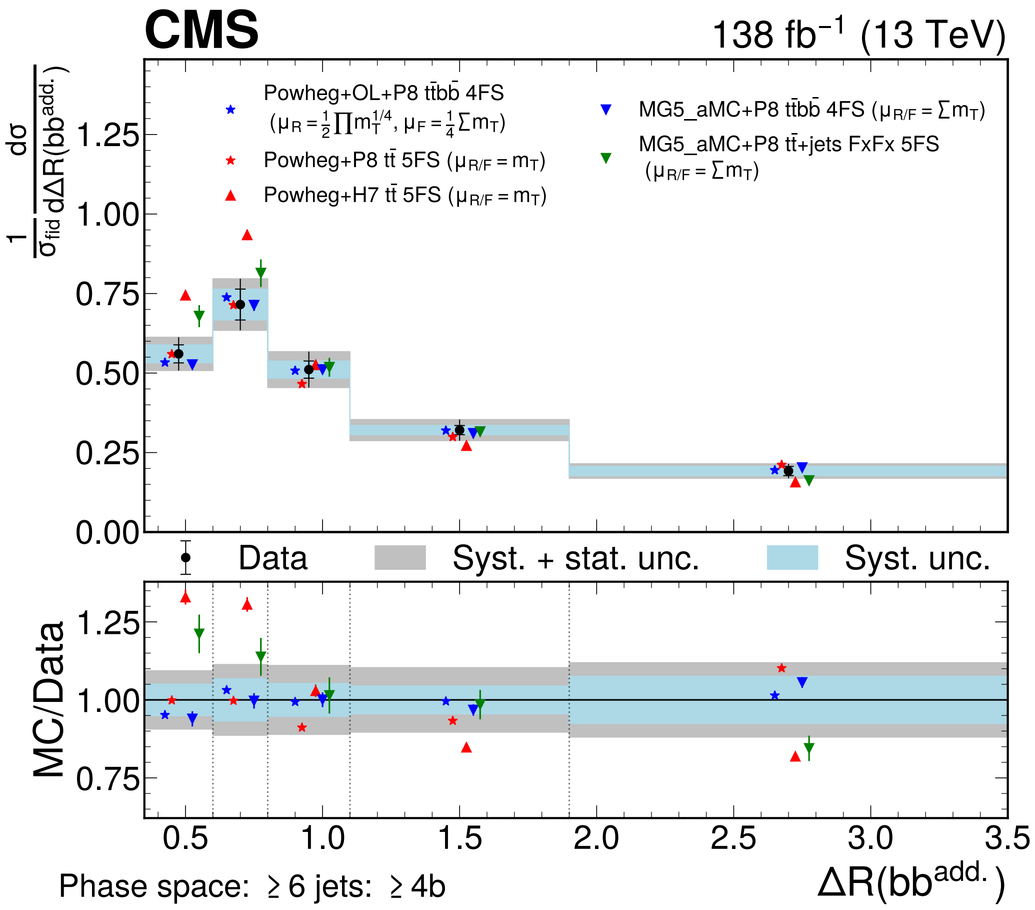
Matching efficiency can go up ~ 60% with surrogate loss functions
EPJ Plus, C. Jang

$t\bar{t}b\bar{b}$ cross section measurement

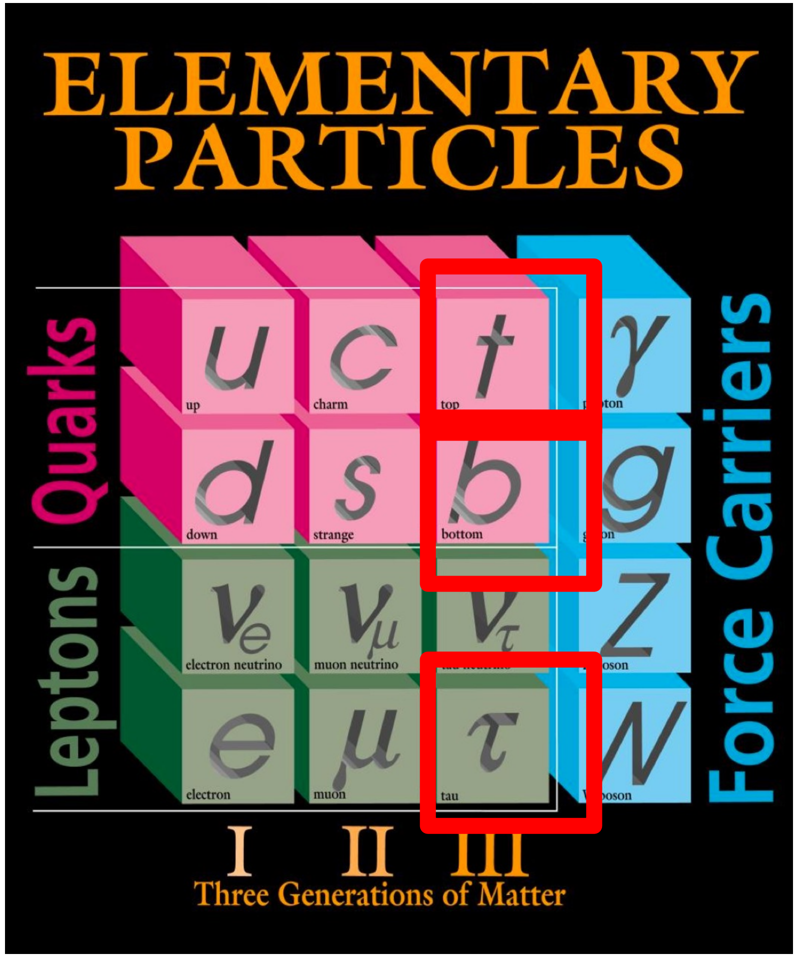
Submitted to JHEP

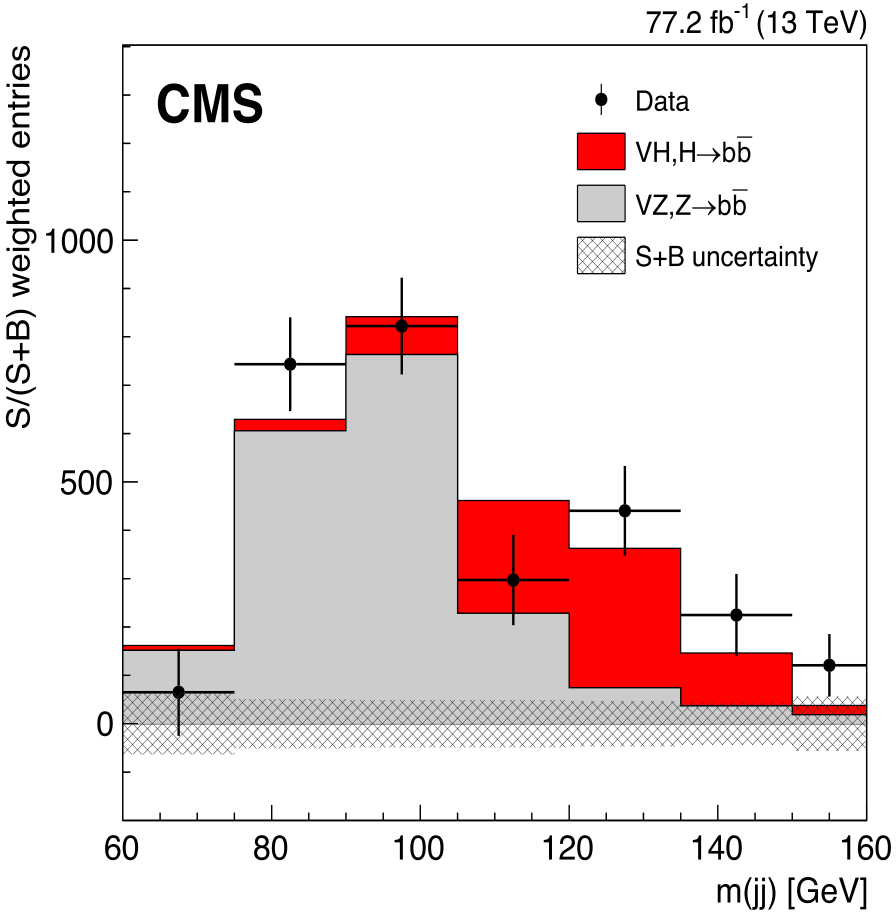
arXiv:2309.14442





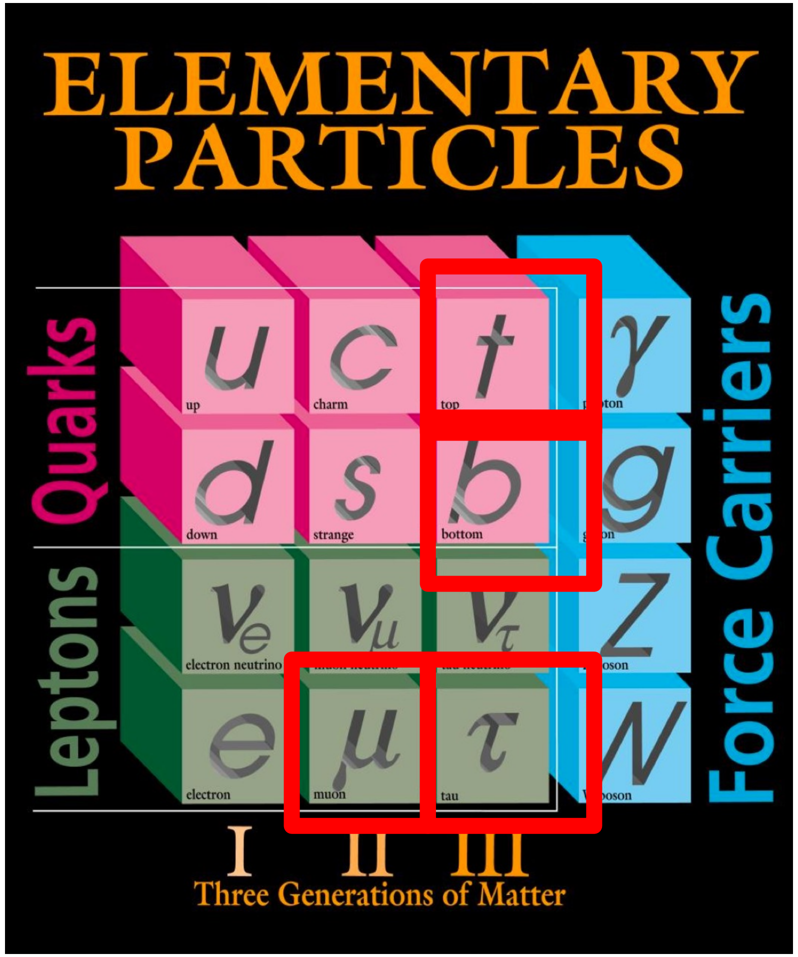
Couling with the third generation particle

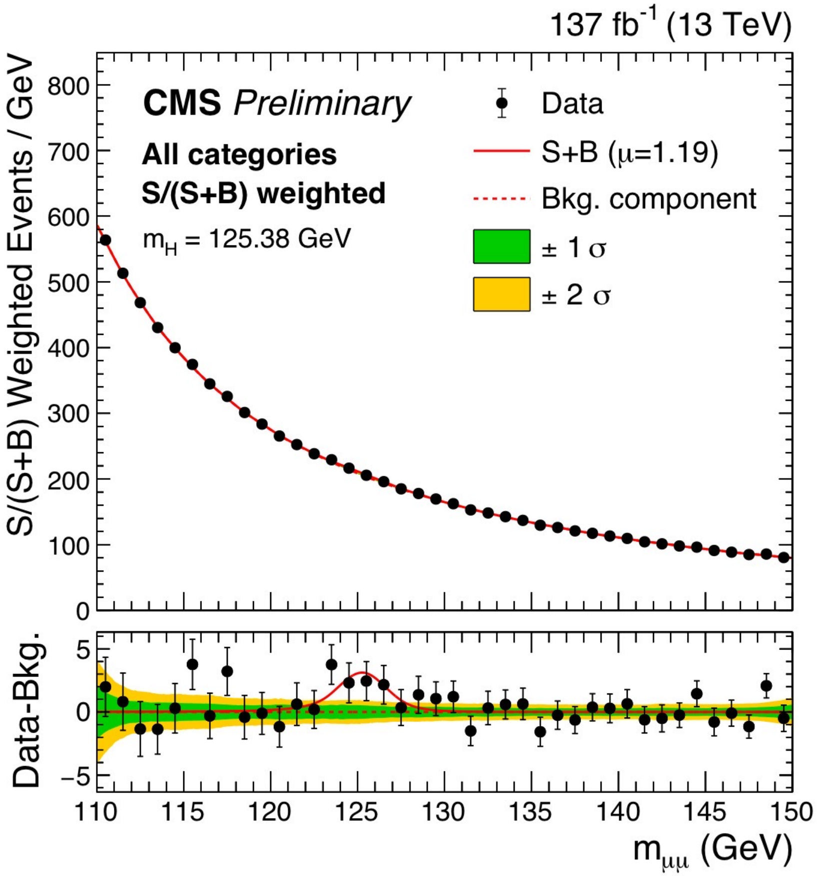




PRL 121 (2018) 121801

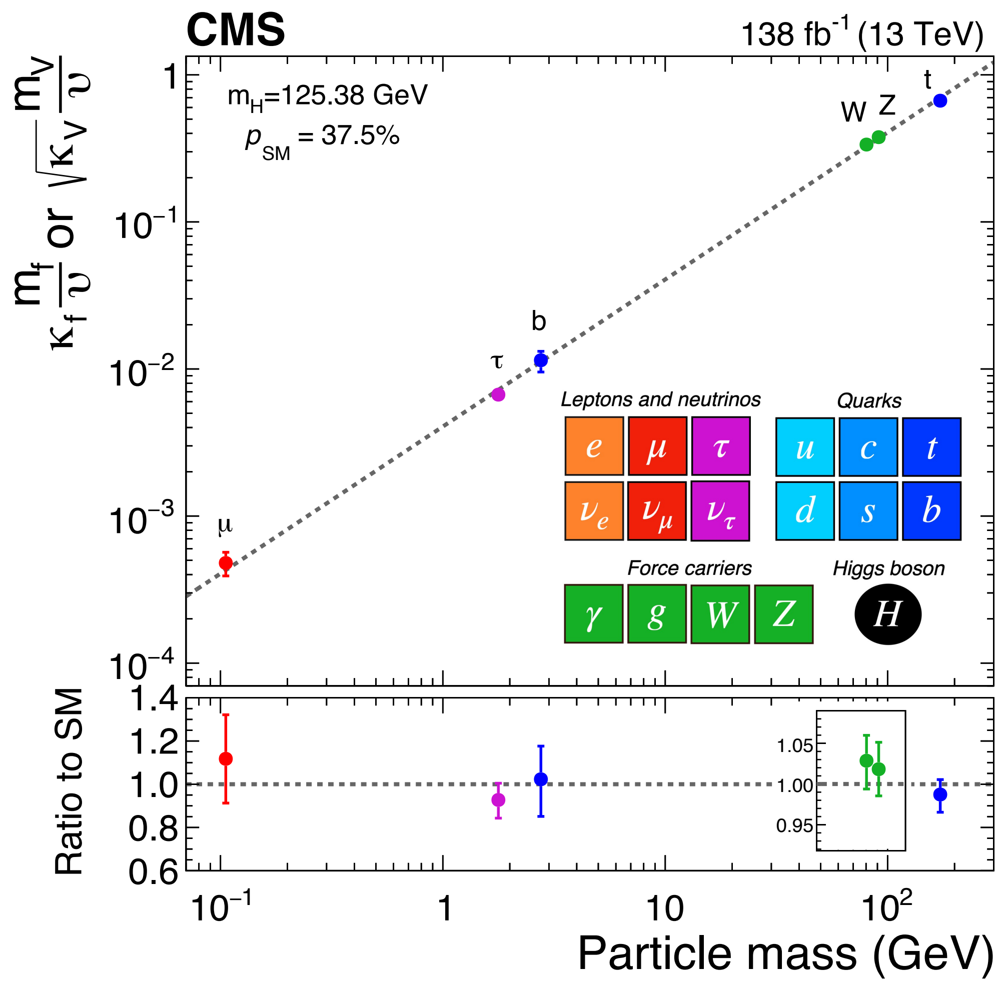
Couling with the second generation particle



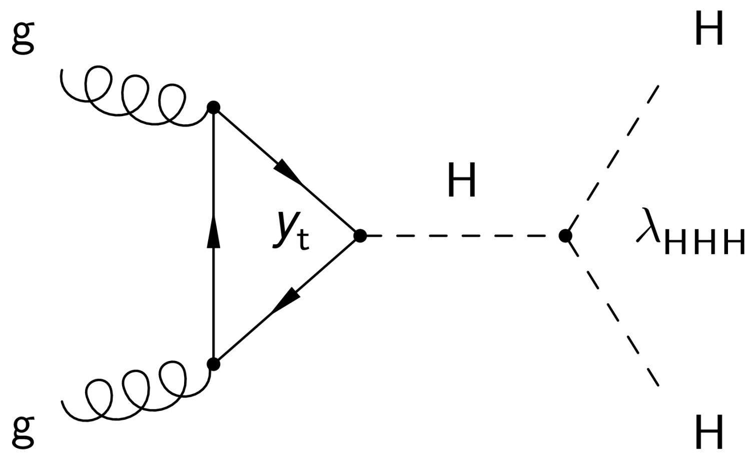


JHEP 01 (2021) 148

Complete Standard Model?



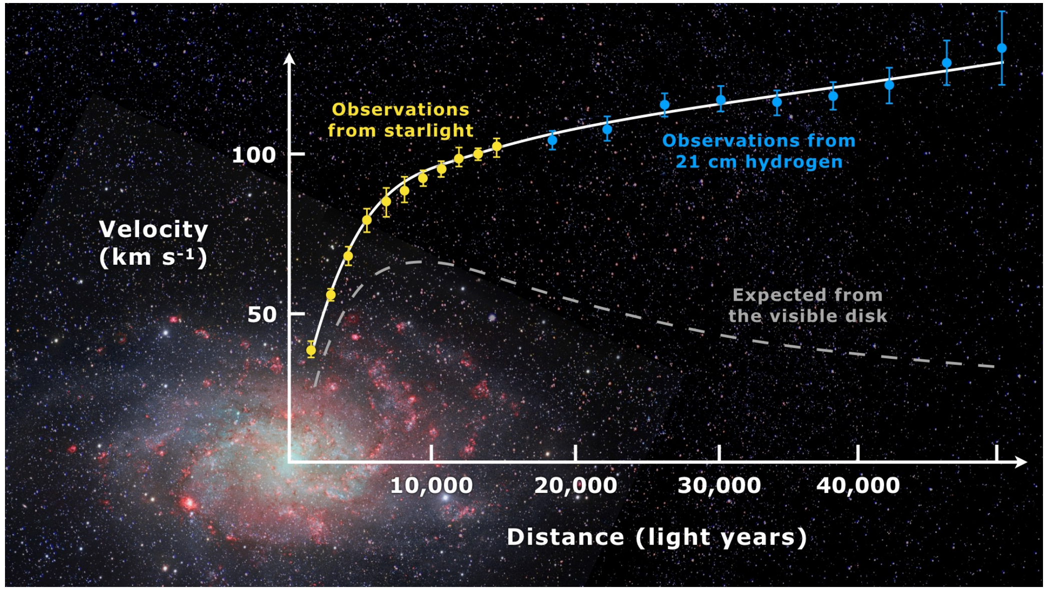
- Higgs self couplig?

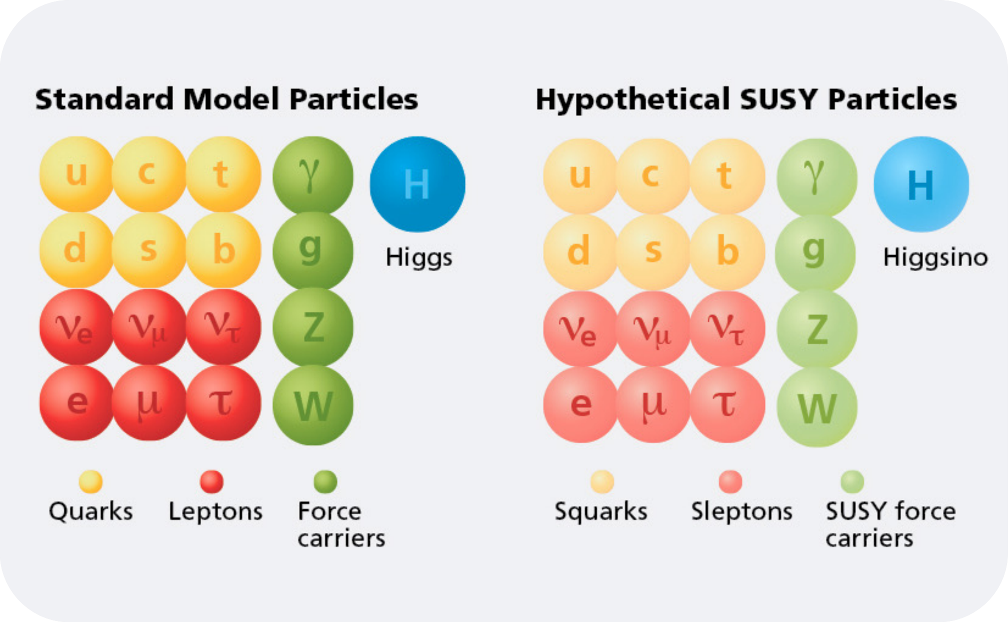


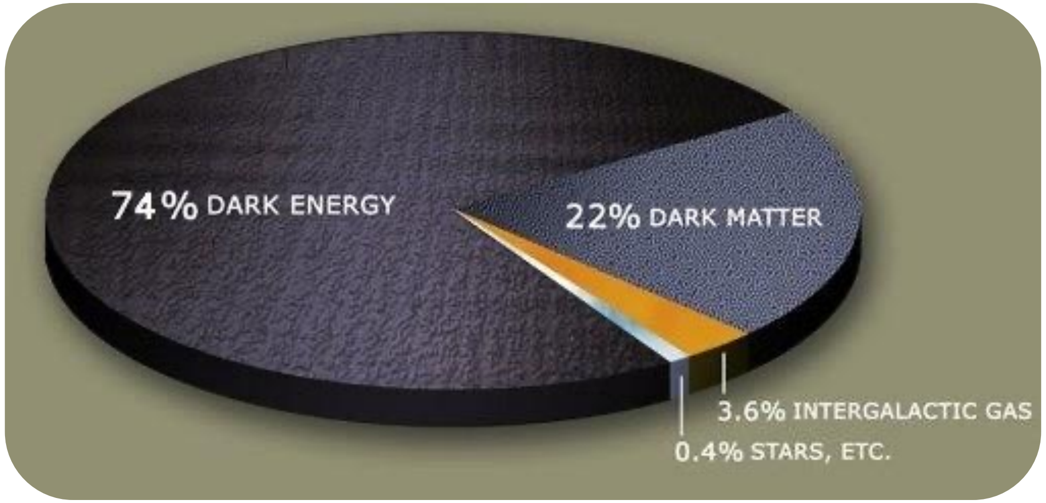
- Many attempts to measure the coupling with AI

Unsolved problems

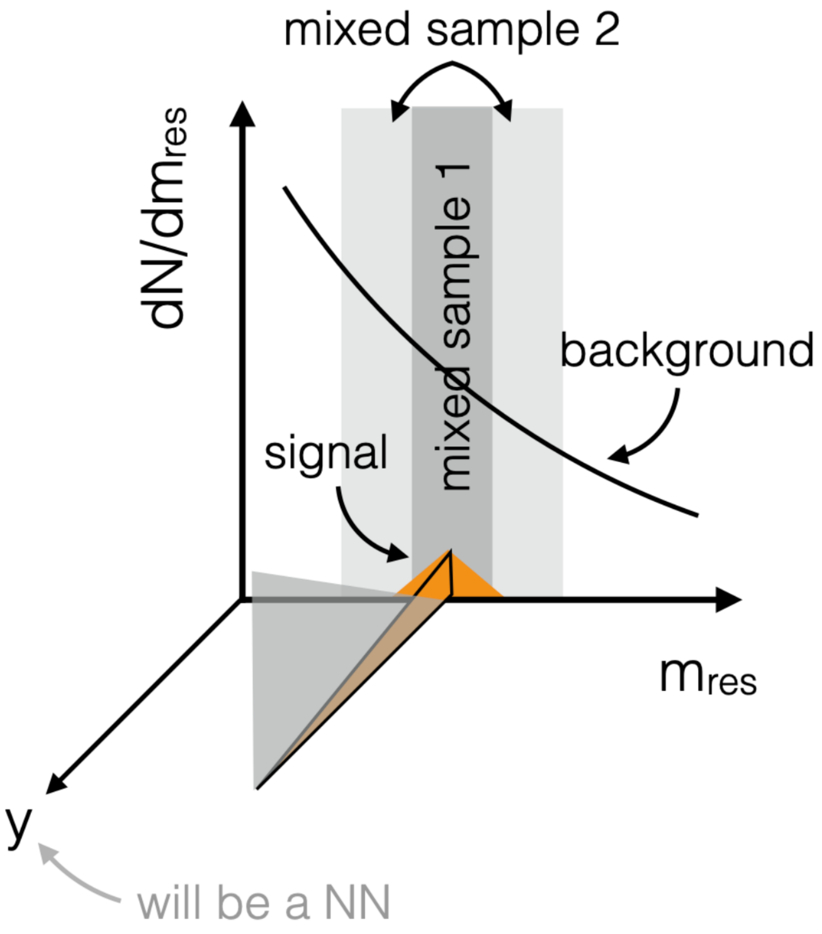
- Dark Matter
- Higgs mass (naturalness problem)
- Matter-Antimatter asymmetry







New physics searches

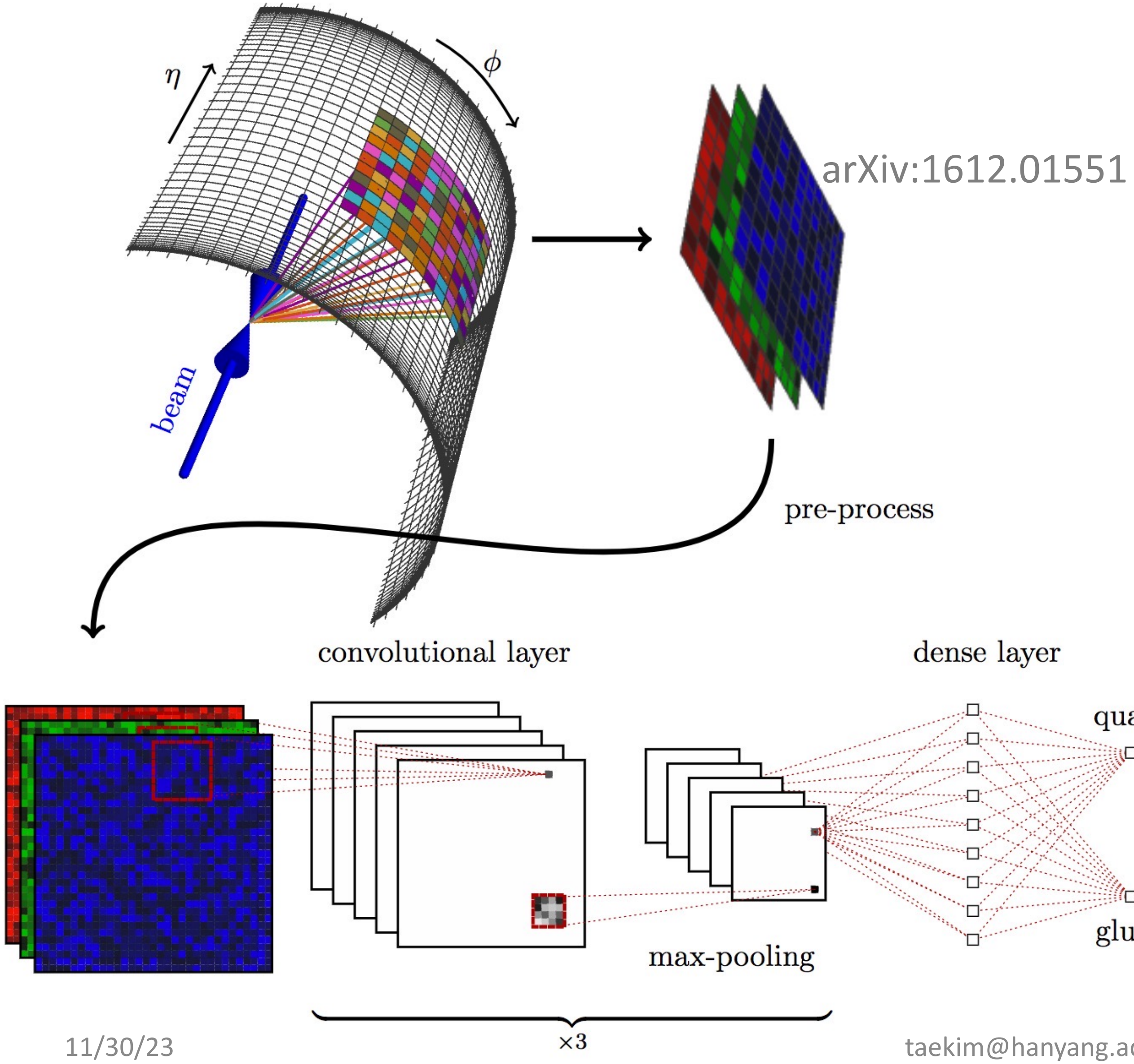


ATLAS article

- Traditional way: Bump hunting
- Combining with technique: Classification without labels (CWoLa)
- Fully data driven machine-learning anomaly detection

taekim@hanyang.ac.kr

gluon and quark jets



Physically motivated observables still necessary?

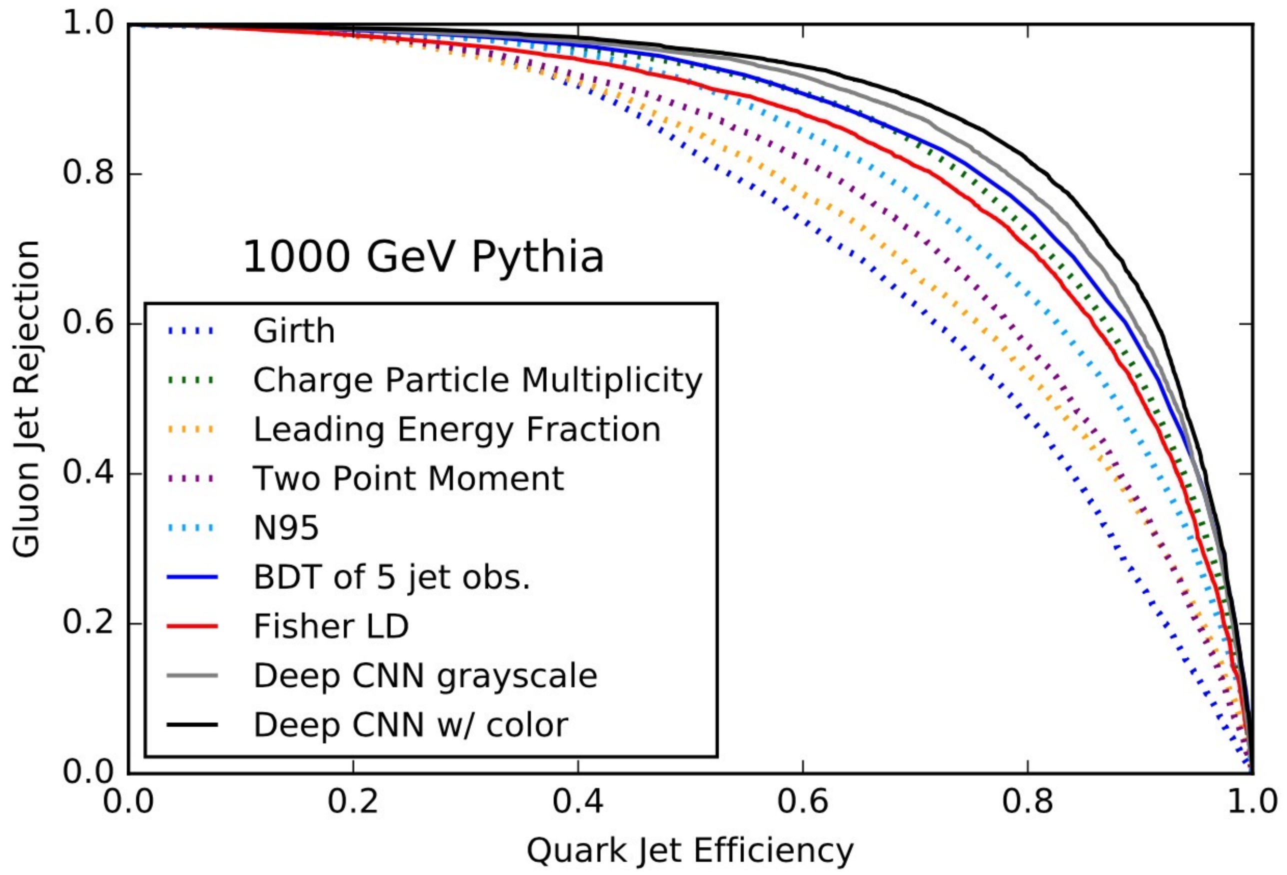
or

Artificial intelligence (deep learning) can find an optimal solution to the quark/gluon discrimination problem?

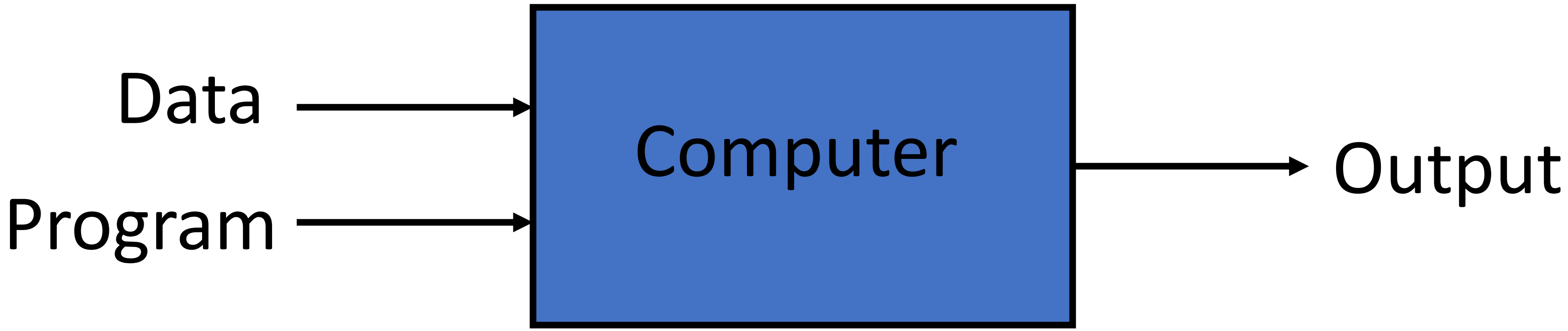
$$qq \to qq,\ q\bar{q} \to q\bar{q} \text{ or } gg \to qq$$

$$gg \to gg \text{ or } q\bar{q} \to gg$$

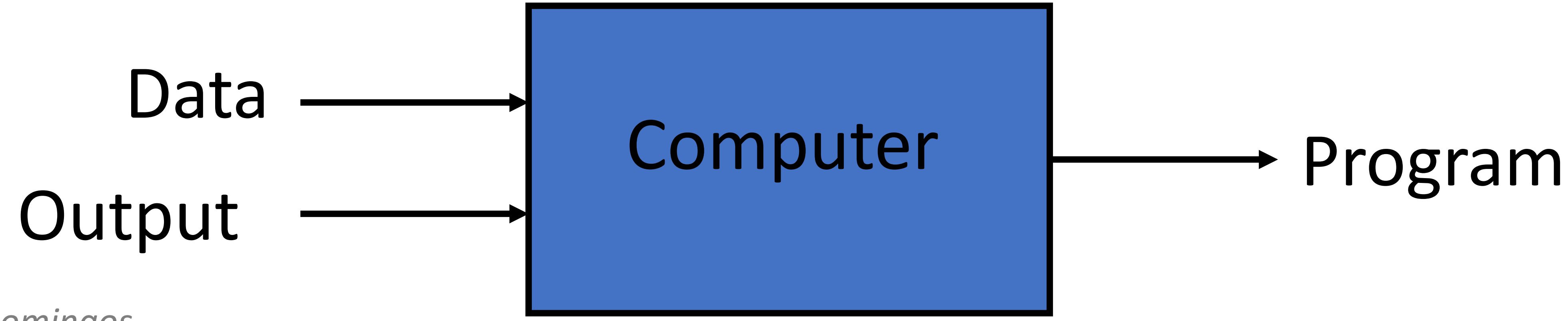
gluon jet vs quark jet discriminator with CNN



Traditional Programming



Machine Learning

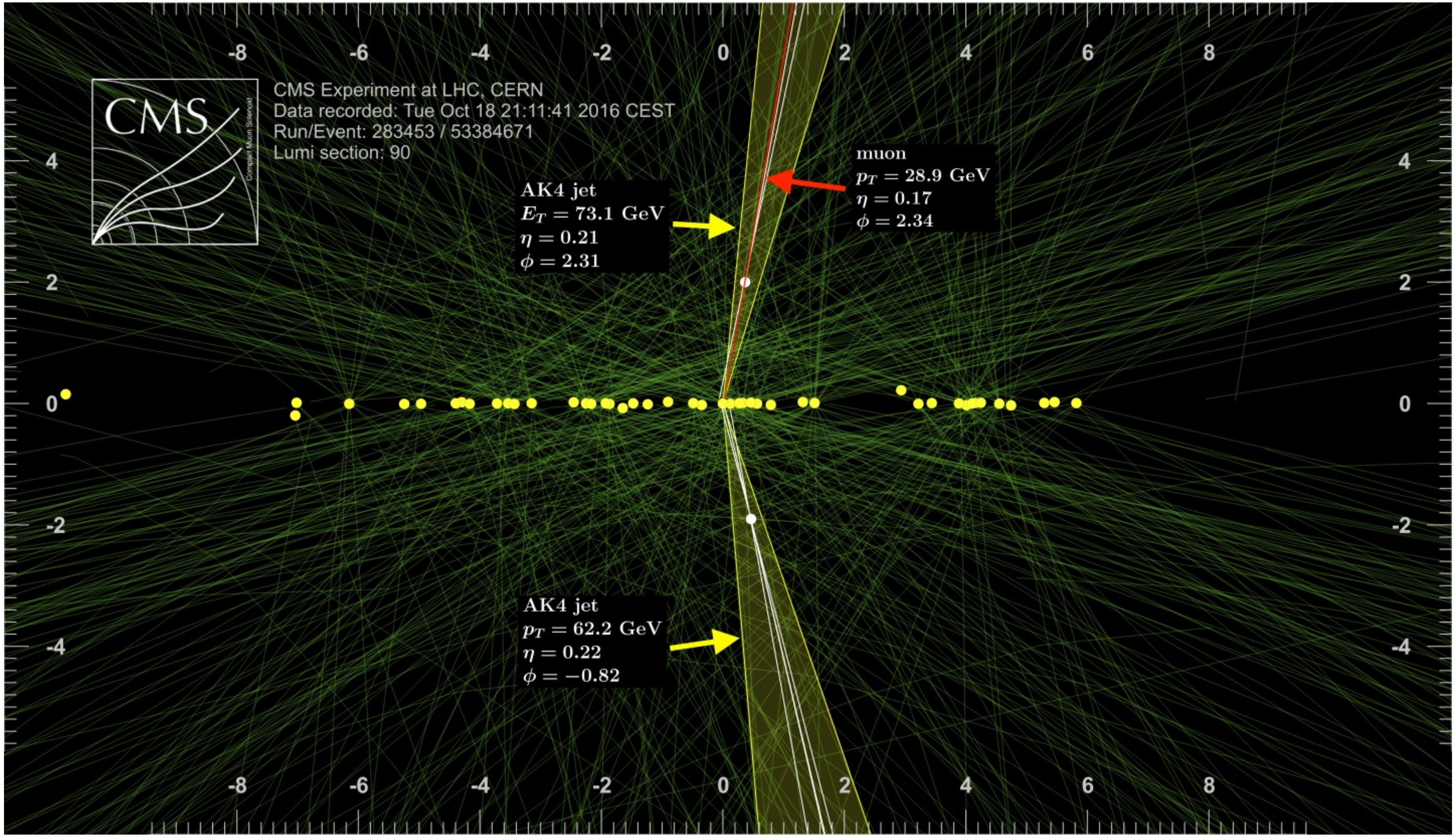


from Pedro Domingos

AI everywhere

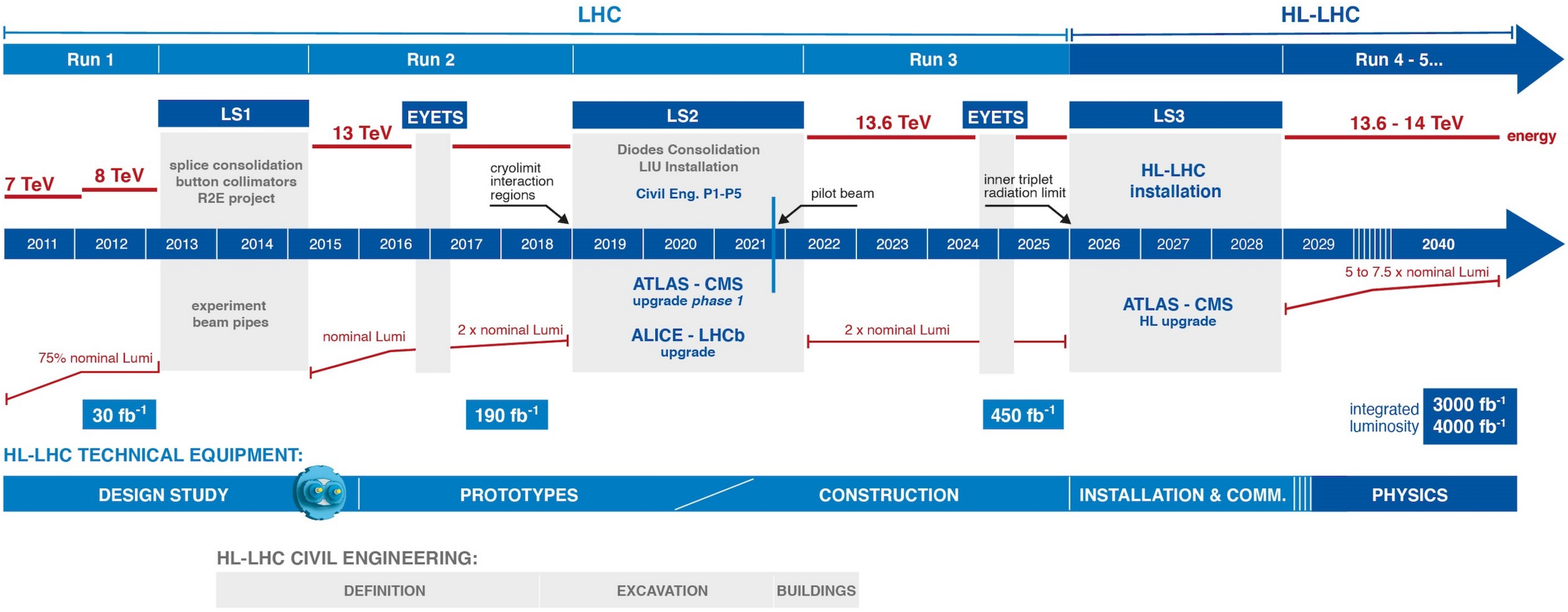
- Trigger
- Object Identification
- Reconstruction
- Precisions measurement
- Search for new physics

Challenges

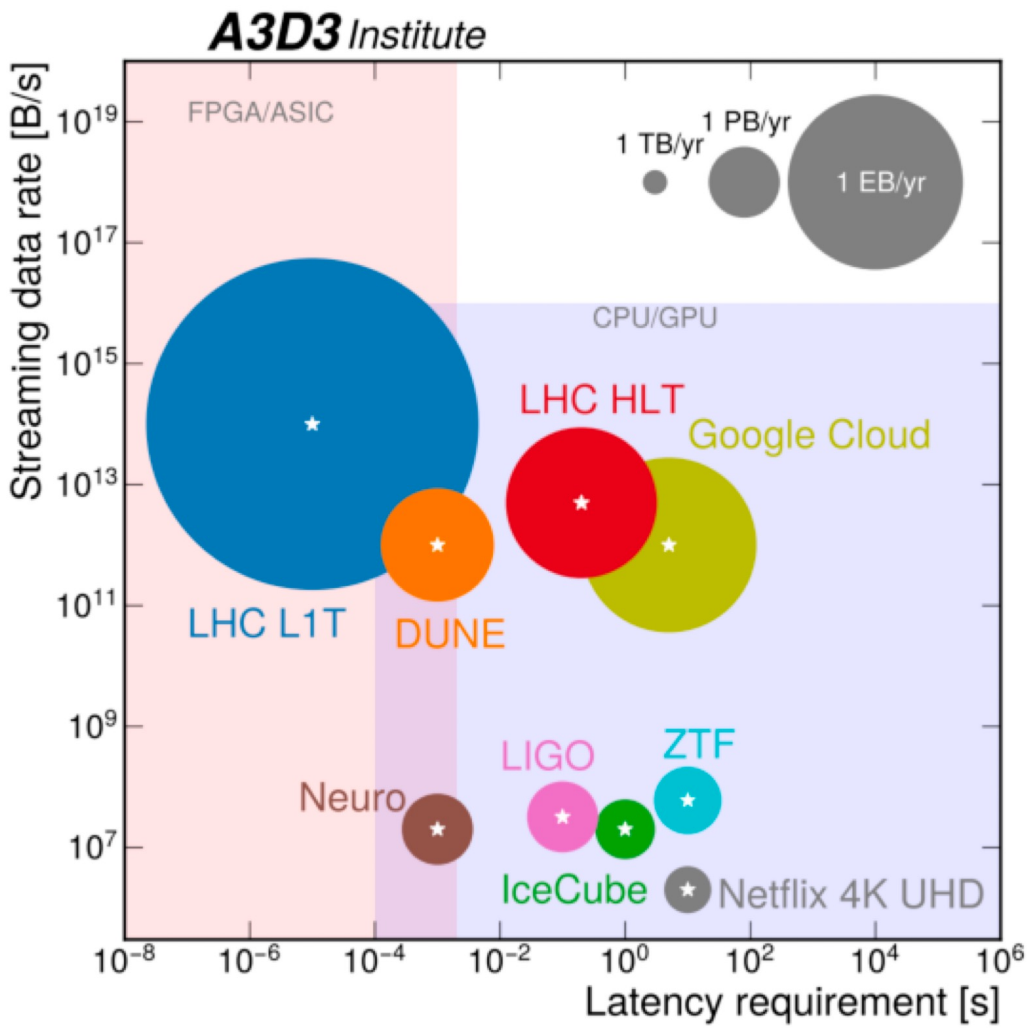


CMS
CMS Experiment at LHC, CERN
Data recorded: Tue Oct 18 21:11:41 2016 CEST
Run/Event: 283453 / 53384671
Lumi section: 90
AK4 jet
$E_T = 73.1$ GeV
$\eta = 0.21$
$\phi = 2.31$
muon
$p_T = 28.9$ GeV
$\eta = 0.17$
$\phi = 2.34$
AK4 jet
$p_T = 62.2$ GeV
$\eta = 0.22$
$\phi = -0.82$

LHC / HL-LHC Plan



Data rate



- Data rate from High-Energy experiments and industry facilities shows wide range of setting
- This gives us hardware challenges

[Snowmass 2021 Computational Frontier Report]

100 TeV Future Circular Collider