CMS Experiment at the LHC, CERN
Data recorded: 2016-Jul-07 12:00:20.388864 GMT
Run / Event / LS: 276495 / 223808853 / 188

Searching for rare Higgs processes with the CMS detector

Irene Dutta (Caltech)

Rising Stars in Experimental Particle Physics Symposium

23 September 2021

The Higgs Boson

- Discovered by both ATLAS and CMS experiments in 2012
- Only scalar particle in the SM, can be represented as a SU(2) doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$$

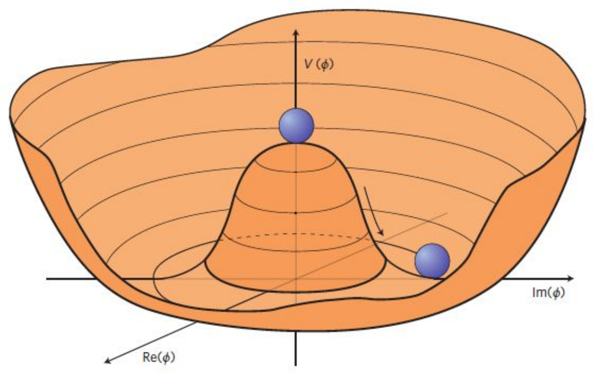
- Higgs field has a mexican hat potential (symmetric under rotations in Φ)

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \qquad \mu^2 < 0 \qquad \lambda > 0$$

- Minimum not at <Φ> = 0

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \qquad v = \frac{|\mu|}{\sqrt{\lambda}}$$

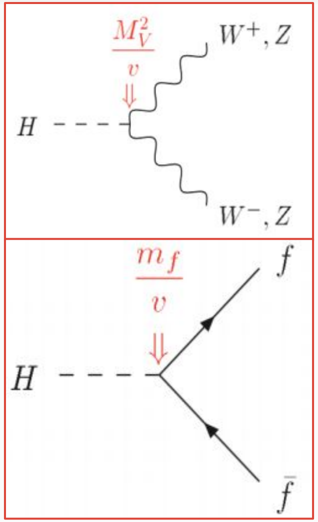
- V.E.V. or v ~246 GeV



Irene Dutta

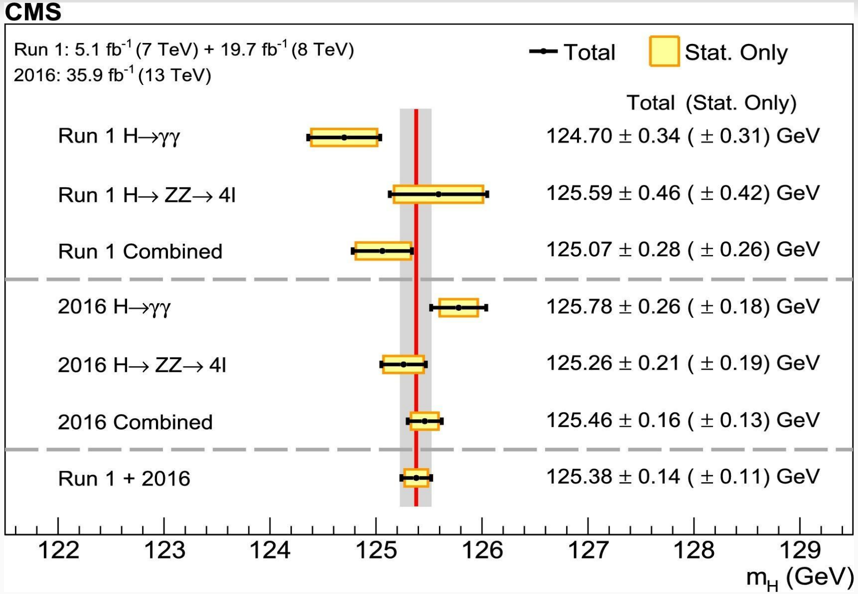
The Higgs Boson

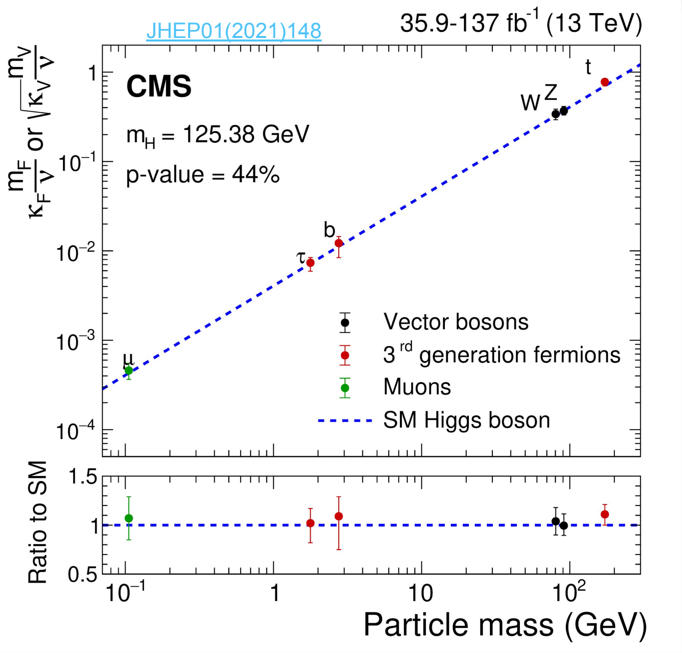
- Fluctuation around v breaks the rotational symmetry - Spontaneous symmetry breaking of the electro-weak vacuum
- Gauge Bosons in the SM acquire mass by electroweak symmetry breaking
- Fermions acquire masses through Yukawa couplings



Credit : PhD Comics

The Higgs Boson properties





Most precise measurement of m_H = 125.38 ± 0.14 GeV
Phys. Lett. B. 805 135425

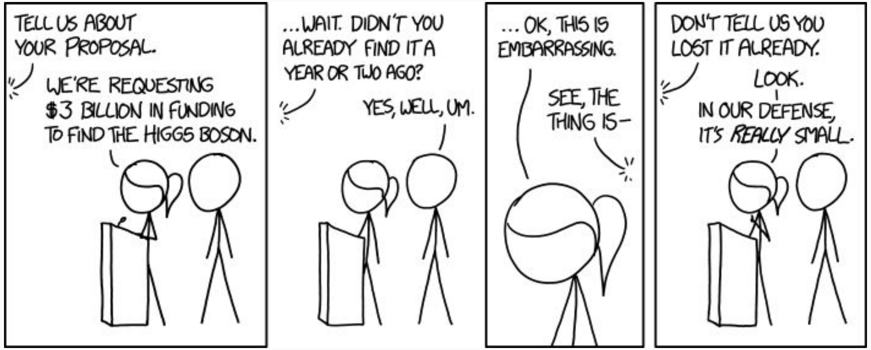
Higgs couplings to the third generation fermions (t/b/τ) and to the SM bosons (W/Z/γ) are well measured

So what’s next ?

Some of the major unanswered questions are

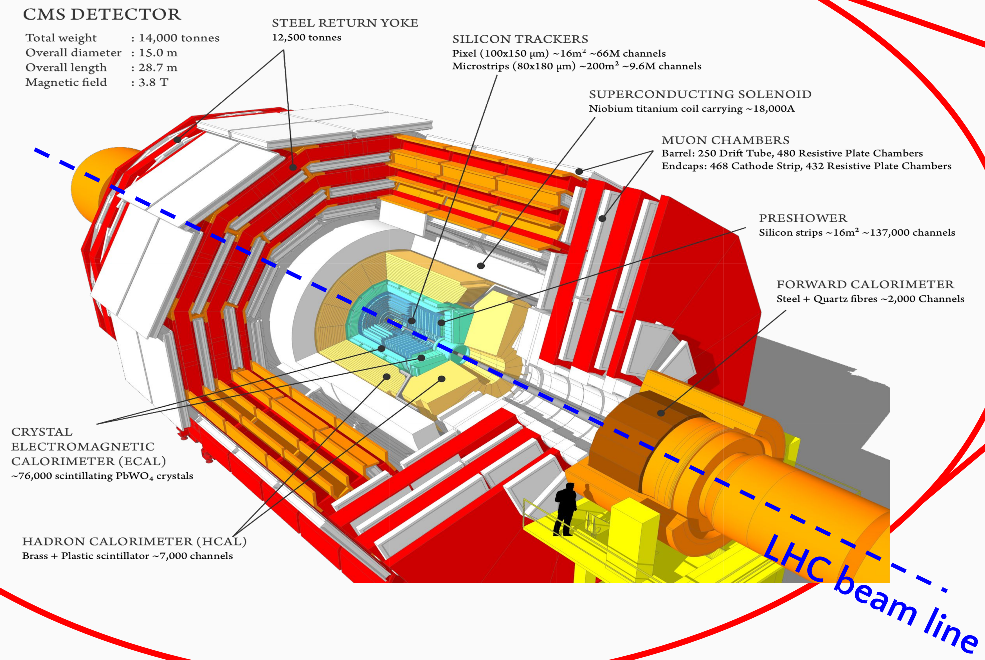
- Yukawa couplings to second generation fermions (muon /charm quark)- any deviation from the SM can indicate the existence of an unknown BSM process!
- Higgs self-coupling - crucial in understanding the shape of the scalar potential at higher scales

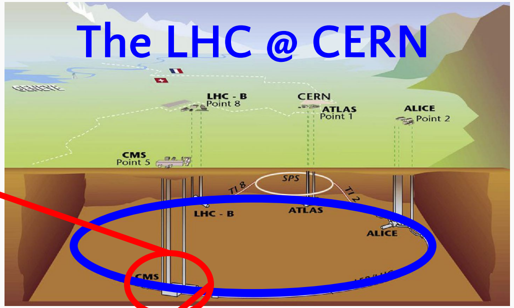
We will discuss both these topics in details in the next few slides

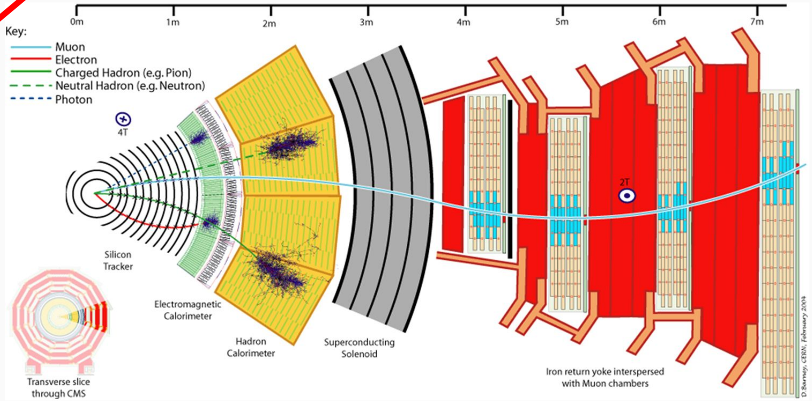


Credit : xkcd comics

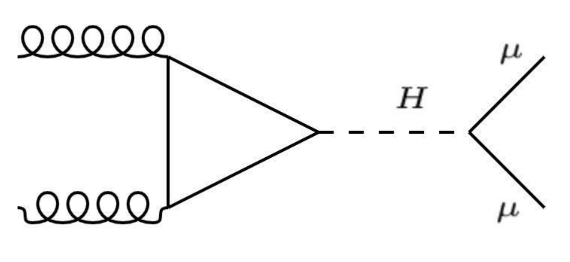
The Compact Muon Solenoid (CMS)







Higgs couplings to 2nd generation fermions

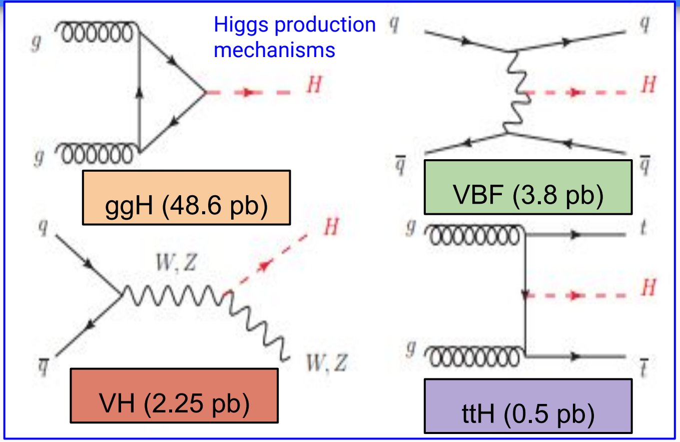


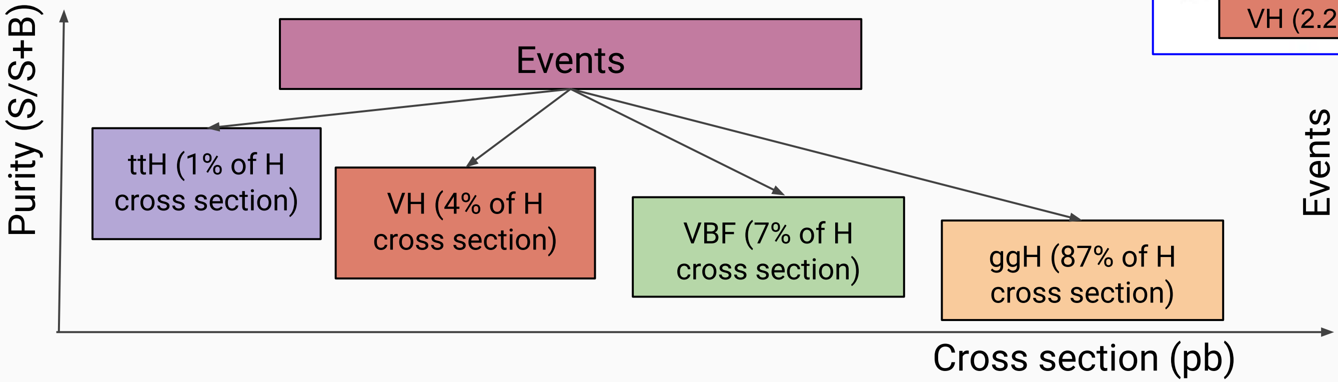
- H→cc has the largest BR, but also more background contamination
- H→μμ is currently the cleanest probe for second generation Yukawa coupling at the LHC
    - BR(H→μμ) ~2.15×10^{-4} for M_H = 125.38 GeV
    - Mass peak resolution : 1.5~2.5 GeV
- Large background (dominated by Drell-Yan Z→μμ, electro-weak Z, others include top, diboson and triboson production) - S/B ~1/500 (very hard to find)!
- Most recent result from CMS: JHEP 01(2021)148
    - 3 σ excess observed in data @ M_H = 125.38 GeV!!

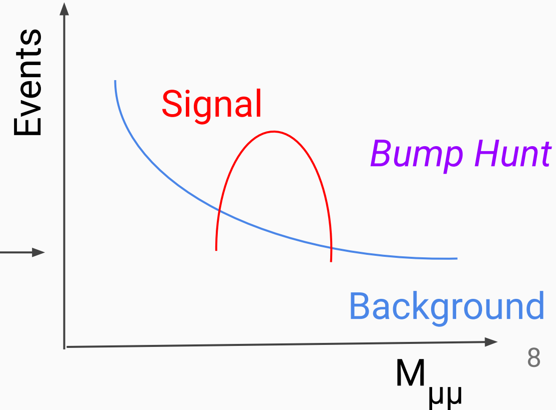
Irene Dutta

Search strategy

- Two oppositely charged muons that are well isolated and have the largest sum p_T.
- Higgs-candidate $M_{\mu\mu} \in [110, 150]$ GeV
- Exploit different kinematic features of different production mechanisms







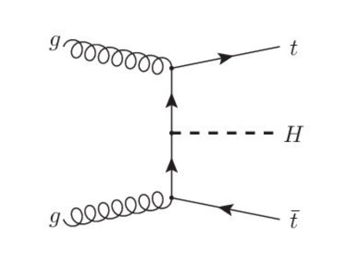
Irene Dutta

Event separation

1)

ttH

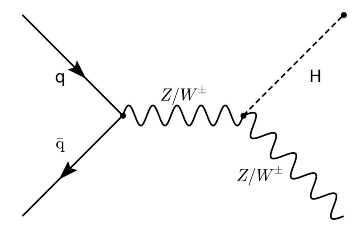
- Look for additional b-jets from top quark decays



2)

VH (leptonic)

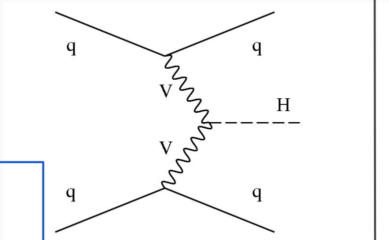
- Look for additional leptons from the W/Z decay



3)

Vector Boson Fusion (VBF)

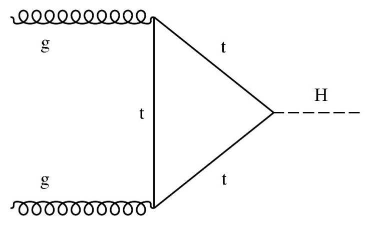
- Look for two high momentum jets with large angular separation and high jet-jet invariant mass



4)

ggH

- Accept all other remaining events after vetoing ttH, VH and VBF events



All categories are exclusive to each other

Fit strategy

ttH | VH (leptonic) | ggH

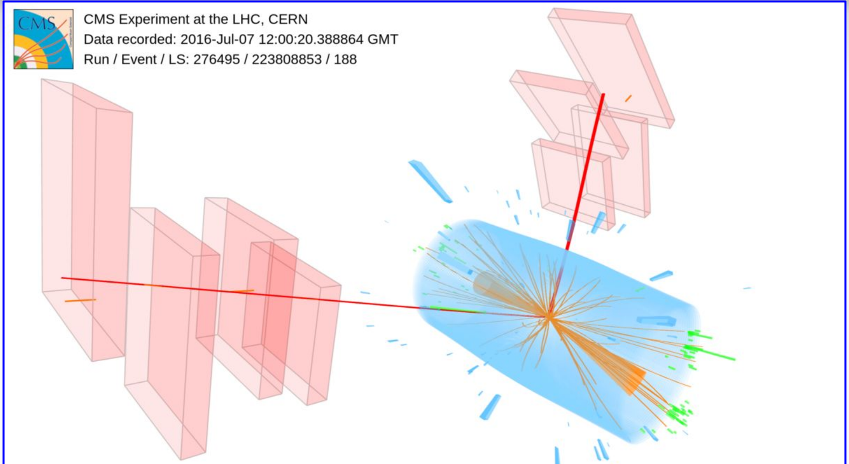
- Train independent BDTs for each region
- Background modelled with discrete likelihood profile of physics inspired and empirical functions.
- Signal peak modelled with a Gaussian function with power-law tails on both sides
- Perform a parametric fit to $M_{\mu\mu}$ spectrum

VBF

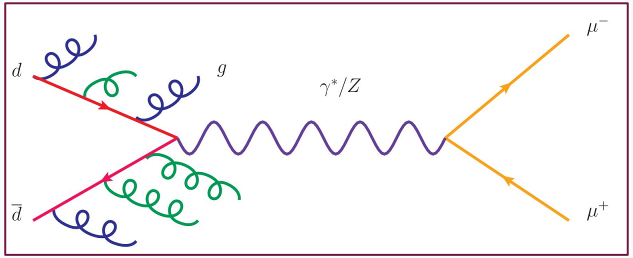
- **Most sensitive category** of this analysis
- Train a deep neural network
- Perform a simulation based template fit to DNN score output
- Personal contribution here: more on this in the next few slides

Template based VBF channel

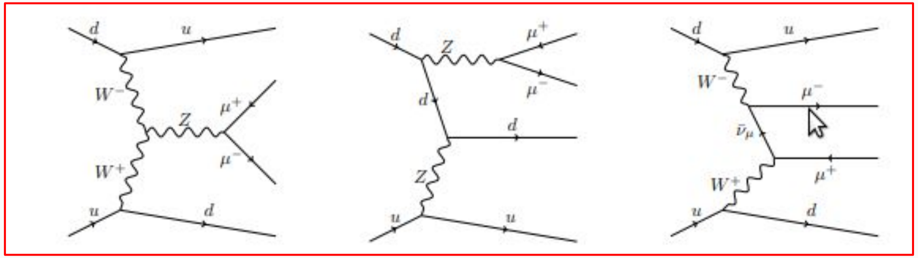
- VBF H→μμ events have distinct signature:
  - 125 GeV peak in dimuon invariant mass
  - Two forward high-p_T jets with high invariant mass
- Major backgrounds are Drell-Yan Z and electroweak production of Z



Candidate VBF H→μμ event



Drell Yan: Z/γ*→ μμ



Electroweak production of Z+jj

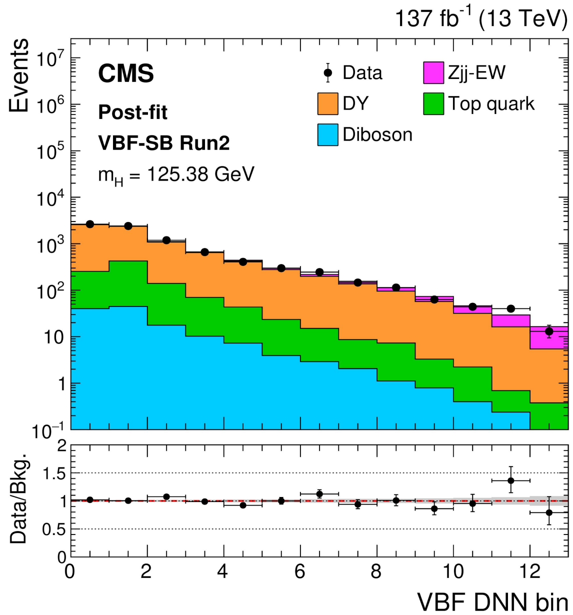
VBF DNN

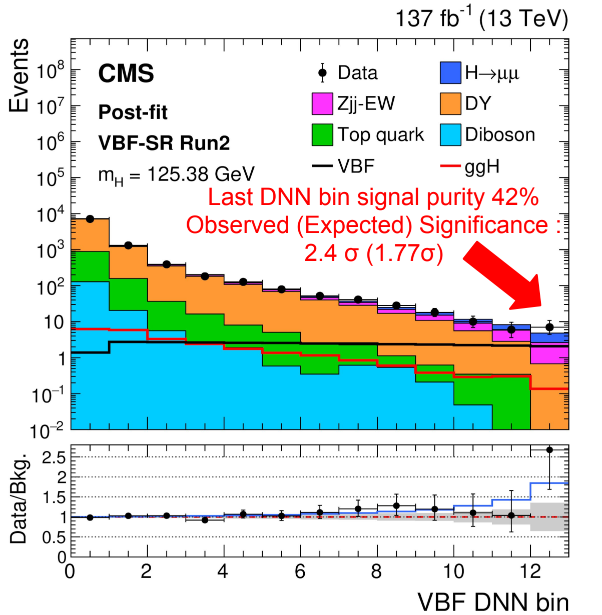
- Train a supervised machine learning classifier (a.k.a deep neural network or DNN)
- Signal and background predictions obtained from MC simulation
- Perform a MC template-based fit to the output score; Performance depends on
    - statistical power of background MC in signal region
    - data/MC agreement
- 20% improvement w.r.t. a data-driven fit approach

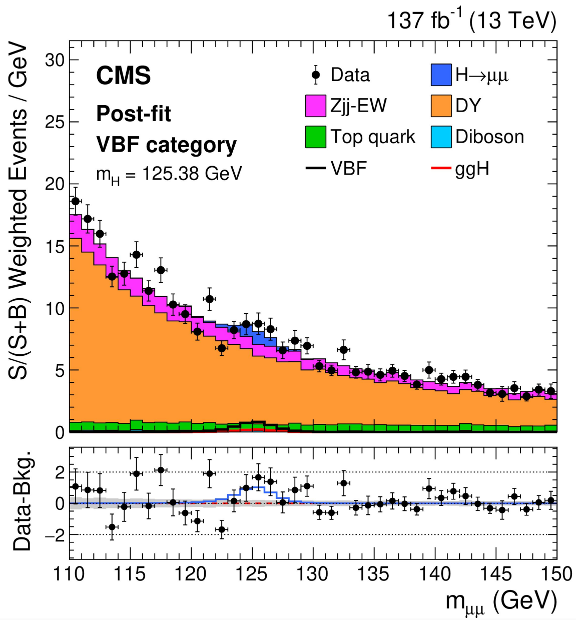
Input features of the DNN :

- Di-muon kinematics
- Di-muon mass and mass resolution
- Di-jet kinematics
- Individual jet kinematics
- Di-muon+Di-jet system kinematics

VBF fitting strategy





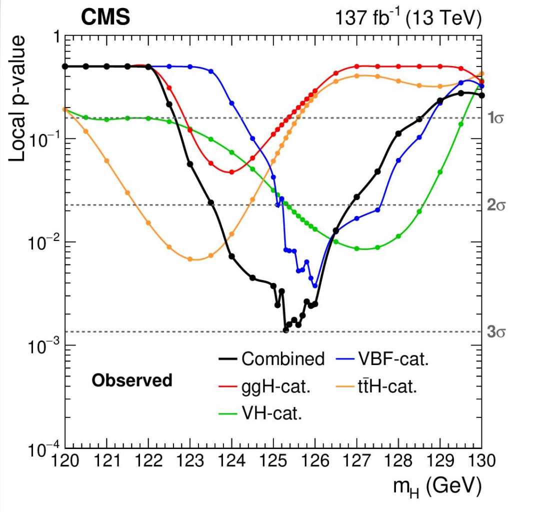


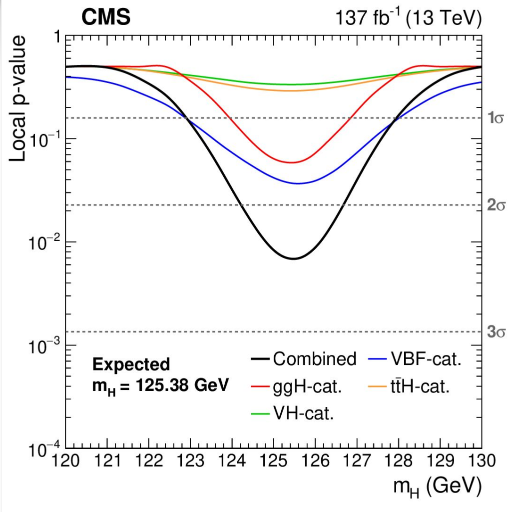
Two analysis regions based on $M_{\mu\mu}$

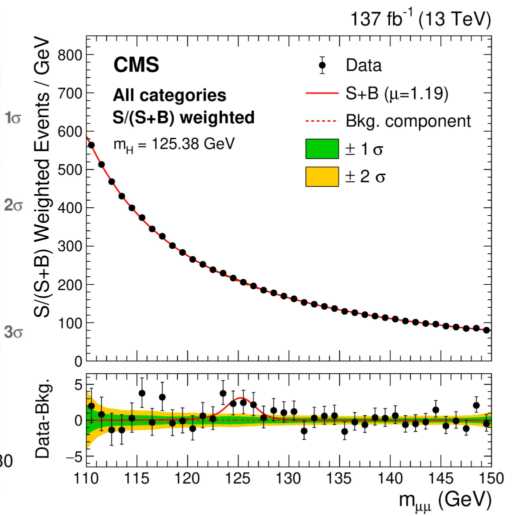
- Higgs peak: $M_{\mu\mu} \in [115,135]$ GeV
- Higgs sideband: $M_{\mu\mu} \in [110, 115]$ GeV ∪ [135, 150] GeV

Mass spectrum for events weighted according to S/(S+B) in a particular DNN bin

Evidence of H→μμ





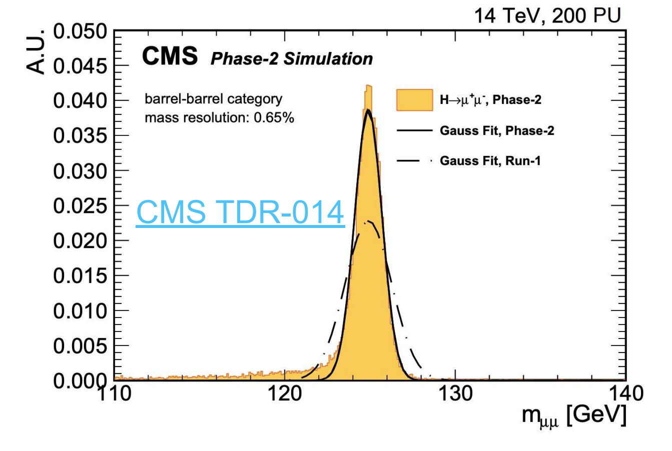


Production category	Observed (expected) Signif.
VBF	2.40 (1.77)
ggH	0.99 (1.56)
$t\bar{t}H$	1.20 (0.54)
VH	2.02 (0.42)
Combined $\sqrt{s} = 13$ TeV	2.95 (2.46)
Combined $\sqrt{s} = 7, 8, 13$ TeV	2.98 (2.48)

Uncertainty source	$\Delta\mu$	
Total uncertainty	+0.44	−0.42
Statistical uncertainty	+0.41	−0.39
Total systematic uncertainty	+0.17	−0.16
Size of simulated samples	+0.07	−0.06
Total experimental uncertainty	+0.12	−0.10
Total theoretical uncertainty	+0.10	−0.11

HL-LHC projections for H→μμ

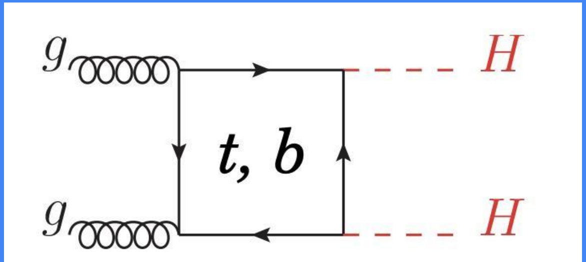
- The HL-LHC will start in 2027 delivering about 3 ab^{-1} of pp collision data at 14 TeV
  - Extreme pileup conditions → 200 concurrent interactions every bunch crossing
- Several detector upgrades planned -
  - New tracker with coverage up to |η|=4 and L1 track trigger - $\sigma(M_{\mu\mu})$ improves by a factor of 2!
  - Upgraded muon system with coverage up to |η|=2.8
  - New high granularity endcap calorimeter (HGCal)
- Precision H→μμ measurement! - overall uncertainty constrained to ~4% (arXiV:1902.00134)



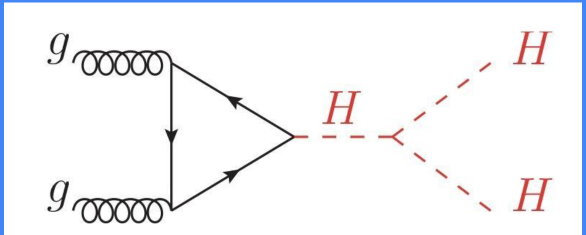
Improvements to H→μμ

- Increased acceptance of muons and improvement in muon p_T resolution
- VBF category: improved jet-energy-resolution and rejection of pileup jets in the endcap and forward region

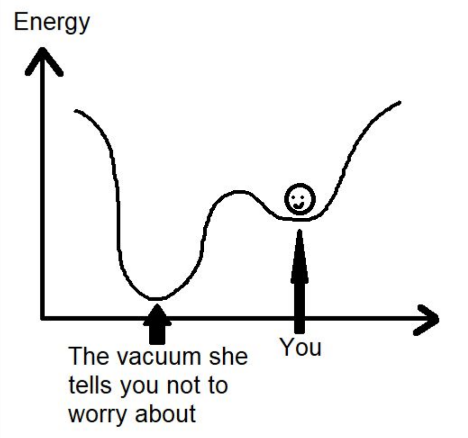
Higgs self coupling



+



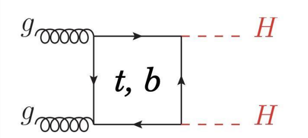
- $V(H) = m_h^2 h^2/2 + m_h^2 h^3/2v + m_h^2 h^4/2v^2$
- Higgs trilinear self coupling is $\lambda = m_h^2/2v$
- Important to study the trilinear coupling
  - probe the structure of the Higgs potential at large scales - metastability of the EW vacuum
  - EW phase transition in the primordial universe might be responsible for origin of matter anti-matter asymmetry

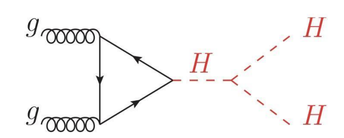


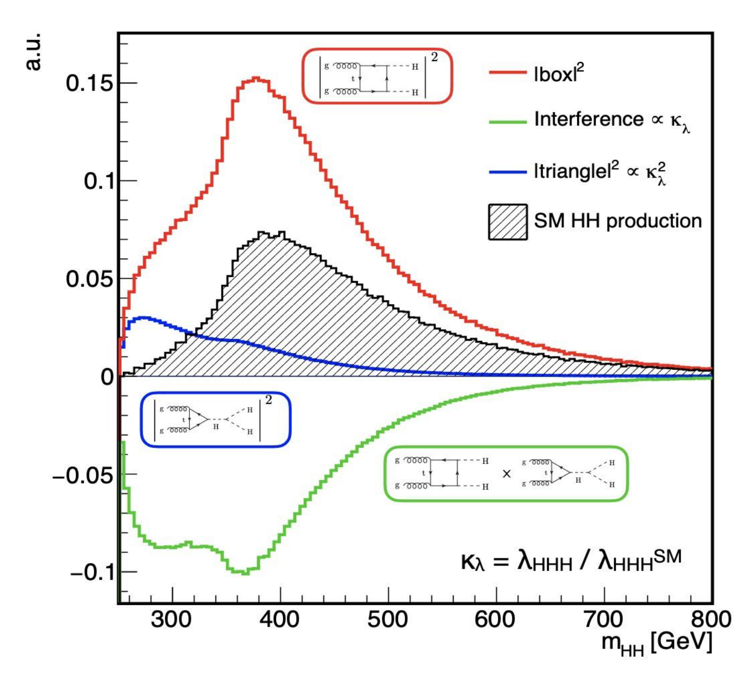
Credit: r/physicsmemes

Higgs self coupling

- Two main leading order di-Higgs production diagrams:
  - Destructive interference
  - Smaller overall cross section
- During Run 2, the LHC produced :
  - 7.5 million single Higgs Bosons
  - 4500 Higgs Boson pairs







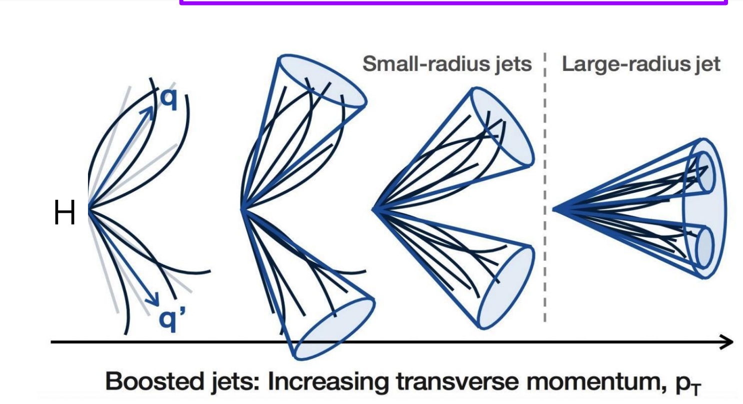
Irene Dutta

Lorentz boosted ggF HH→(bb)(bb)

- HH production cross section through ggF ~31 fb at 13 TeV at NNLO
- HH→(bb)(bb) has the highest branching ratio i.e. 33.9 % of HH decay
- However very largely dominated by QCD multi-jet background and poor decay channel resolution
- Most studies target HH→(bb)(γγ) for better resolution and higher S/B (0.26 % of HH decay)
- To reduce background for HH→(bb)(bb) , explore the regime where both Higgs are boosted
    - Exploit jet sub-structures for better S/B

HH branching ratios

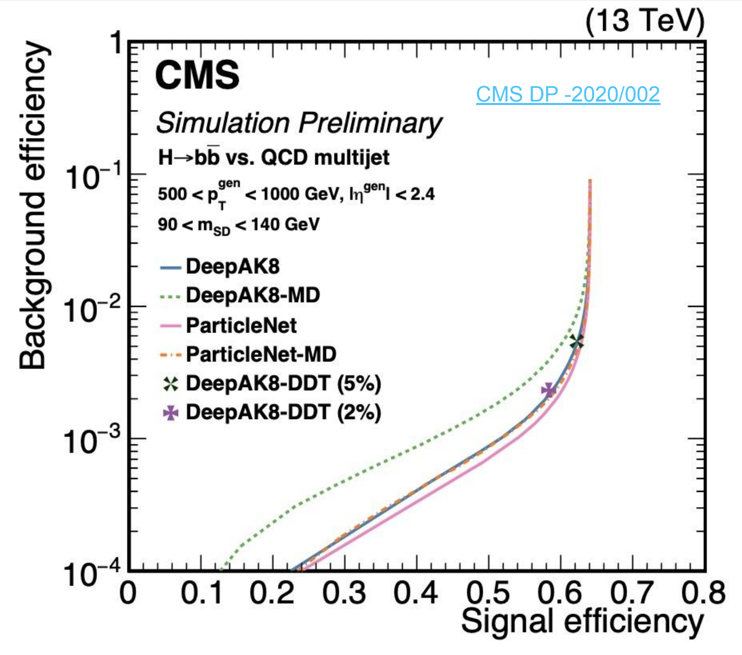
	bb	WW	ττ	ZZ	γγ
bb	33.9%				
WW	24.9%	4.6%			
ττ	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.07%	
γγ	0.26%	0.10%	0.03%	0.01%	



ParticleNet jet tagger for large radius jets

ParticleNet: A multi-class jet classifier for top, Higgs, W or Z tagging for large radius jets

- low-level jet information (collections of particles inside jet, secondary vertices from b-quark decays etc.) as inputs
- Dynamic Graph Convolutional Neural Networks (DGCNN) as ML architecture, details in Phys. Rev. D 101, 056019



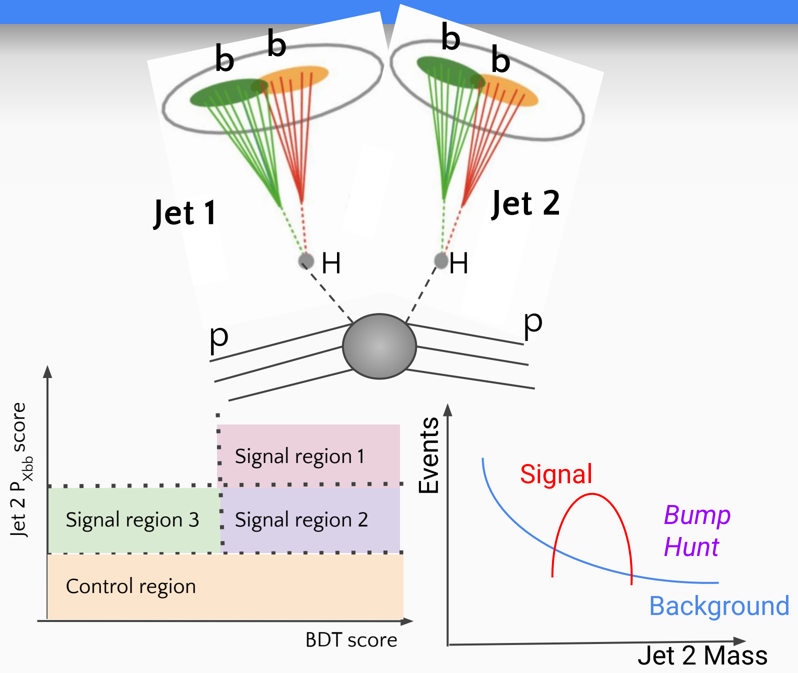
ParticleNet-MD: The mass-decorrelated version

- Agnostic to the jet mass
- output scores: X→bb, X→cc, X→qq, QCD

Irene Dutta

Analysis Strategy

- Major backgrounds : QCD and Top
- Design a Boosted Decision Tree (BDT): discriminate HH signal events from QCD and top quark background events. Inputs include:
  - Jet 1 kinematics
  - Dijet kinematics
- Define event categories optimized based on a 2D grid of BDT score and Jet 2 P_{Xbb} score
- Use control regions (very little signal) to estimate background shapes in the final fit regions
- Fit the Mass of Jet 2 in all categories



Analysis is still under internal review – stay tuned for public results!

Future of HH

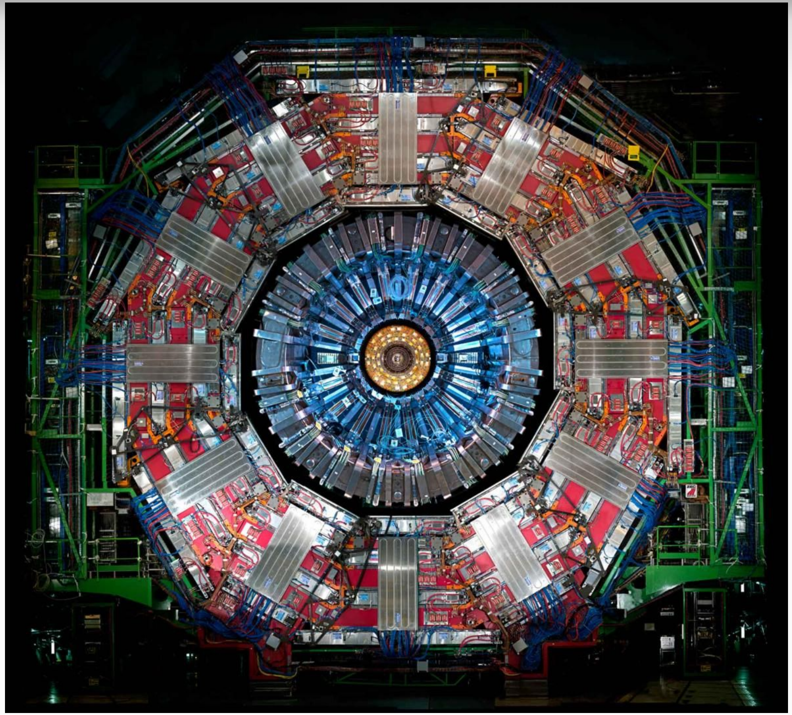
- Current existing best limit for HH : 3.6 (7.3) X SM Observed (Expected) in resolved HH→4b CMS analysis : CMS-PAS-HIG-20-005
    - The Boosted ggHH→4b results are competitive with these limits
    - Stay tuned for public results by end of 2021 /early 2022.

HL-LHC prospects for HH production

- Projected combined ATLAS +CMS expected significance at ~4σ (CERN-LPCC-2018-04)
- Can see further improvements from new analysis methods/ constraining systematics

Summary

- The Higgs Boson is a rather elusive particle
- Several of its properties are still not well understood
- Current and future colliders will help answer several open questions about the role of the Higgs in our universe



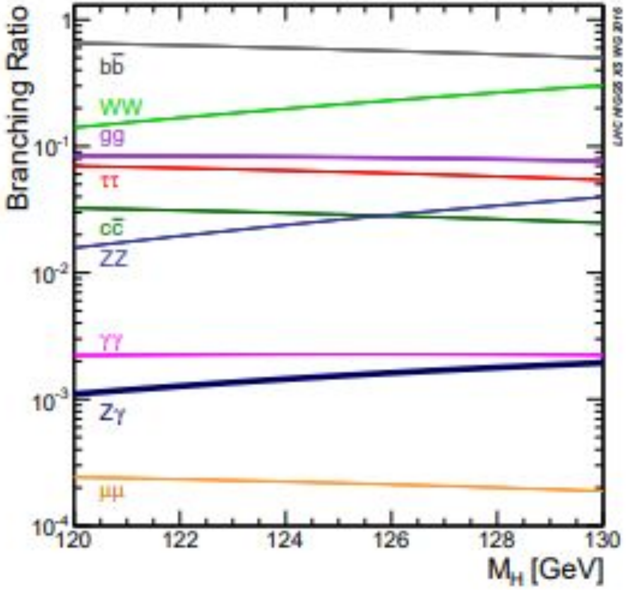
A giant eye for tiny particles

Backup

HIGGS DECAY CHANNELS

Table 11.4: The five principal decay channels for low mass SM Higgs boson searches at the LHC. The numbers reported are for $m_H = 125\,\text{GeV}$.

Decay channel	Mass resolution
$H \to \gamma\gamma$	1–2%
$H \to ZZ \to \ell^+\ell^-\ell'^+\ell'^-$	1–2%
$H \to W^+W^- \to \ell^+\nu_\ell\ell'^-\bar{\nu}_{\ell'}$	20%
$H \to b\bar{b}$	10%
$H \to \tau^+\tau^-$	15%



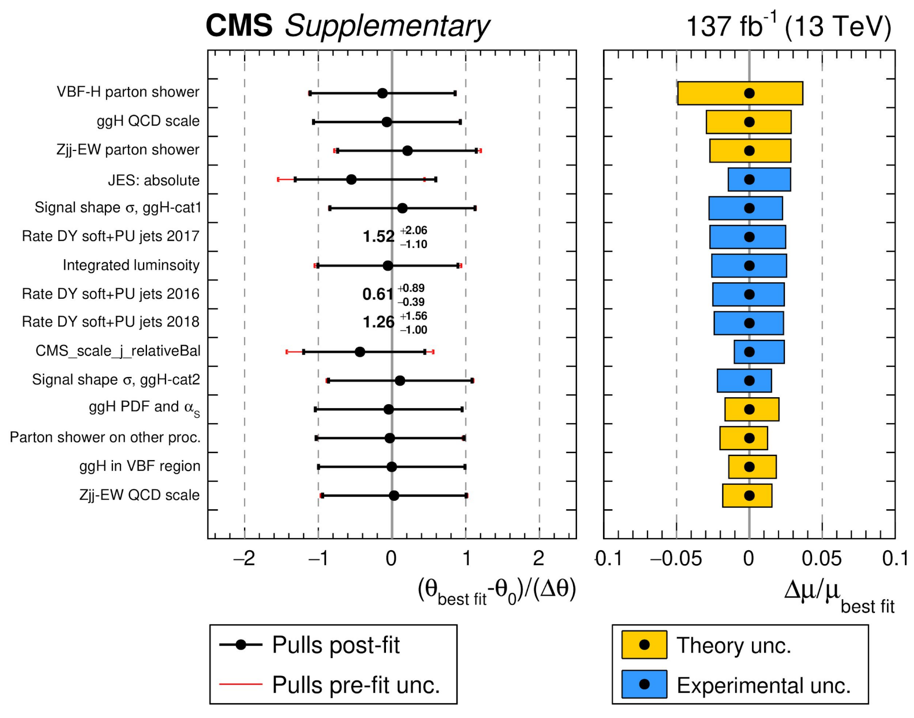
Irene Dutta

Systematic uncertainties on Run 2 combination

Largest systematic uncertainty impacts from

- limited statistics in data
- the signal and background theory modeling
- Main experimental uncertainties include jet energy scale and resolution uncertainties

Uncertainty source	$\Delta\mu$	
Total uncertainty	+0.44	−0.42
Statistical uncertainty	+0.41	−0.39
Total systematic uncertainty	+0.17	−0.16
Size of simulated samples	+0.07	−0.06
Total experimental uncertainty	+0.12	−0.10
Total theoretical uncertainty	+0.10	−0.11



Improvements in sensitivity w.r.t. 2016 CMS result

2016 CMS result: Phys.Rev.Lett.122,021801

- Targeting ggH and VBF production - single BDT.
- Classic bump hunt in $M_{\mu\mu}$ with data-driven fits
- Expected sensitivity with 2016 data: 0.9σ

Run 2 CMS result

- Embedding $M_{\mu\mu}$ resolution in BDT training to improve measurement
- New channels targeting VH and ttH.
- New exclusive VBF channel with redesigned simulation based strategy.
- More robust background modeling
- Specific muon p_t corrections: final state radiation and muon-track impact parameter w.r.t beamspot

Run 2 result improves upon the extrapolation of the 2016 only result (by luminosity) by about 35%

The HL-LHC

- Starting in 2027
- Increase beam intensity
  - Baseline: $L = 5.0 \times 10^{34}$ $cm^{-2}s^{-1}$ (140 PU)
  - Ultimate: $L = 7.5 \times 10^{34}$ $cm^{-2}s^{-1}$ (200 PU)
- Target is 3000 - 4000 fb^{-1} data
- More p-p collisions per bunch crossing
- Experiments will need to upgrade to maintain current level of particle reconstruction efficiencies.

HL-LHC starts running

Currently almost at the end of LS2

Install detector upgrades