

Observation of Higgs Boson Decay to Bottom Quarks When Produced in Association with a Weak Vector Boson

Sean-Jiun Wang
March 12, 2019

Chair: Jacobo Konigsberg
Co-Chair: Ivan Furic

Outline

- Introduction
- Theory Background
- The CMS Detector
- Event Reconstruction
- Motivation for VH($b\bar{b}$) Searches
- 2017 VH($b\bar{b}$) Analysis
- Results
- Conclusions

Introduction

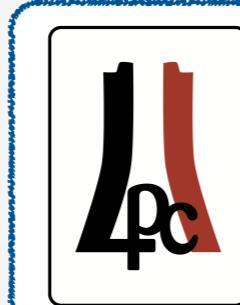
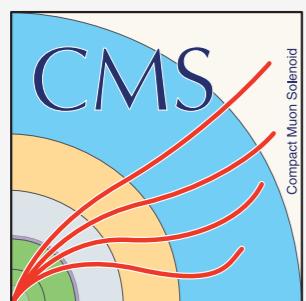
Introduction

Overview



2009

2019



- 2 Years at FNAL
- 2016 LPC Graduate Scholar
- Science Advocacy in D.C.
- Finish Ph.D. Research

2011

2018

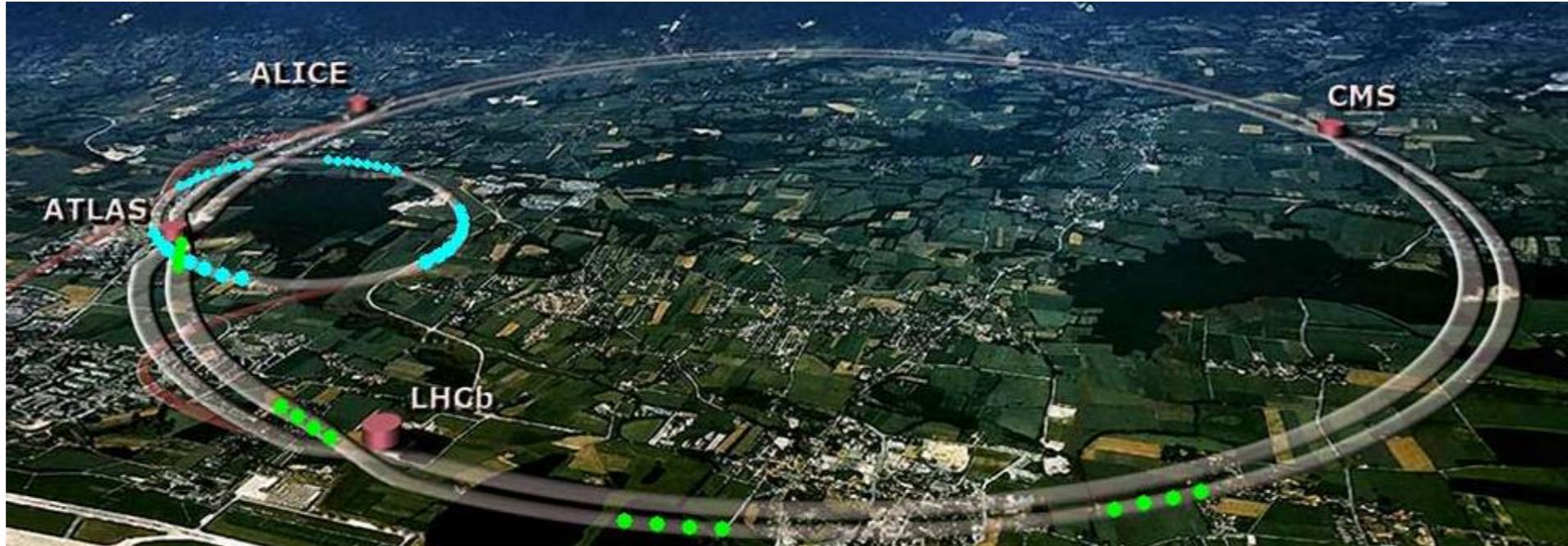
Undergraduate Research



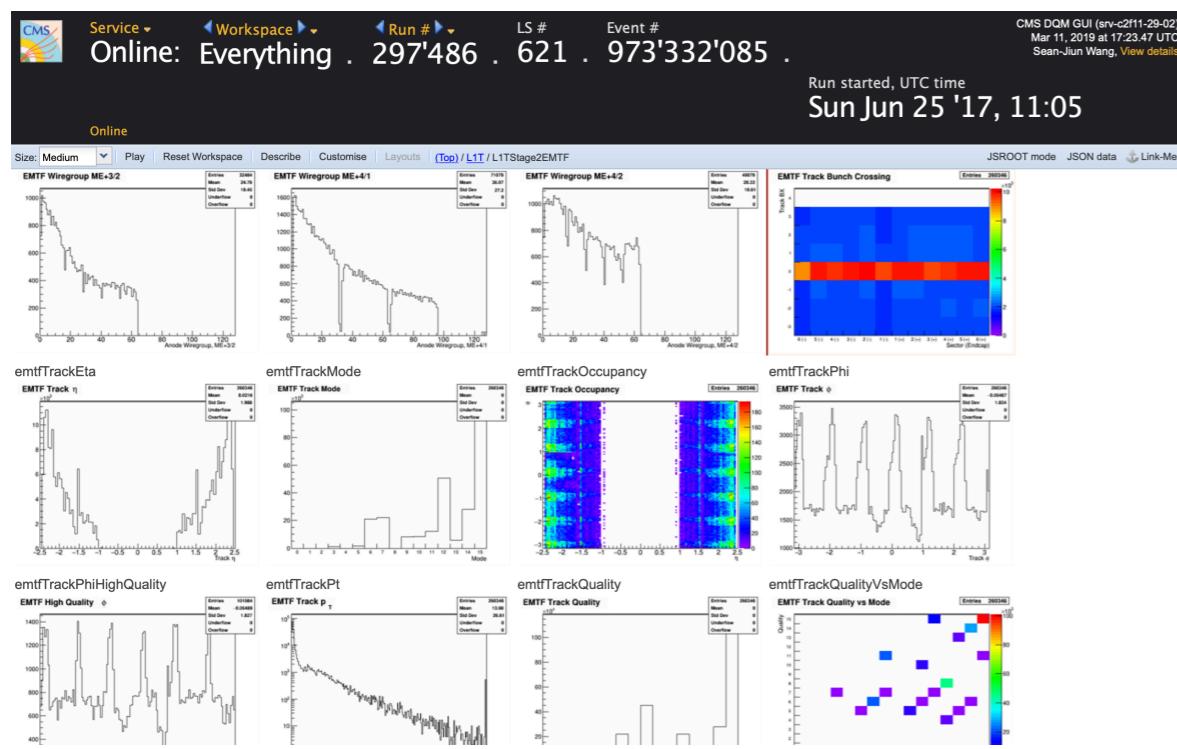
- 1 Year at CERN
- Begin Ph.D. Research
- Data Quality Monitoring

Introduction

Time at CERN



Physics Letters B 780 (2018) 501–532



Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb



Evidence for the Higgs boson decay to a bottom quark–antiquark pair

The CMS Collaboration*

CERN, Switzerland

ARTICLE INFO

Article history:

Received 21 September 2017
Received in revised form 17 February 2018
Accepted 20 February 2018
Available online 27 February 2018
Editor: M. Doser

Keywords:
CMS
Physics
Higgs

ABSTRACT

A search for the standard model (SM) Higgs boson (H) decaying to $b\bar{b}$ when produced in association with an electroweak vector boson is reported for the following processes: $Z(\nu\nu)H$, $W(\mu\nu)H$, $W(e\nu)H$, $Z(\mu\mu)H$, and $Z(ee)H$. The search is performed in data samples corresponding to an integrated luminosity of 35.9 fb^{-1} at $\sqrt{s} = 13 \text{ TeV}$ recorded by the CMS experiment at the LHC during Run 2 in 2016. An excess of events is observed in data compared to the expectation in the absence of a $H \rightarrow b\bar{b}$ signal. The significance of this excess is 3.3 standard deviations, where the expectation from SM Higgs boson production is 2.8. The signal strength corresponding to this excess, relative to that of the SM Higgs boson production, is 1.2 ± 0.4 . When combined with the Run 1 measurement of the same processes, the signal significance is 3.8 standard deviations with 3.8 expected. The corresponding signal strength, relative to that of the SM Higgs boson, is $1.06^{+0.31}_{-0.29}$.

© 2018 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>). Funded by SCOAP³.

Introduction

Time at FNAL



121

PRL 121 (12), 120401–129902, 21 September 2018 (332 total pages)

12

Published by
American Physical Society

**PHYSICAL
REVIEW
LETTERS**

Articles published week ending 21 SEPTEMBER 2018

The journal cover for Physical Review Letters (PRL) Volume 121, Number 12, dated 21 September 2018. The cover features a green and white geometric design on the left. The title "PHYSICAL REVIEW LETTERS" is prominently displayed in the center in a serif font. Below the title, it says "Articles published week ending 21 SEPTEMBER 2018". On the right side, there is a circular plot with radial lines and data points, overlaid with a red magnifying glass icon pointing towards the center.

Volume 121, Number 12

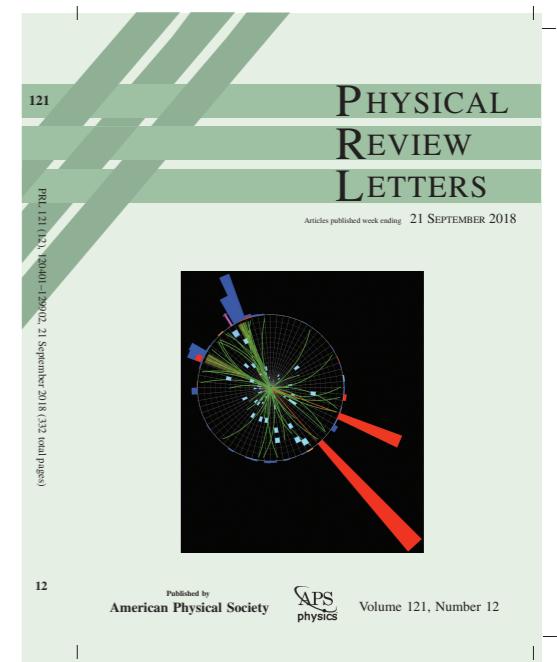
6

Introduction

PhD Research

- Joined the VH($b\bar{b}$) analysis group during transition to Run 2.
- Lead analyst for the 0-lepton channel.
- Main Contributions
 - Redesign of 0-lepton analysis
 - Analysis software development and open source packages
 - Fast data preprocessing
 - Automated hyperparameter tuning
 - Investigate novel approach to invariant mass analyses

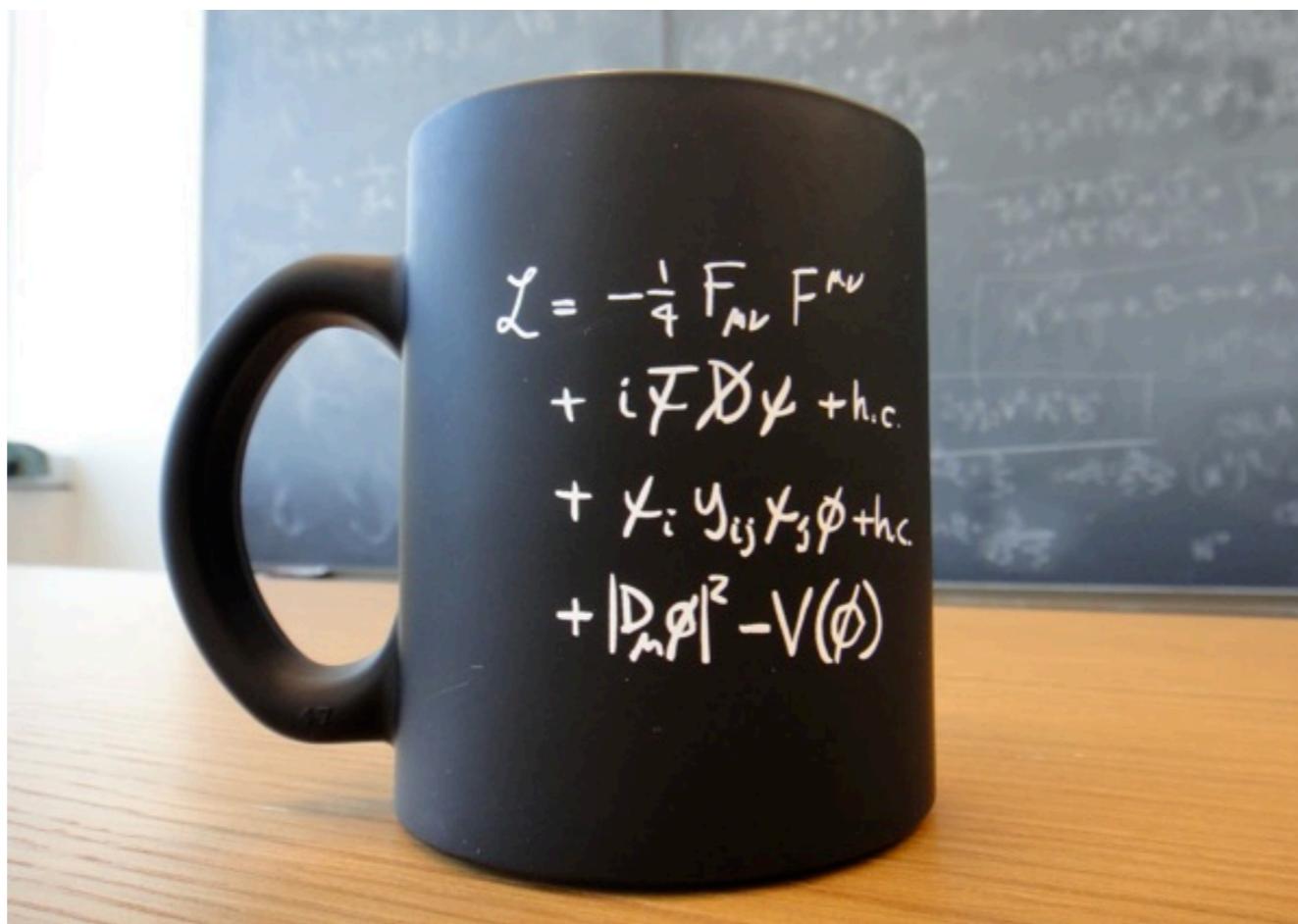
From evidence to observation!



Theory Background

The Standard Model

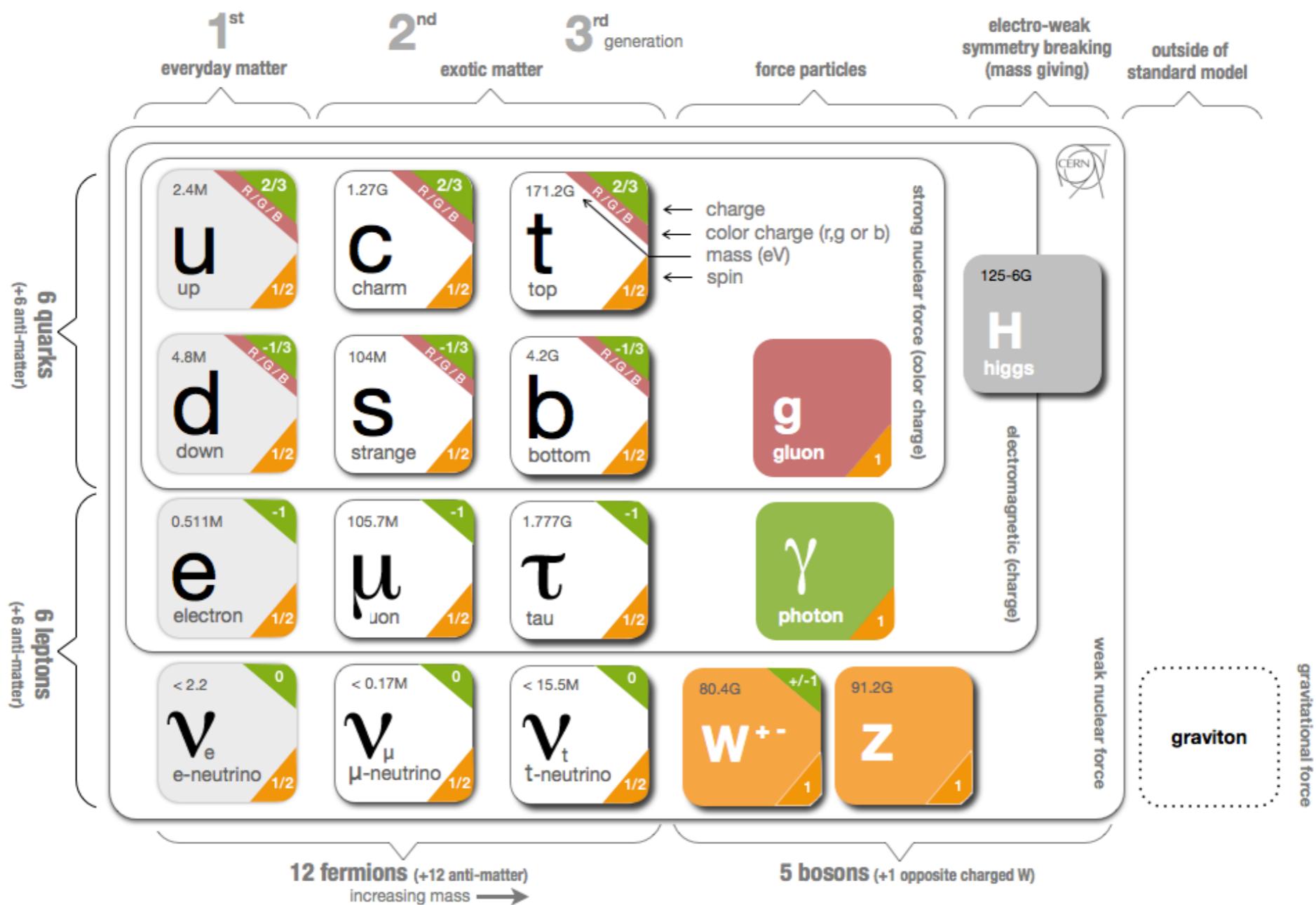
- The quantum field theory encapsulating our understanding of the fundamental constituents of our universe and their interactions.
- Derived from the symmetries that give rise to the **electromagnetic** and the **weak** and **strong** nuclear forces: **SU(3) × SU(2) × U(1)**.



Theory Background

The Standard Model

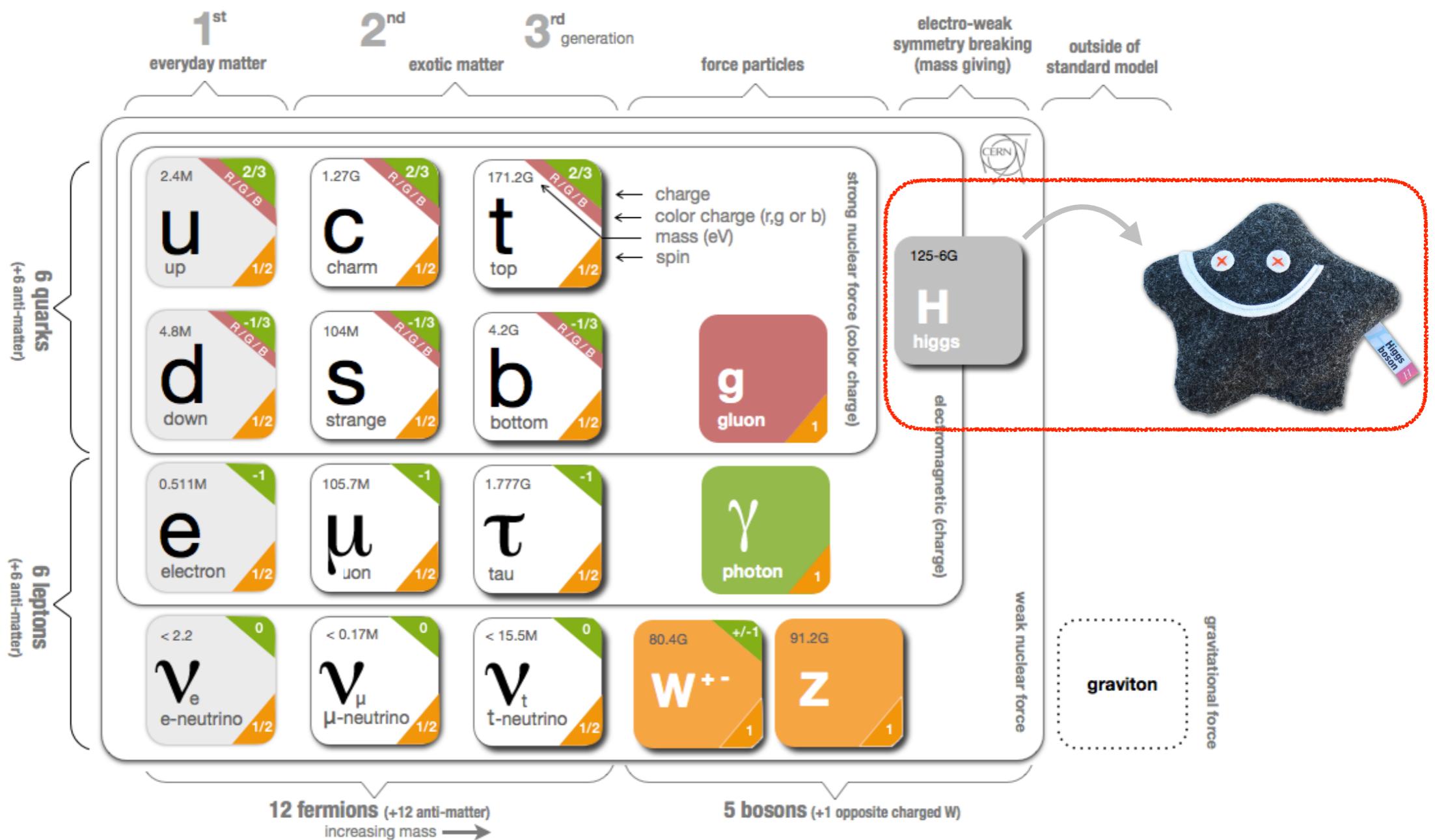
The fundamental particles are classified by their properties and interactions.



Theory Background

The Standard Model

The fundamental particles are classified by their properties and interactions.



Theory Background

The Higgs Mechanism

- Under SU(2) gauge invariance, the model predicts massless gauge bosons...
- ... but the weak vector bosons are **massive!**
- The Higgs mechanism introduces *spontaneous symmetry breaking* to reconcile the gauge invariance of the theory with the massive nature of the weak vector bosons.

VOLUME 13, NUMBER 9

PHYSICAL REVIEW LETTERS

31 AUGUST 1964

BROKEN SYMMETRY AND THE MASS OF GAUGE VECTOR MESONS*

F. Englert and R. Brout

Faculté des Sciences, Université Libre de Bruxelles, Bruxelles, Belgium

(Received 26 June 1964)

VOLUME 13, NUMBER 16

PHYSICAL REVIEW LETTERS

19 OCTOBER 1964

BROKEN SYMMETRIES AND THE MASSES OF GAUGE BOSONS

Peter W. Higgs

Tait Institute of Mathematical Physics, University of Edinburgh, Edinburgh, Scotland

(Received 31 August 1964)

VOLUME 13, NUMBER 20

PHYSICAL REVIEW LETTERS

16 NOVEMBER 1964

GLOBAL CONSERVATION LAWS AND MASSLESS PARTICLES*

G. S. Guralnik,[†] C. R. Hagen,[‡] and T. W. B. Kibble

Department of Physics, Imperial College, London, England

(Received 12 October 1964)

The Higgs Mechanism

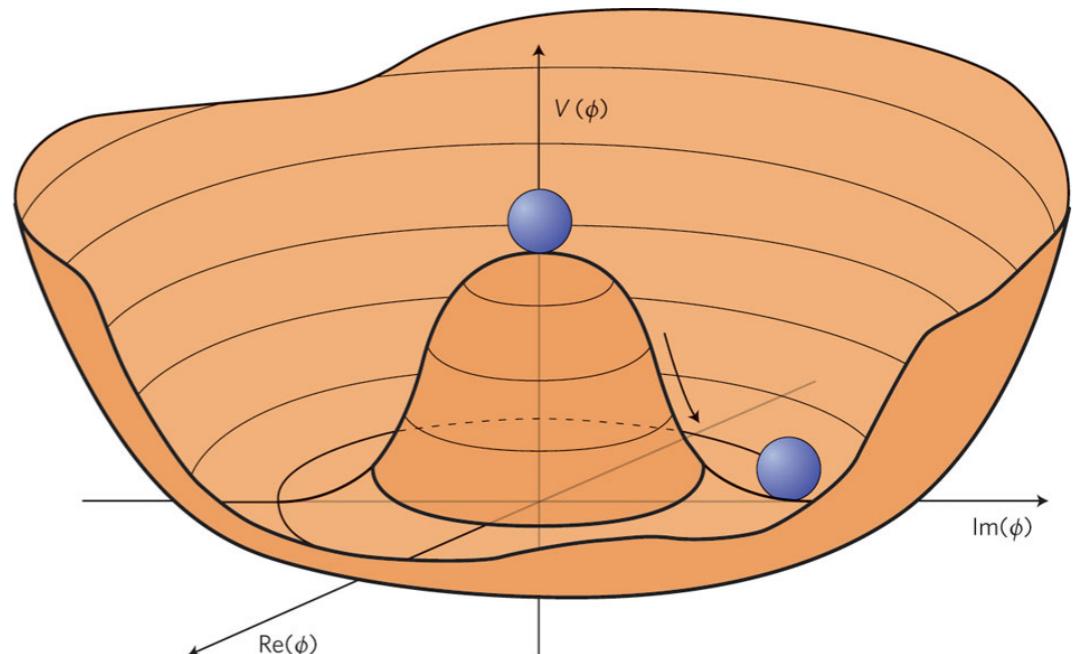
- Suppose a scalar field is present with a potential of the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda}{2} (\phi^\dagger \phi)^2$$

- When $\mu^2 > 0$, the potential is parabolic with a minimum at zero. When $\mu^2 < 0$, the potential takes on a “sombrero-like” shape which attains its minimum at

$$v = \left(\frac{-\mu^2}{\lambda} \right)^{1/2}$$

- The field now attains a vacuum expectation value in its ground state and its symmetry has been spontaneously broken.



The Higgs Mechanism

- The gauge boson masses arise from the kinetic term of the Lagrangian

$$\mathcal{L}_\phi = \frac{1}{2}(\partial_\mu H)^2 + \frac{g^2}{8}(W_\mu^1)^2 v^2 + \frac{g^2}{8}(W_\mu^2)^2 v^2 + \frac{1}{8} (g'B_\mu - gW_\mu^3)^2 v^2$$

$$m_W = \frac{gv}{2}$$

$$m_Z = \frac{\sqrt{g^2 + g'^2}v}{2}$$

- The fermions acquire mass through their Yukawa interactions with the Higgs field

$$m_\ell = \frac{\lambda_\ell v}{\sqrt{2}}$$

Leptons

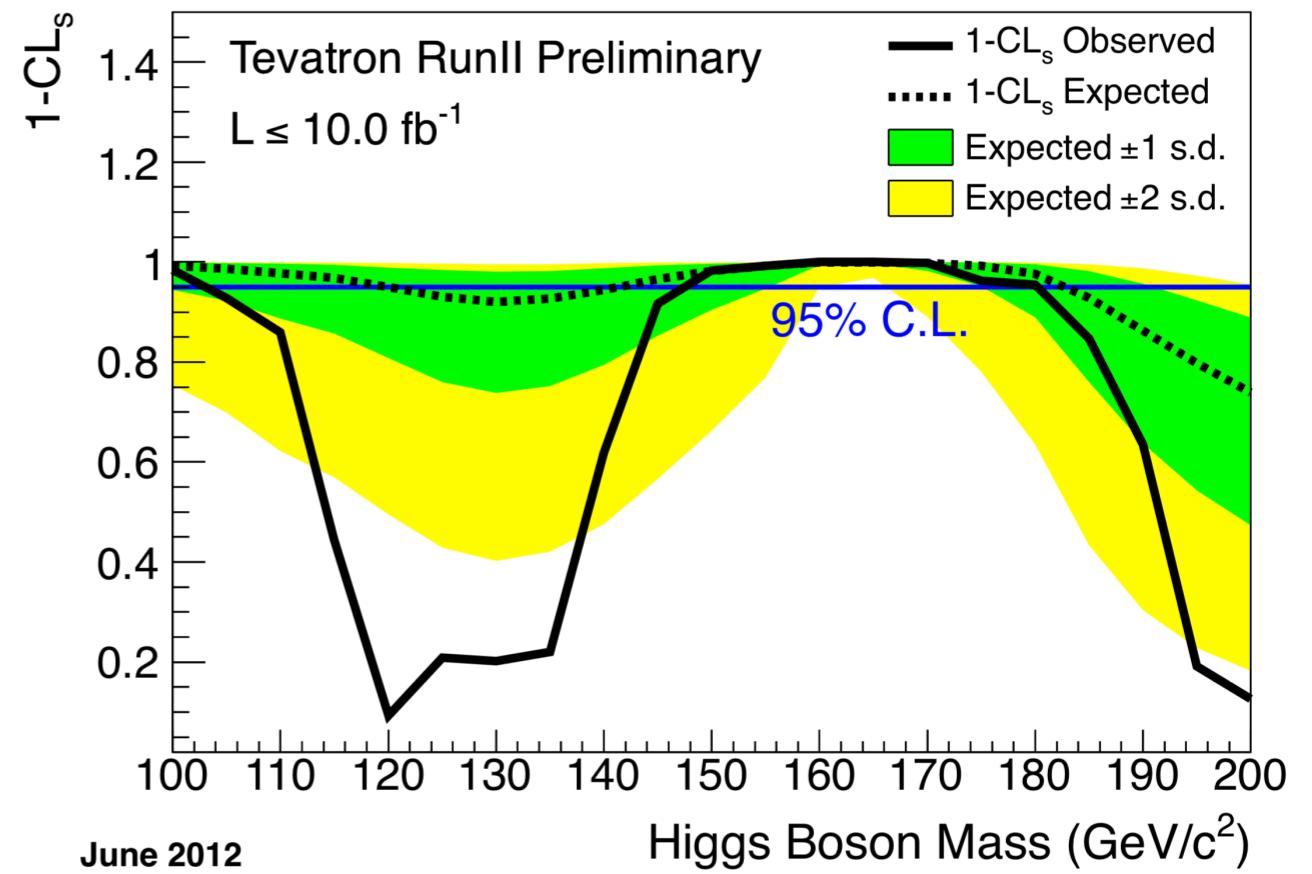
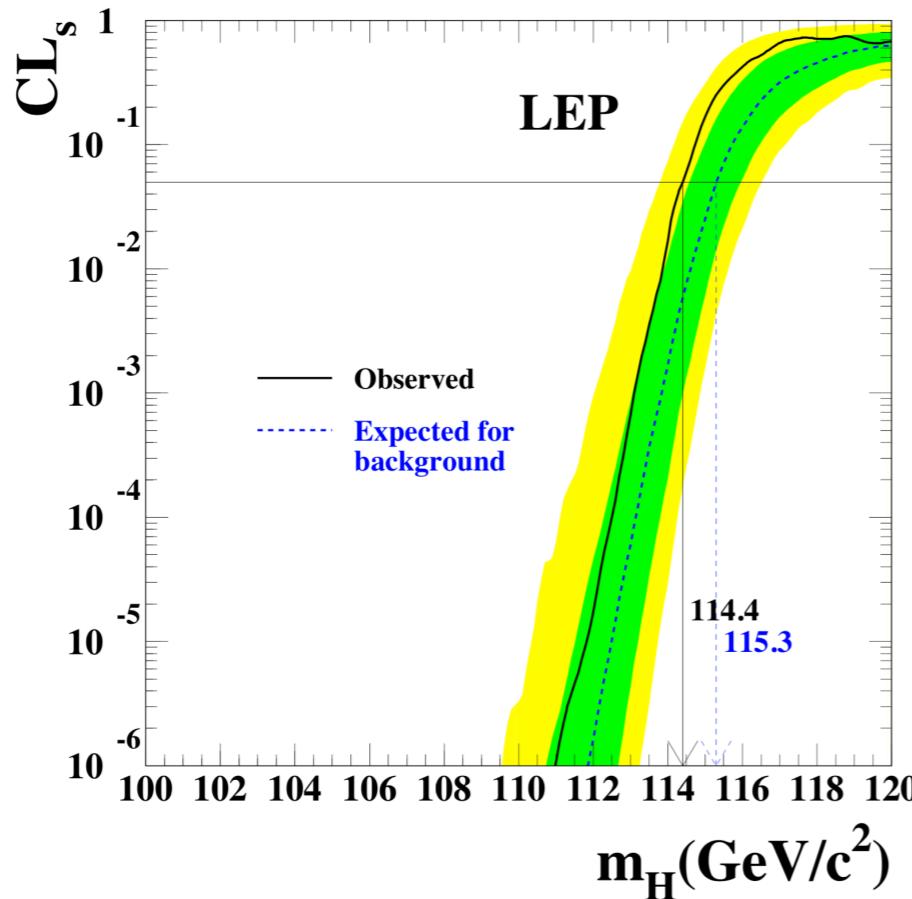
$$M_d^{ij} = \frac{\Lambda_d^{ij} v}{\sqrt{2}} , \quad M_u^{ij} = \frac{\Lambda_u^{ij} v}{\sqrt{2}}$$

Quarks

Theory Background

The Higgs Boson

- The excitations of the Higgs field manifest as a particle: a Higgs boson.
- Its mass is an unknown parameter of the Standard Model:
 - Theoretical constraints place it within a range of 100 GeV to 1 TeV.
 - Searches for a light Higgs at LEP established a lower bound of 114.4 GeV.
 - Analysis of the full Tevatron dataset placed it between 103 GeV and 147 GeV.

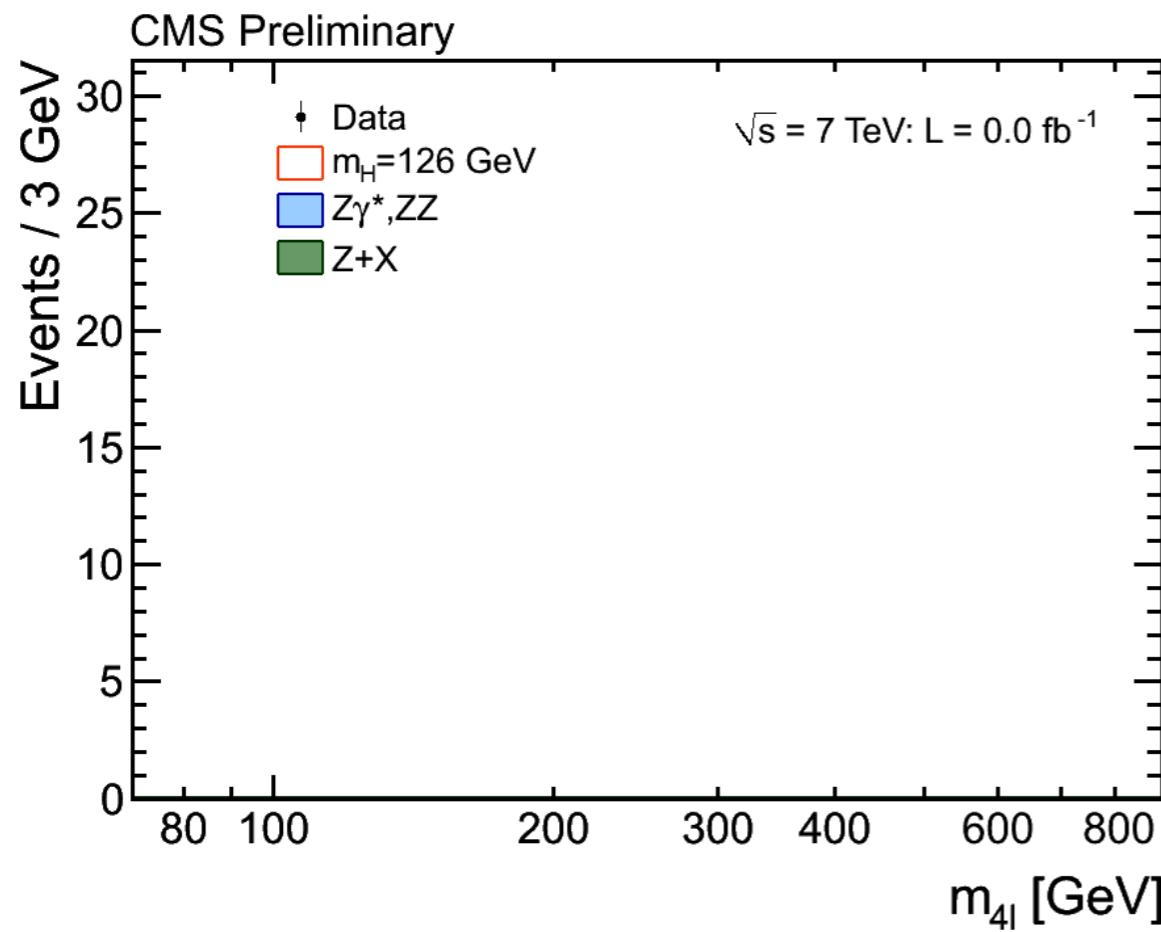


June 2012

The Higgs Boson

On July 4, 2012, CERN announced that the Higgs boson was jointly discovered at the LHC by the ATLAS and CMS collaborations with a mass of $m_H = 125.09$ GeV.

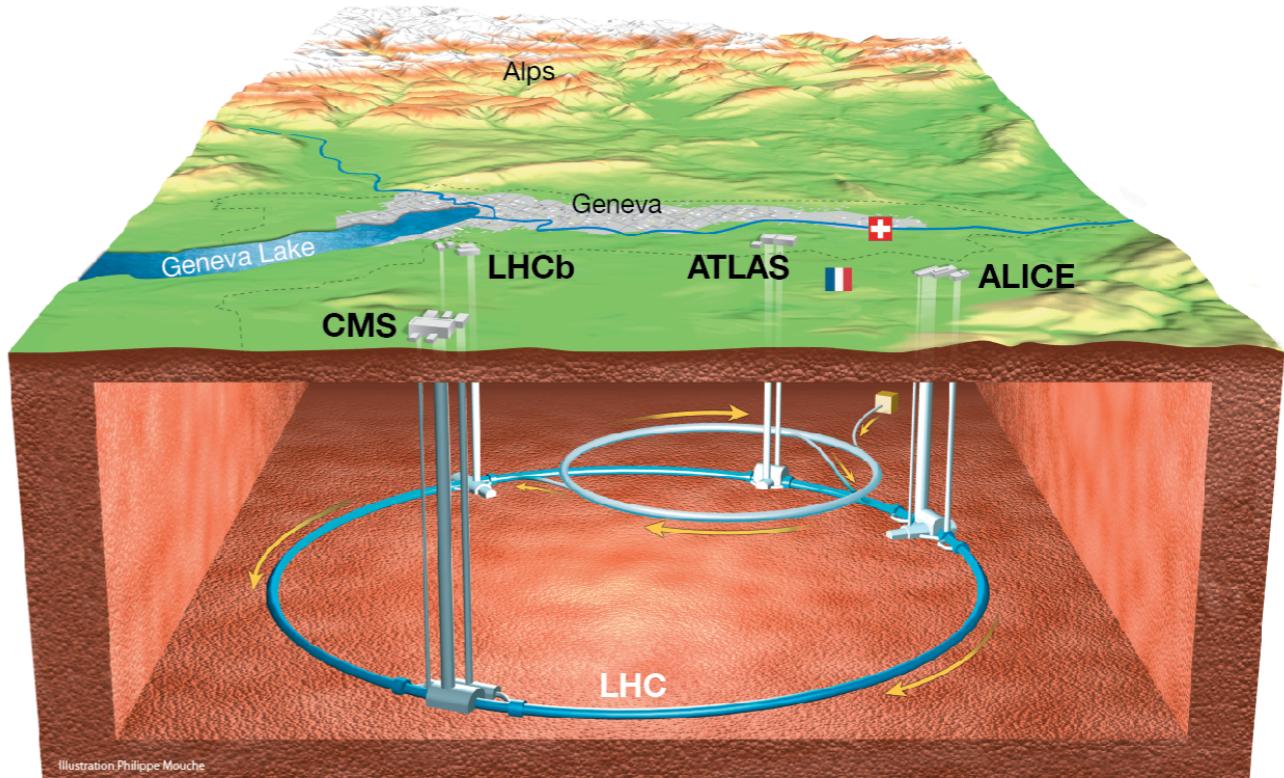
- Within a year, its spin, parity, and several couplings were found to be consistent with predictions.
- The 2013 Nobel Prize in Physics was awarded jointly to François Englert and Peter Higgs.



The CMS Detector

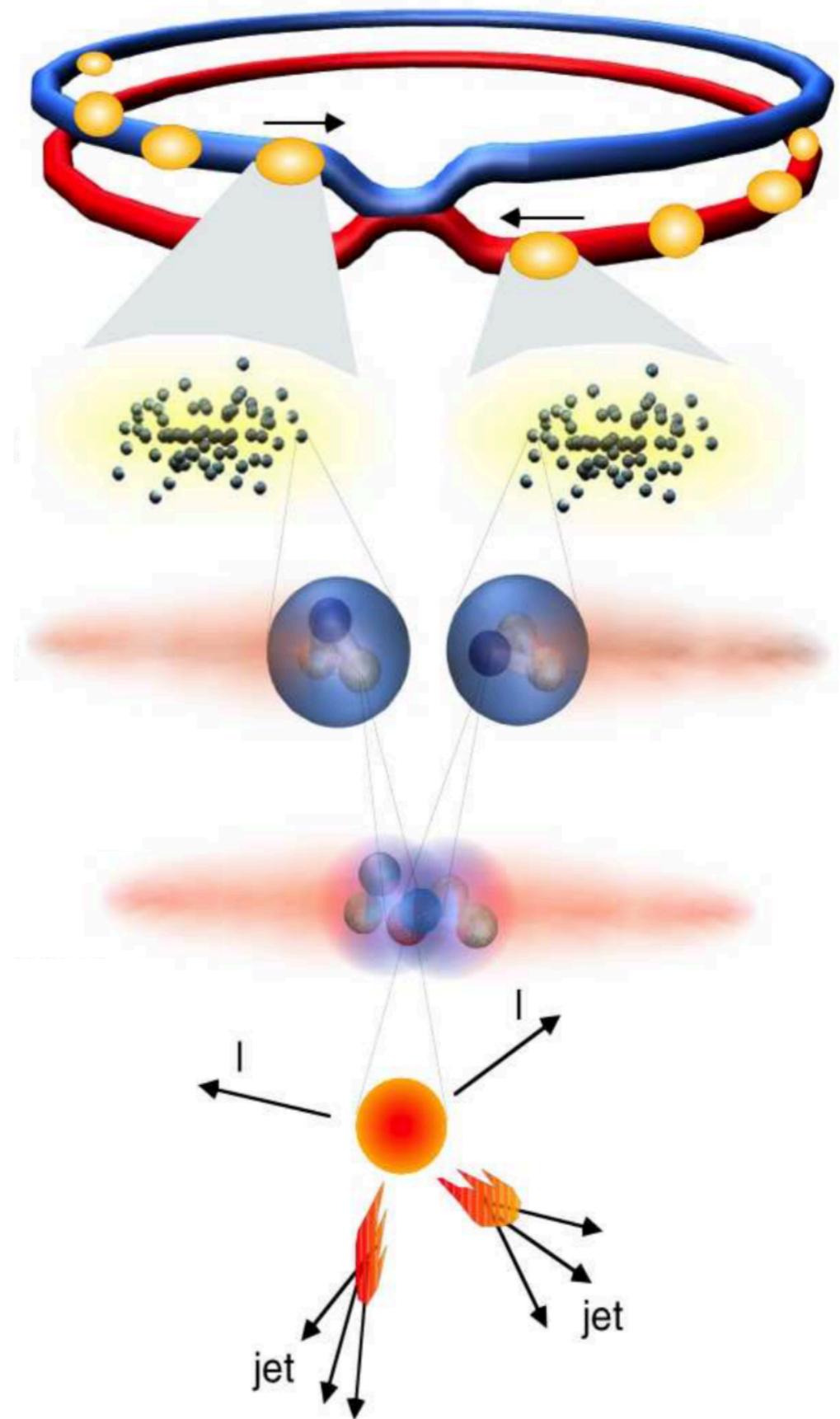
The Large Hadron Collider

- Final stage of CERN's accelerator complex, located 100 meters beneath the Franco-Swiss border.
- Two circulating beams along its 17 mile circumference, guided by 1,232 bending magnets and 392 focusing magnets.
- Four main experiments located at designated collision points: ALICE, ATLAS, CMS, and LHCb.



Proton-Proton Collisions

- Two proton beams with energies of 6.5 TeV.
- 2556 bunches per beam, spaced 25 ns apart.
- Each bunch has around $\sim 10^{11}$ protons.
- Instantaneous luminosity of $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$.
- Produces one billion collisions per second which are filtered and recorded by one of the particle detectors for data analysis.



The CMS Detector

The CMS Detector

CMS DETECTOR

Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel ($100 \times 150 \mu\text{m}$) $\sim 16\text{m}^2 \sim 66\text{M}$ channels
Microstrips ($80 \times 180 \mu\text{m}$) $\sim 200\text{m}^2 \sim 9.6\text{M}$ channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying $\sim 18,000\text{A}$

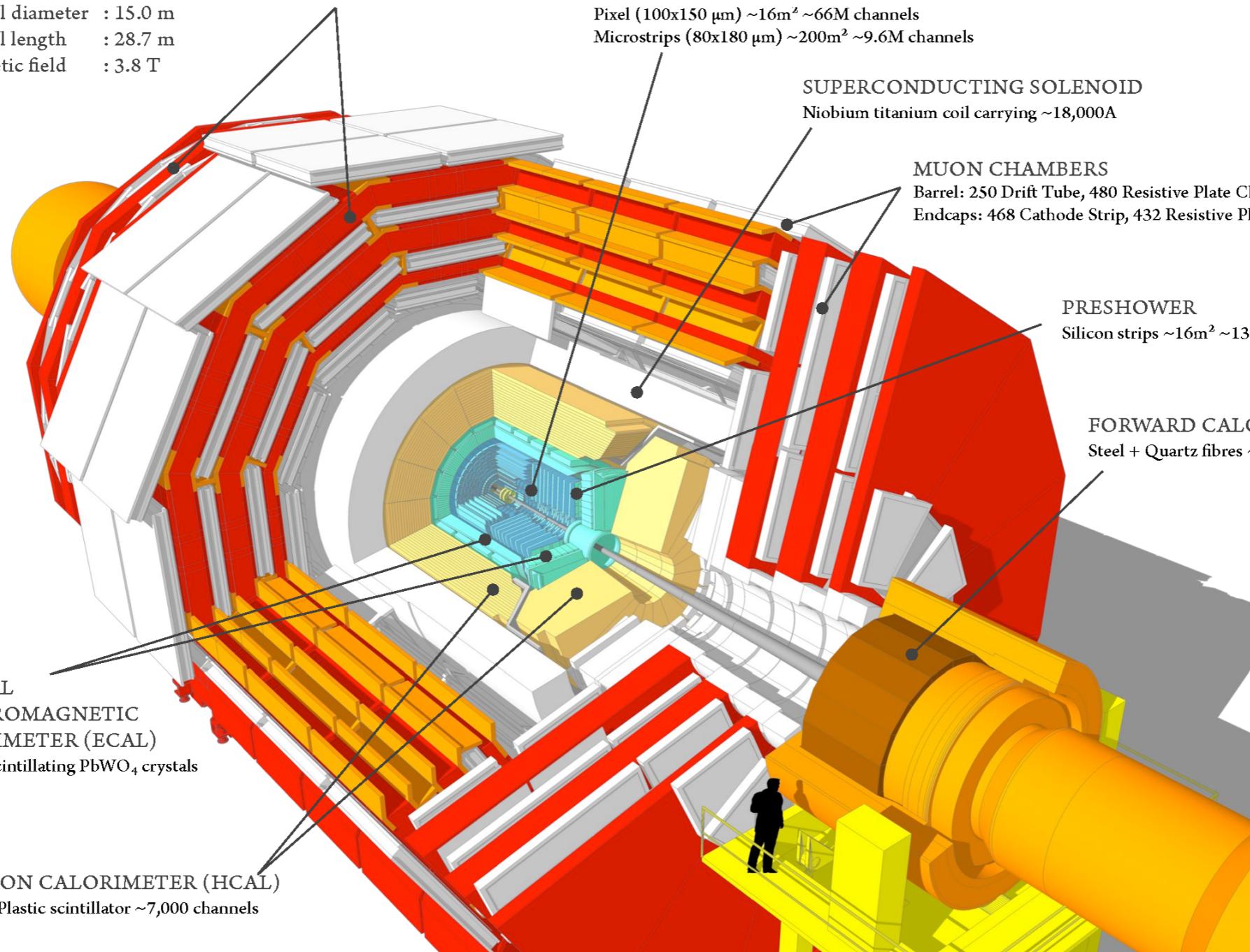
MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

PRESHOWER
Silicon strips $\sim 16\text{m}^2 \sim 137,000$ channels

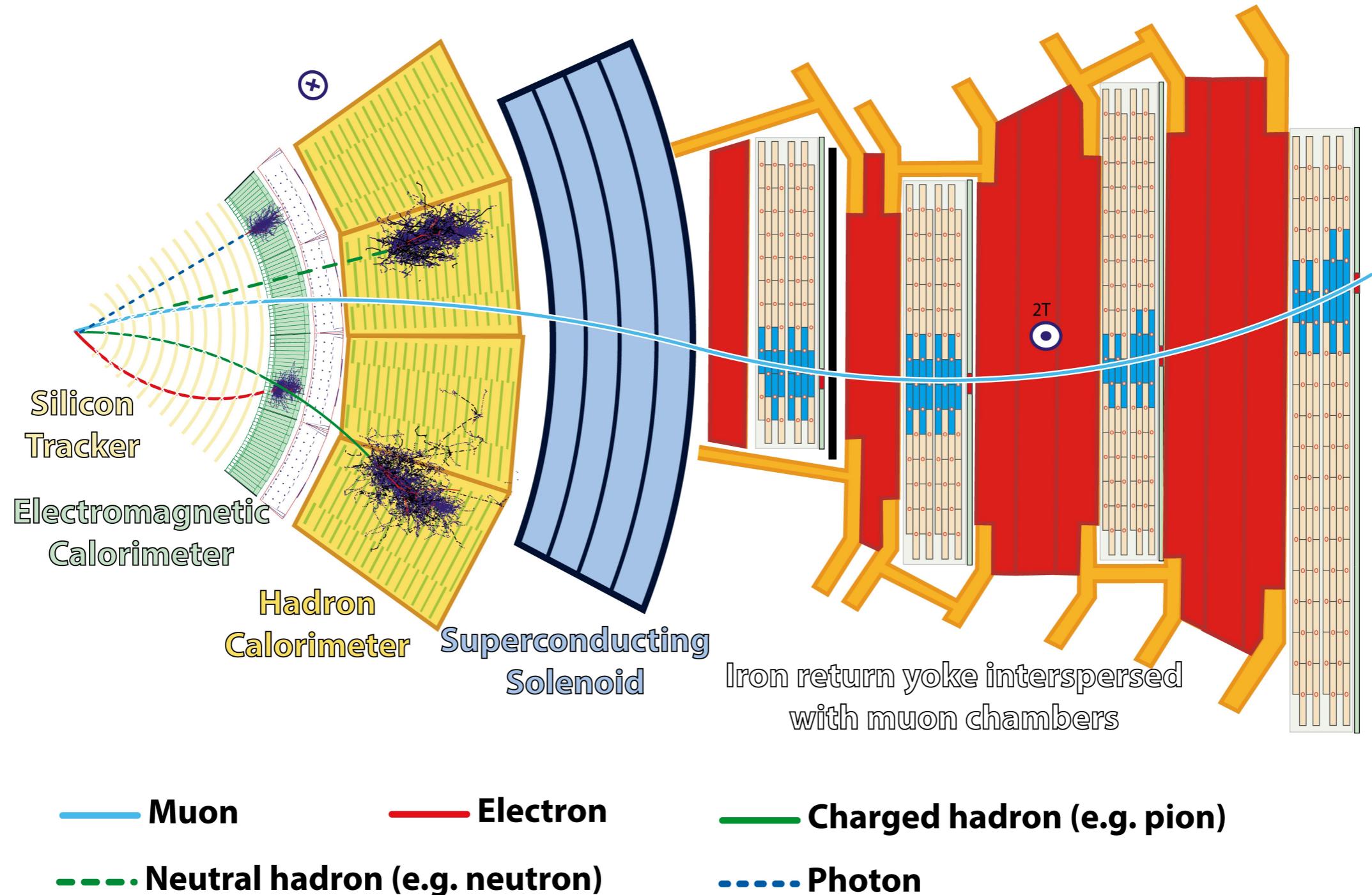
FORWARD CALORIMETER
Steel + Quartz fibres $\sim 2,000$ Channels

CRYSTAL
ELECTROMAGNETIC
CALORIMETER (ECAL)
 $\sim 76,000$ scintillating PbWO_4 crystals

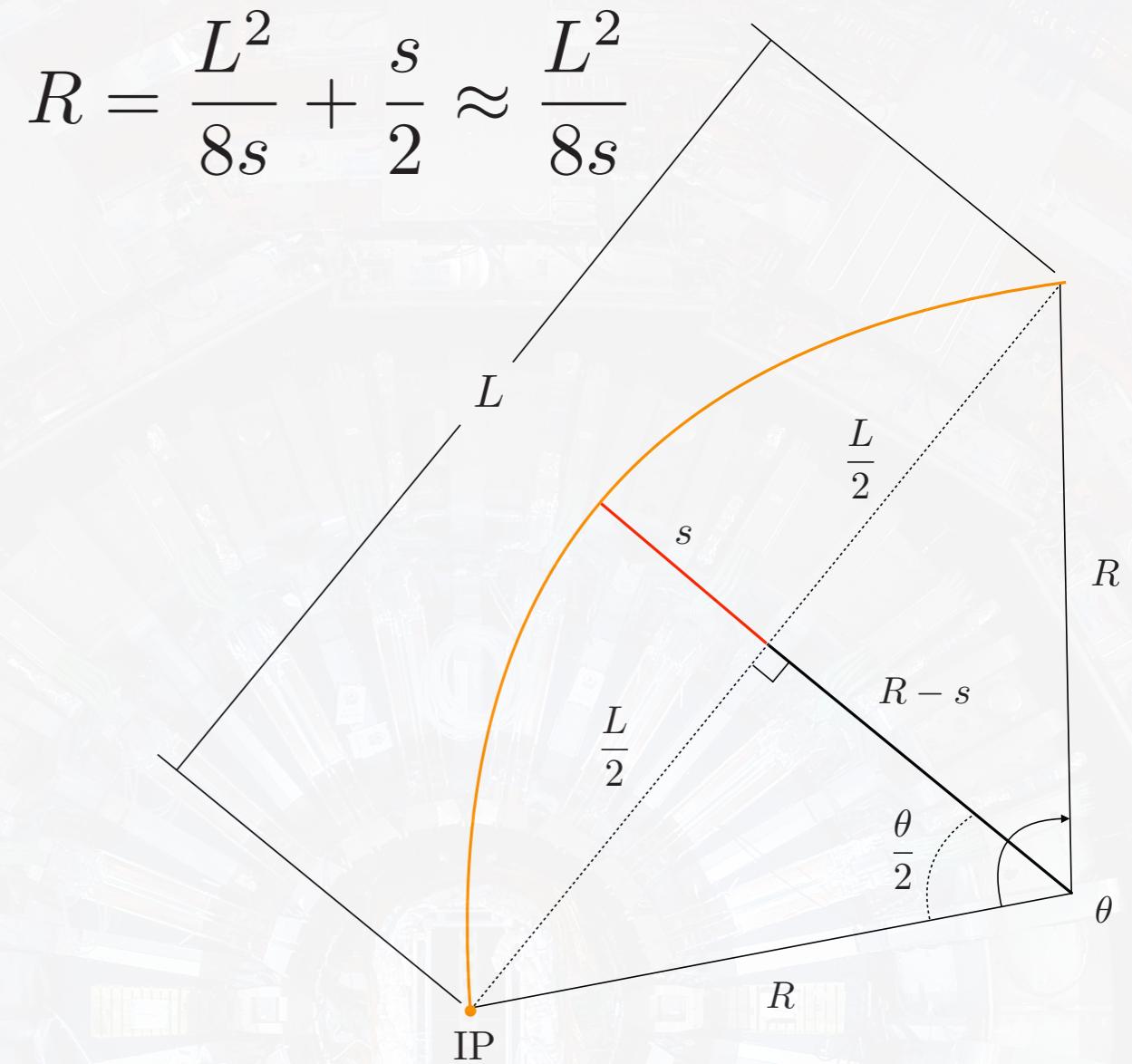
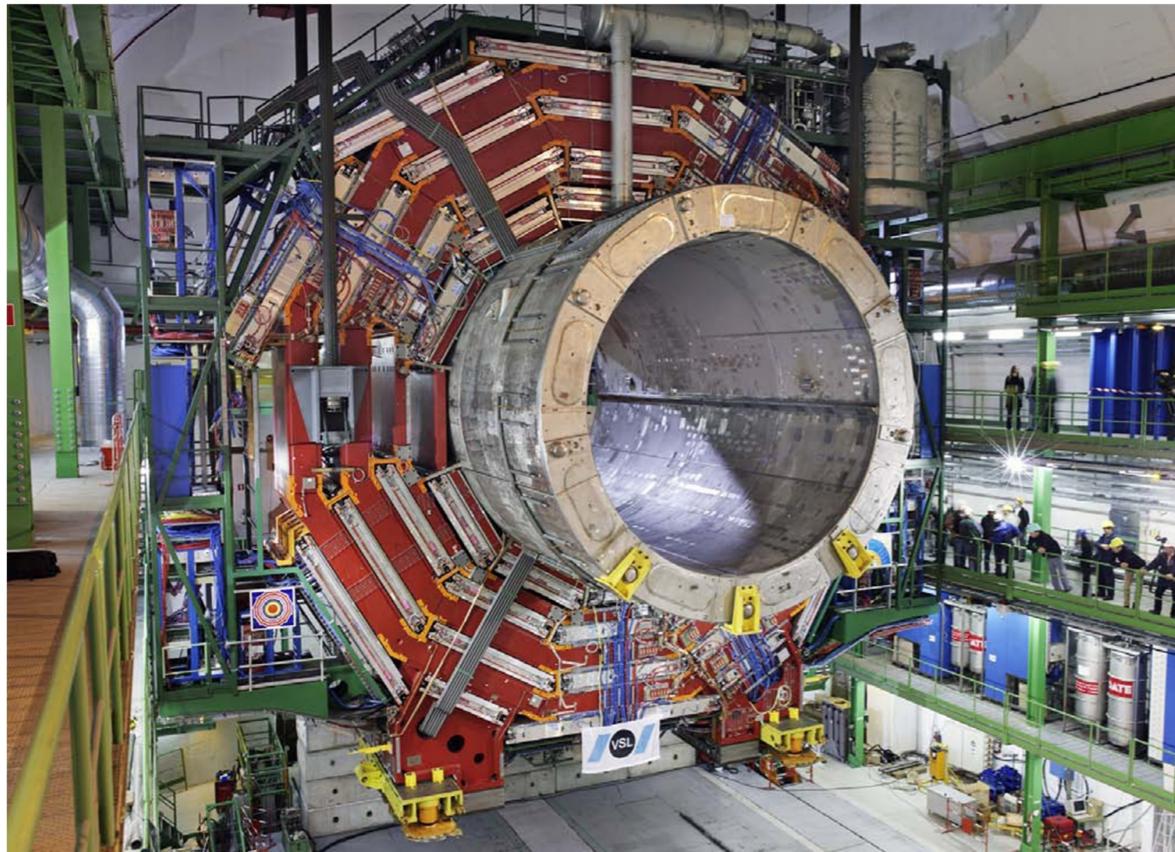
HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator $\sim 7,000$ channels



Layered Design



Superconducting Solenoid



A particle's transverse momentum can be measured from their "track".

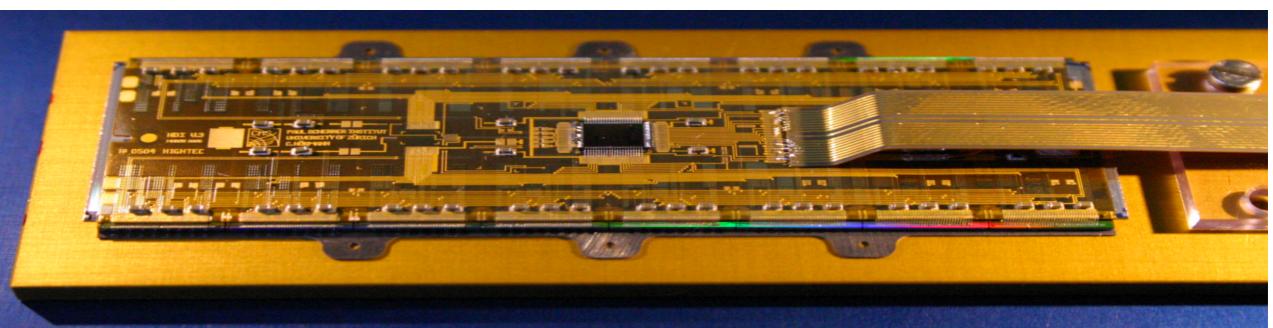
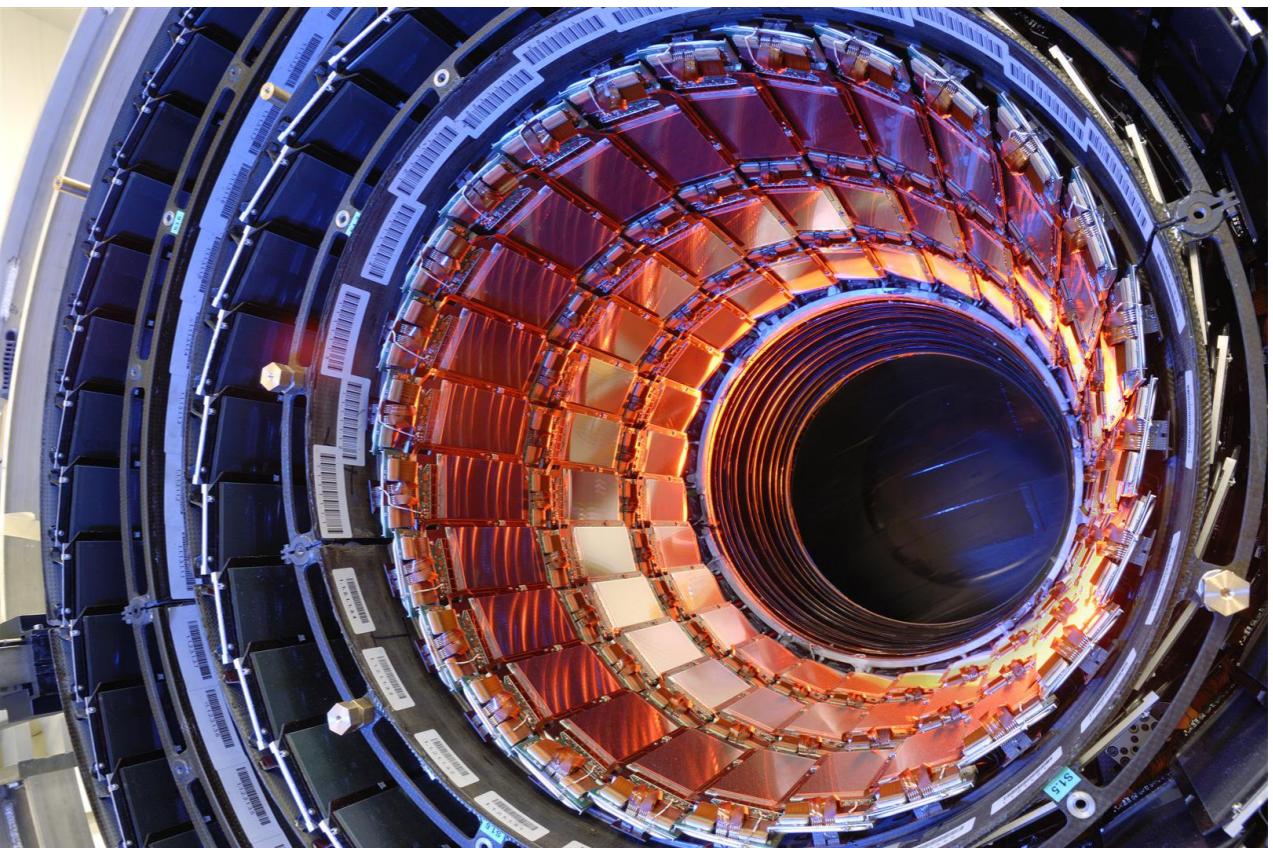
- Niobium-titanium coil cooled to 4.2 K and carrying 18,160 A.
- Produces a 3.8 T magnetic field that bends particle trajectories.

$$p_T = qRB = \frac{qBL^2}{8s}$$

$$\frac{\sigma_{p_T}}{p_T} \sim p_T \times \frac{\sigma_s}{BL^2}$$

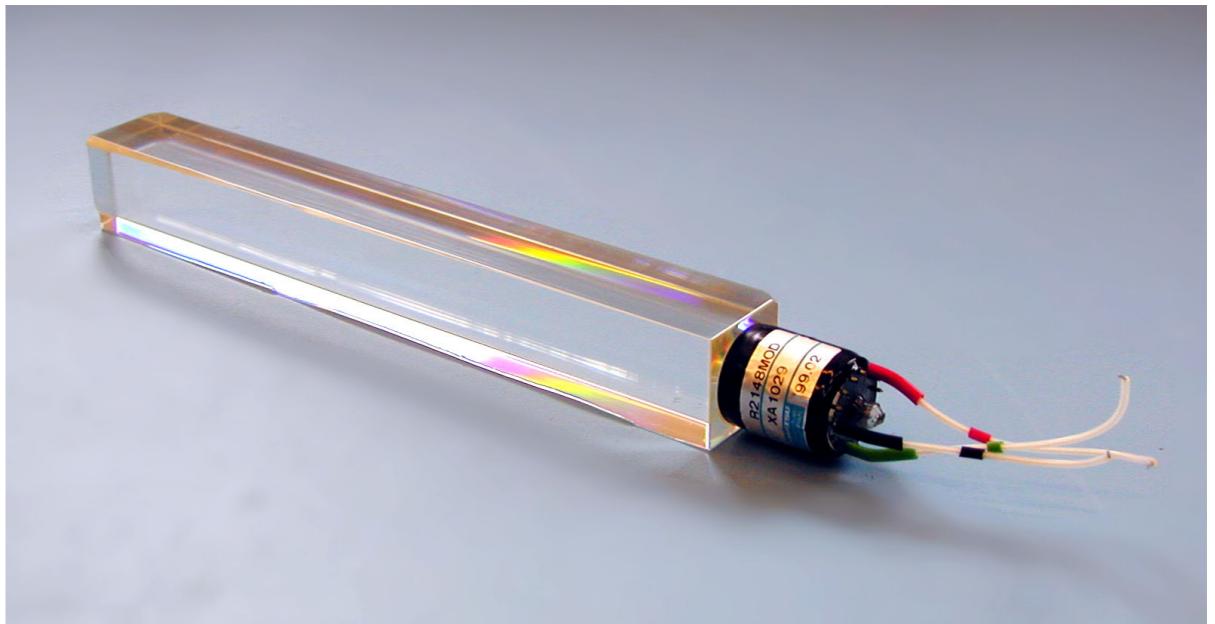
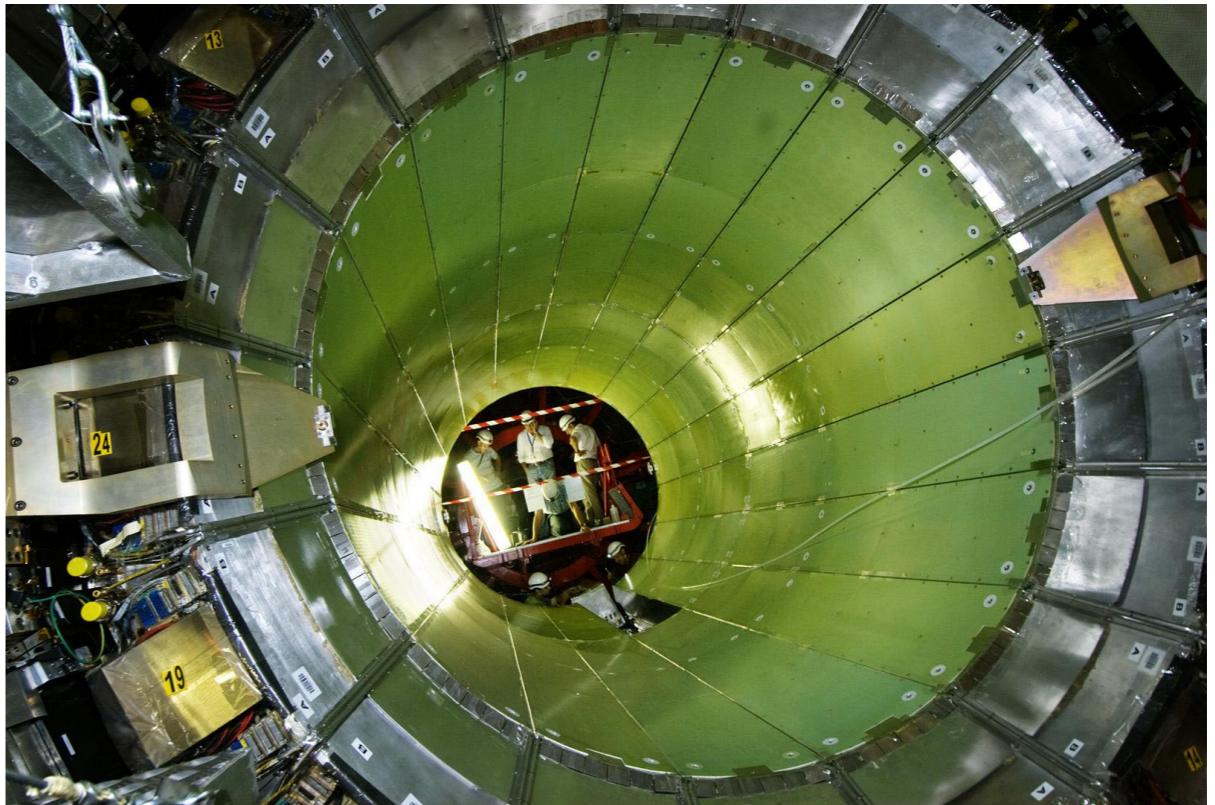
Silicon Tracker

- The innermost layer surrounding the interaction point requires
 - fast response time,
 - excellent spatial resolution,
 - and high radiation tolerancein order to reconstruct vertices and charged particle tracks.
- Composed of 16,588 silicon sensor modules:
 - 66 million pixels
($100 \mu\text{m} \times 150 \mu\text{m}$)
 - 9.6 million microstrips
(180 μm wide, 10 to 25 cm long)
- p_T resolution $\approx 1.5\%$ for charged particles with $1 \text{ GeV} < p_T < 10 \text{ GeV}$



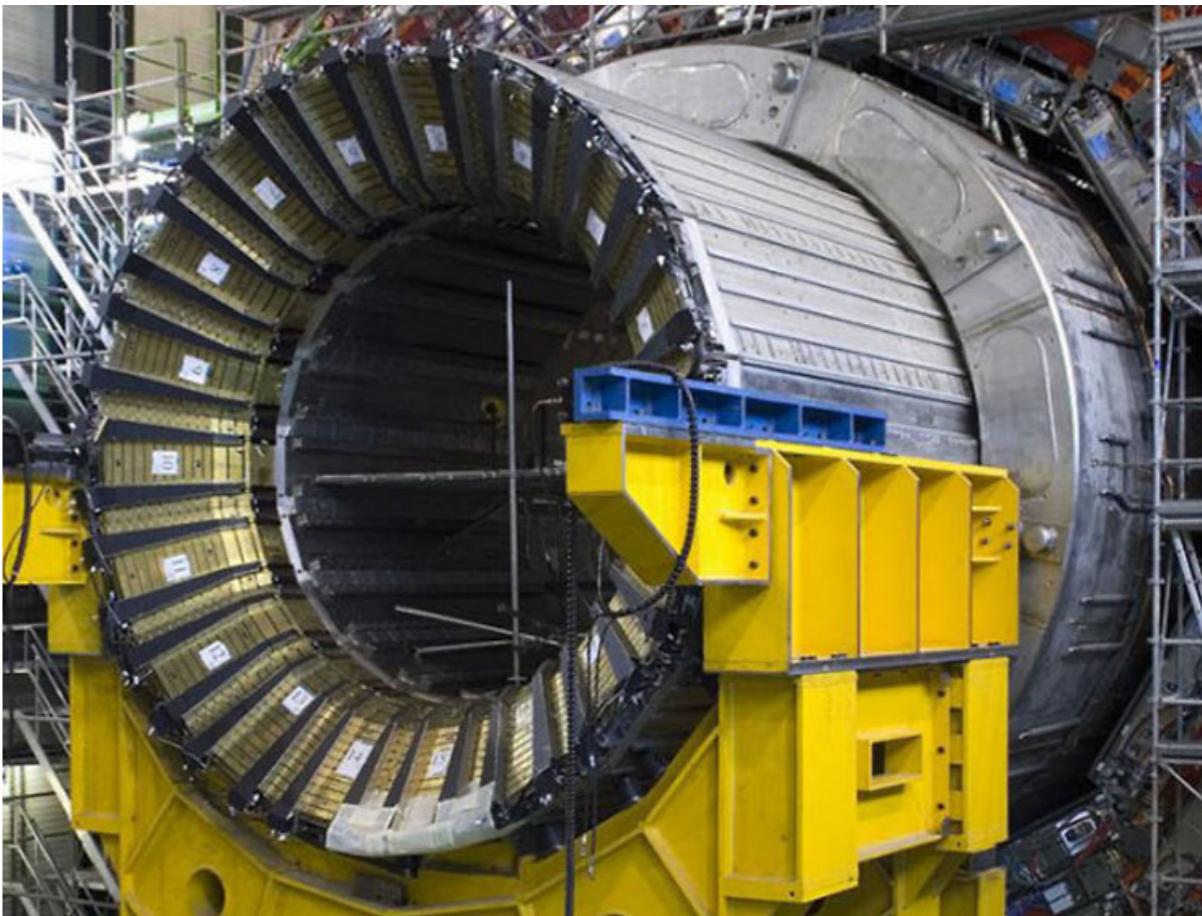
Electromagnetic Calorimeter

- The ECAL requires
 - a short radiation length and Molière radius,
 - fast light yield,
 - and high radiation tolerance
- in order to measure the energy deposits of electrons and photons.
- Consists of 75,848 PbWO₄ crystals.
- Lead and silicon pre-shower layer used to identify neutral pion decays.
- Energy resolution of 2-5% for electrons and 1.1-5% for photons.



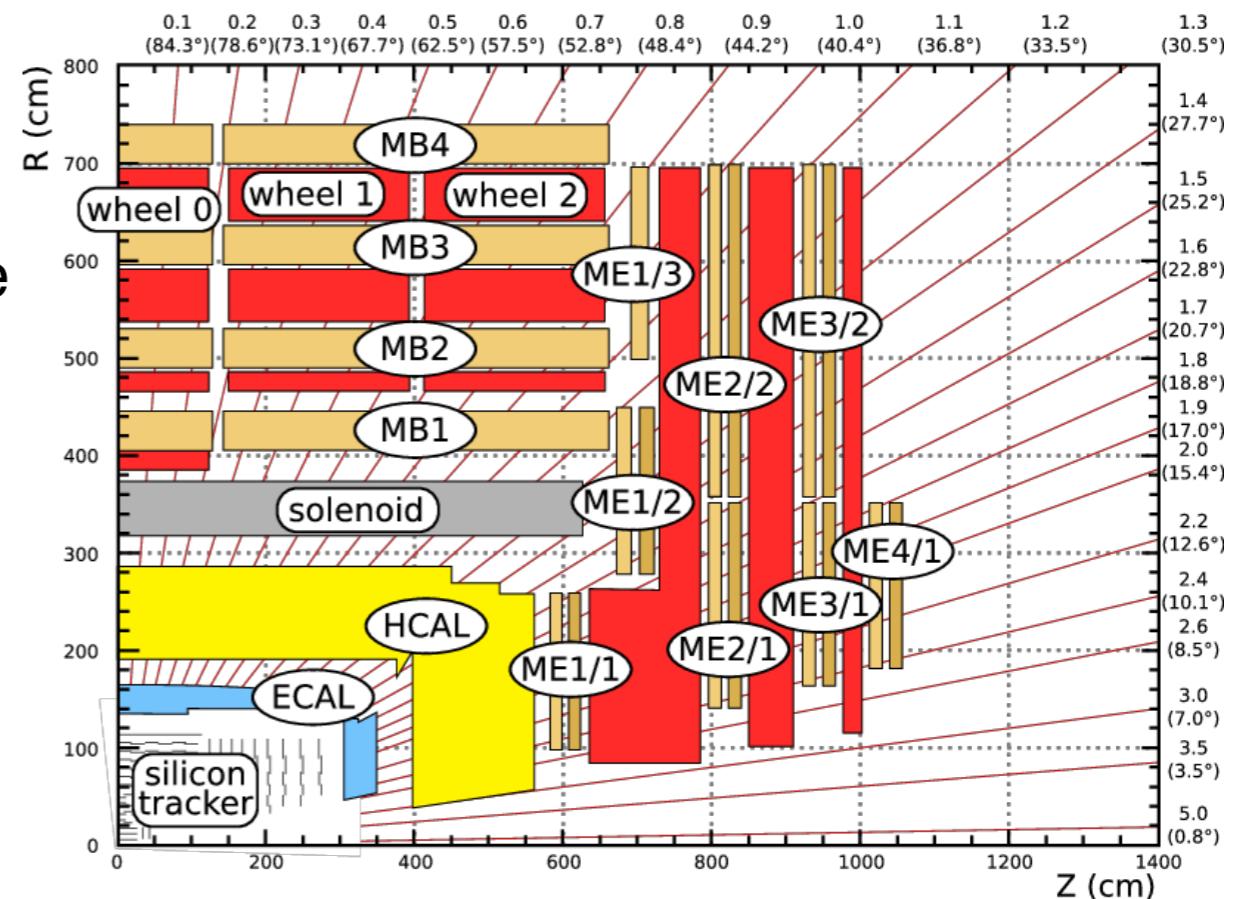
Hadronic Calorimeter

- The HCAL requires a short nuclear interaction length in order to measure the energy deposits of charged and neutral hadrons.
- Alternating layers of brass absorbers and plastic scintillators. Half of the brass came from WW2 Russian artillery shells.
- Optical signals are carried by wavelength shifting fibers and amplified by hybrid photodiodes.
- Hadronic Forward detectors made of steel absorbers and quartz fibers extend the coverage from $| \eta | < 3.0$ to $| \eta | < 5.0$.



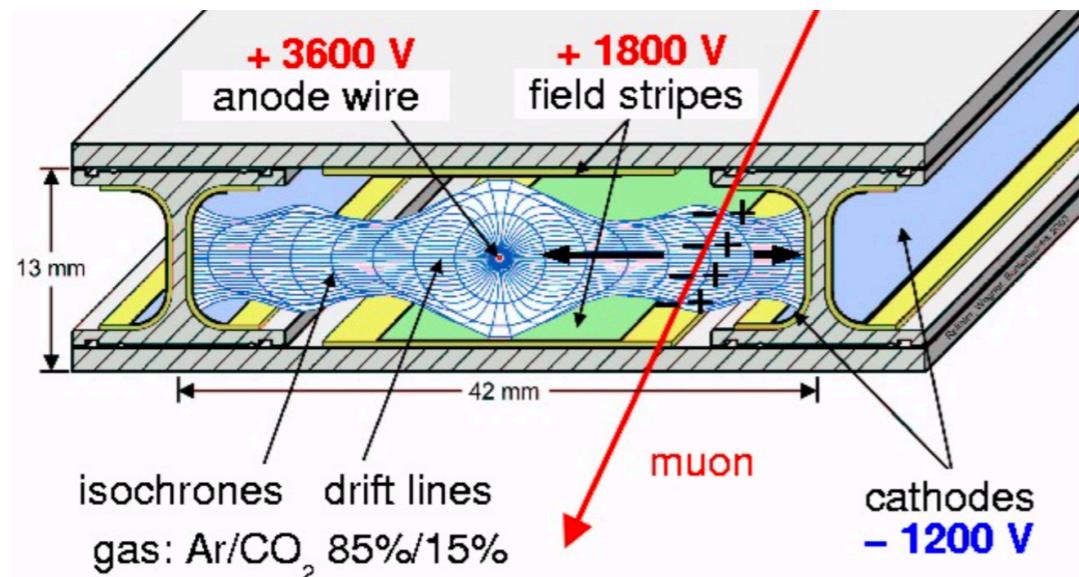
Muon Detectors

- Muons are weakly interacting and easily pass through the inner layers.
- Outermost layer of the detector uses three different types of gas ionization chambers that reconstruct muon tracks:
 - Drift Tubes (DTs)
 - Resistive Plate Chambers (RPCs)
 - Cathode Strip Chambers (CSCs)

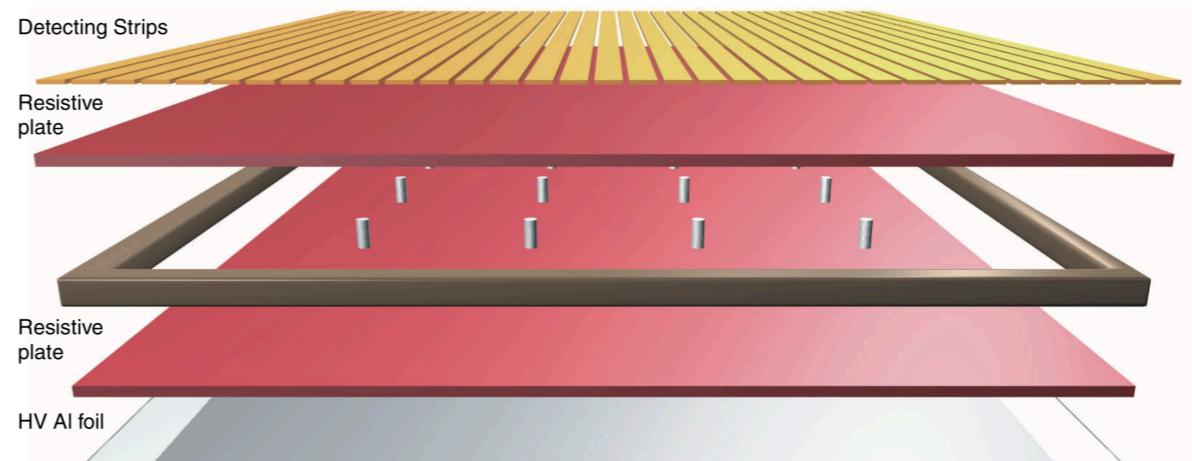


Muon Subsystems

Drift Tubes



Resistive Plate Chambers

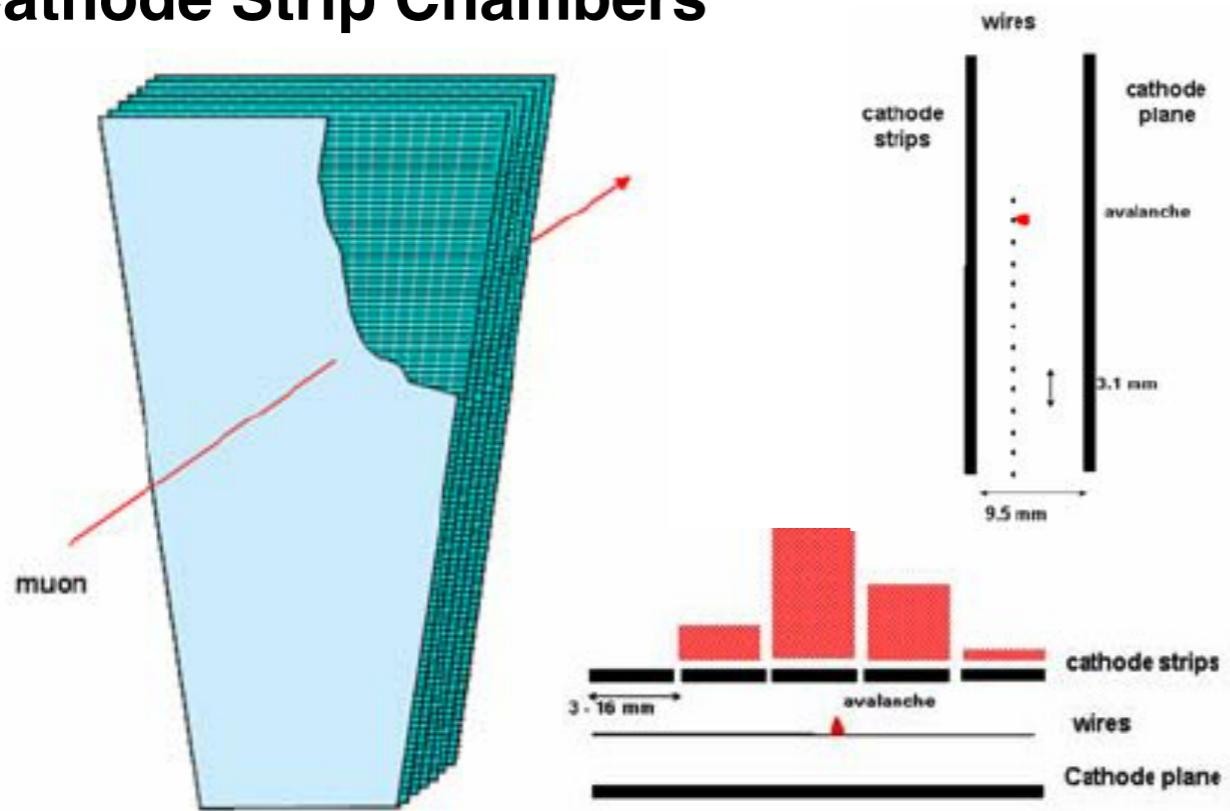


- Exclusive to the barrel region.
- 4 cm wide tubes of aluminum plates and I-beams enclosing a positive wire.
- Measures two spatial coordinates with a resolution of 80-120 μm .

- Used in both barrel and endcap regions.
- Two resin plates surrounding a 2 mm gap filled with tetrafluoroethane (95%) and isobutane (5%)
- Achieves a 1 ns time resolution and a spatial resolution of 0.8-1.2 cm

Muon Subsystems

Cathode Strip Chambers

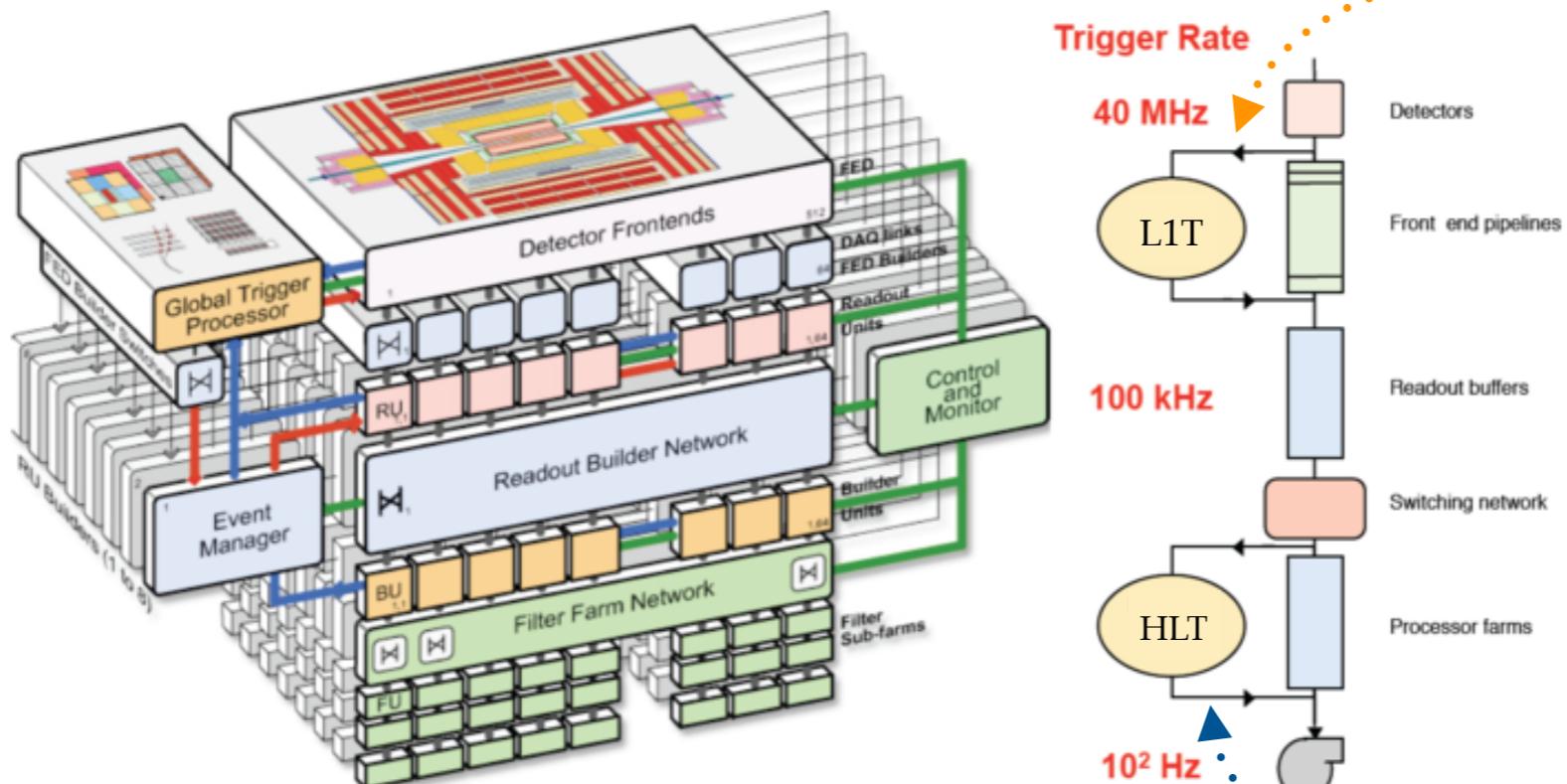


- Exclusive to the endcap region.
- Trapezoidal chamber with 6 gas gaps of 40% Ar, 50% CO₂, and 10% CF₄ containing radial cathode strips and perpendicular anode wires.
- Measures r-φ positions with a resolution of 40-150 μm.



Trigger and DAQ Systems

- Each collision event takes ≈ 1 MB of storage, generating almost 1 PB/s of data across millions of channels!
- Data readout, processing, and storage is handled by hardware and software.



- Accepted events are sent to Tier-0 computing center for storage and further reconstruction.
- Propagated to Tier-1 and Tier-2 data centers for analysis.

Level 1 Trigger

- Custom FPGAs and ASICs with fixed latency of 4 μ s.
- Local triggers and global trigger accept or reject events based on trigger primitives.
- Reduces event rate to a maximum of 100 kHz.

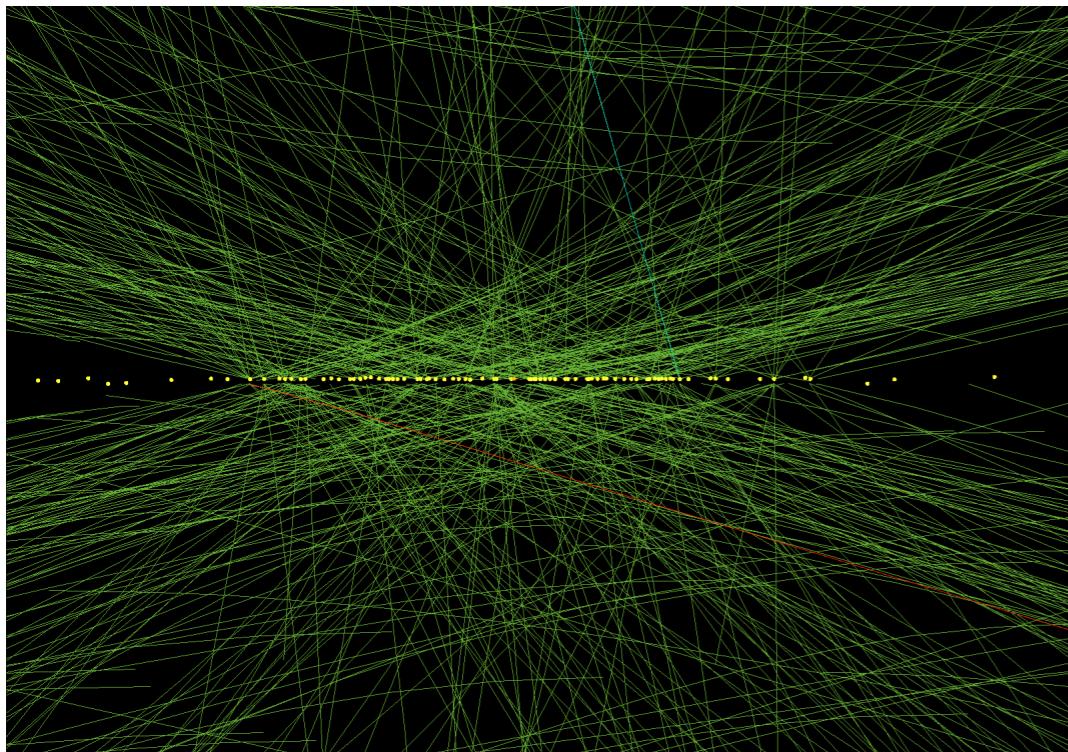
High Level Trigger

- Software executed on a cluster with 26,000 cores.
- Average processing time of 260 ms per event.
- Several hundred algorithms for specific physics goals
- Reduces event rate to a maximum of 100 Hz.

Event Reconstruction

Tracks and Vertices

- The trajectories of charged particles are reconstructed from their signals.
- In-time and out-of-time pileup pose a challenge for time complexity and efficiency.



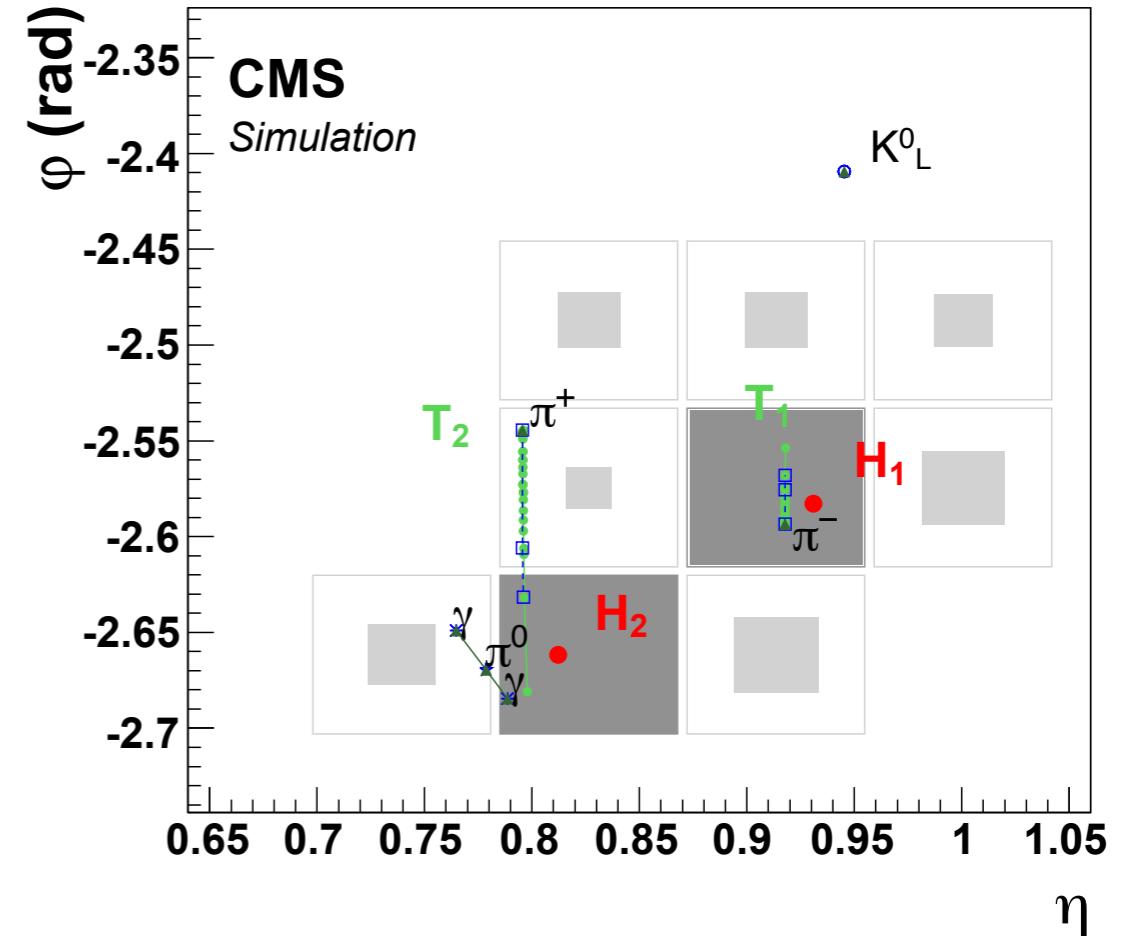
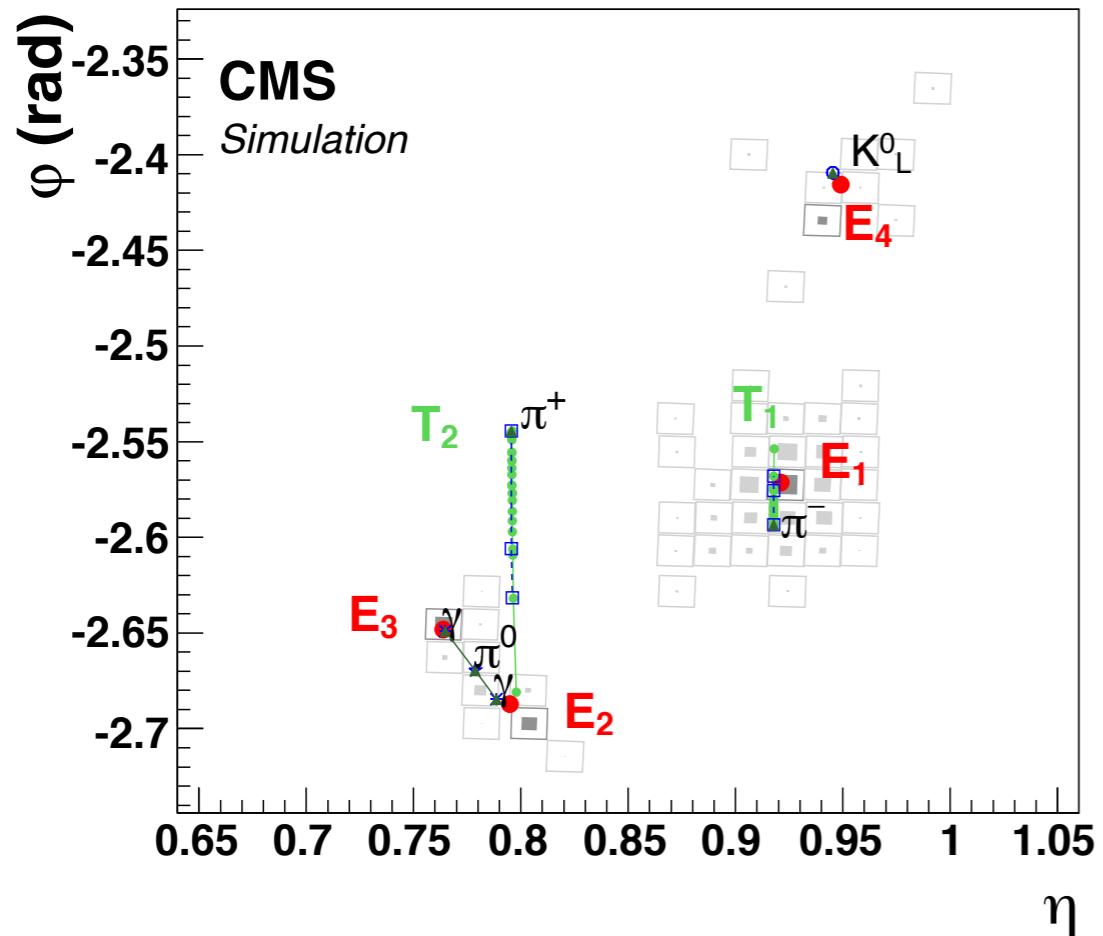
Iteration	Seed Type	Reconstructed Tracks
1	pixel triplets	prompt, high p_T
2	pixel triplets	from b -hadron decays, $R \lesssim 5$ cm
3	pixel triplets	prompt, low p_T
4	pixel pairs	recover high p_T
5	pixel+strip triplets	displaced, $R \lesssim 7$ cm
6	strip triplets/pairs	very displaced, $R \lesssim 25$ cm
7	strip triplets/pairs	very displaced, $R \lesssim 60$ cm
8	pixel+strip pairs	inside high p_T jets
9	muon-tagged tracks	muons
10	muon detectors	muons

- Iterative tracking algorithm applies a Kalman filter based combinatorial track finder for extrapolation to compatible hits.
- 90% reconstruction efficiency for tracks with p_T between 1 GeV and 10 GeV.
- Primary vertex candidates formed by clustering tracks with a deterministic annealing algorithm then applying an adaptive vertex fit.

Event Reconstruction

Calorimeter Clusters

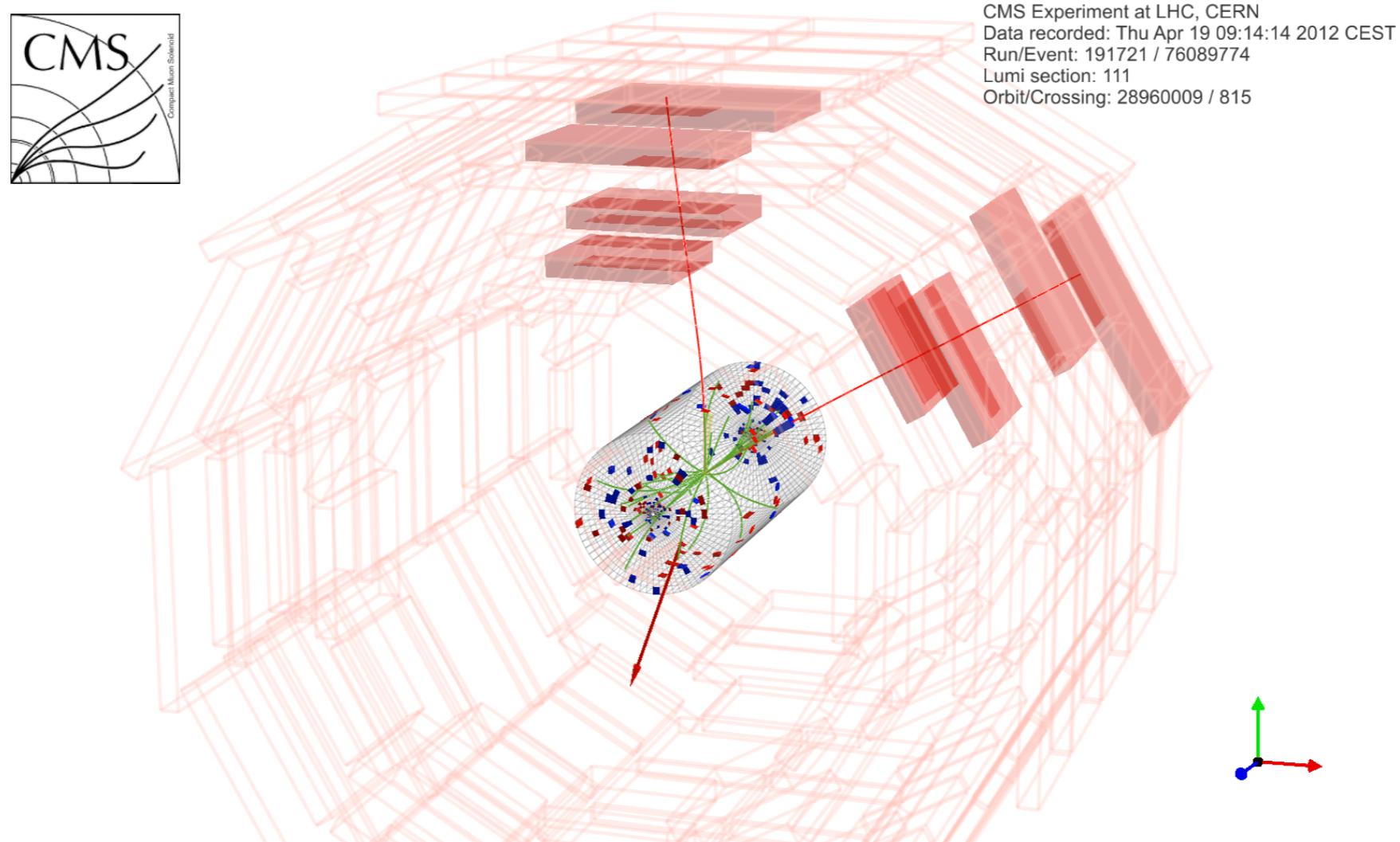
- The energy deposits of electromagnetic or hadronic showers form clusters.



- Seeded by calorimeter cells featuring local energy maxima.
- Clustering algorithms search for adjacent cells with energies above a noise threshold to propose topological clusters, which themselves gathered into superclusters.

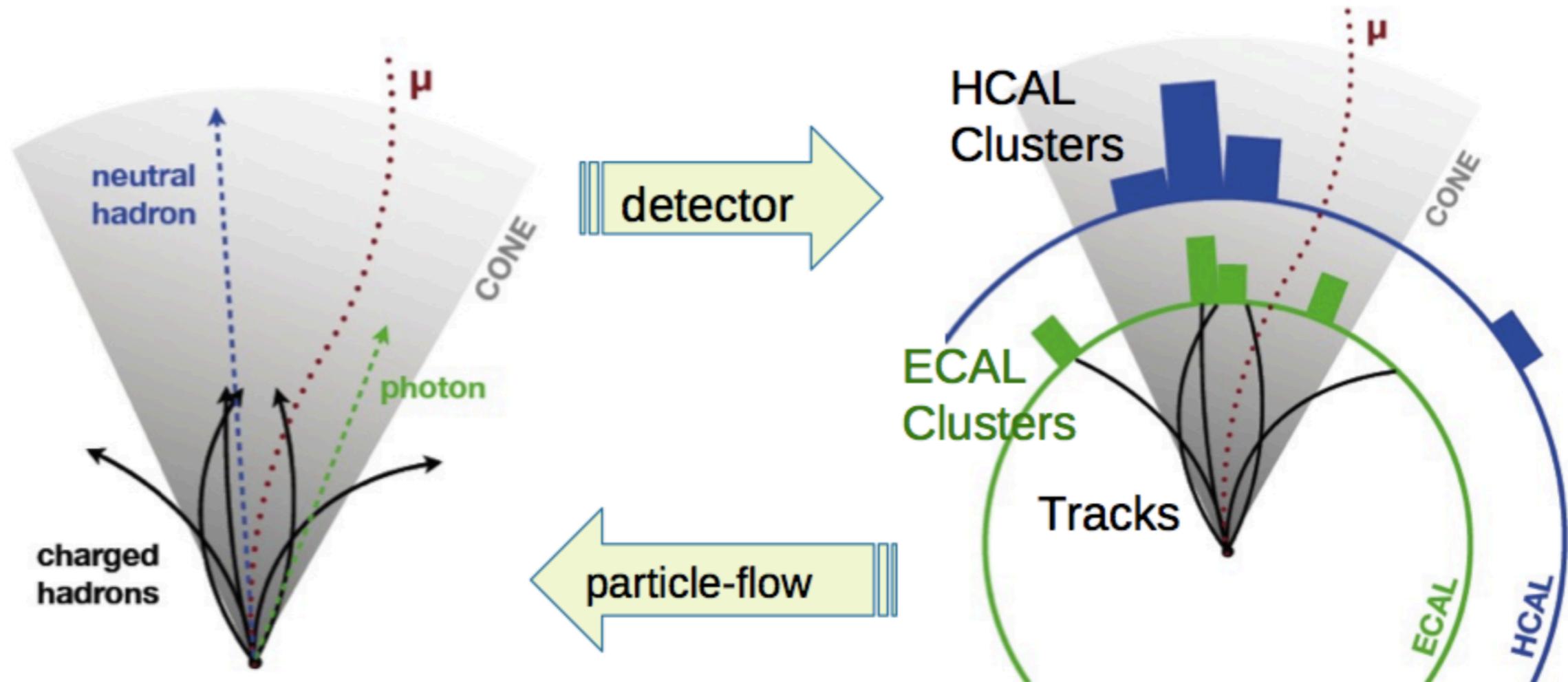
Standalone Muon Tracks

- The trajectories of muons are also reconstructed from their signals.



- Seeded by the inner-most track segments reconstructed by a DT or CSC.
- Extrapolation to and from compatible track segments and hits found in outer layers.

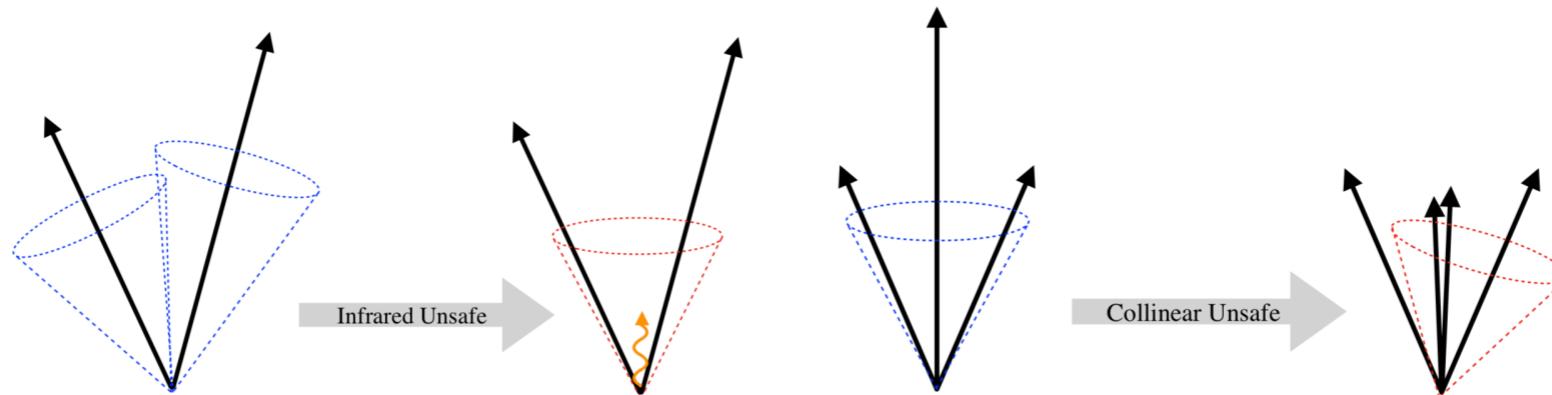
Particle Flow Reconstruction



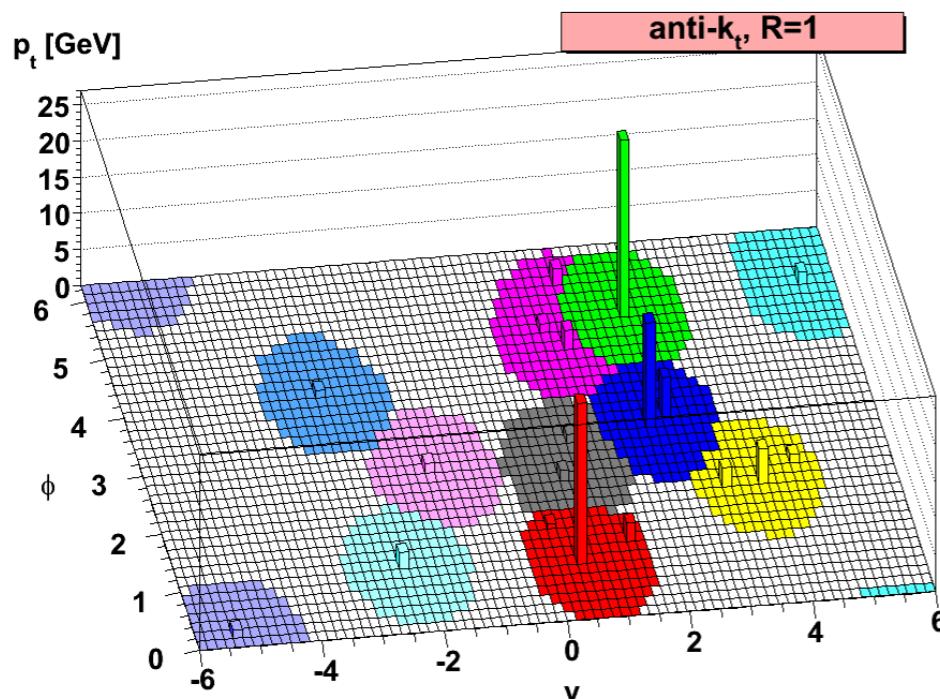
1. A linking algorithm creates blocks of compatible elements (detector-level objects) across detector subsystems.
2. Particle candidates are identified and reconstructed within each block: muons, electrons, photons, and charged and neutral hadrons.
3. Proposed candidates are ready to use in analyses, but additional post-processing is typical.

Jet Clustering

- Experimental signature of quarks and gluons are called jets.
- Reconstruction of jets must be handled carefully.



- PF candidates clustered using the anti- k_T sequential recombination algorithm.

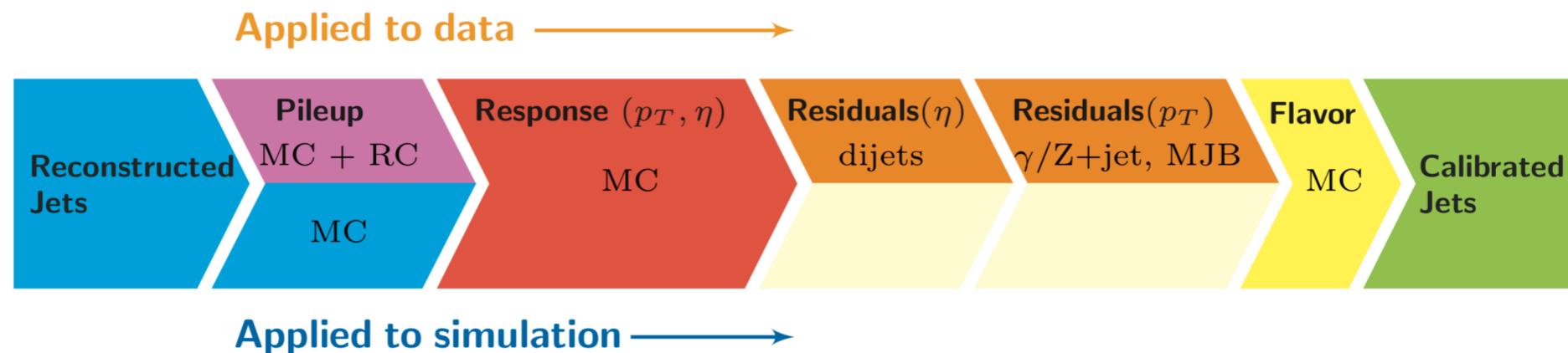


$$d_{ij} = \min \left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2} \right) \times \frac{R_{ij}^2}{R} \quad d_{iB} = \frac{1}{p_{Ti}^2}$$

$$R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$$

Jet Energy Scale and Resolution

- Constituent particles encounter energy losses traveling through detector material and their measured energies may be biased by non-linear detector response.
- Calibrate jet energies by applying jet energy scale corrections.

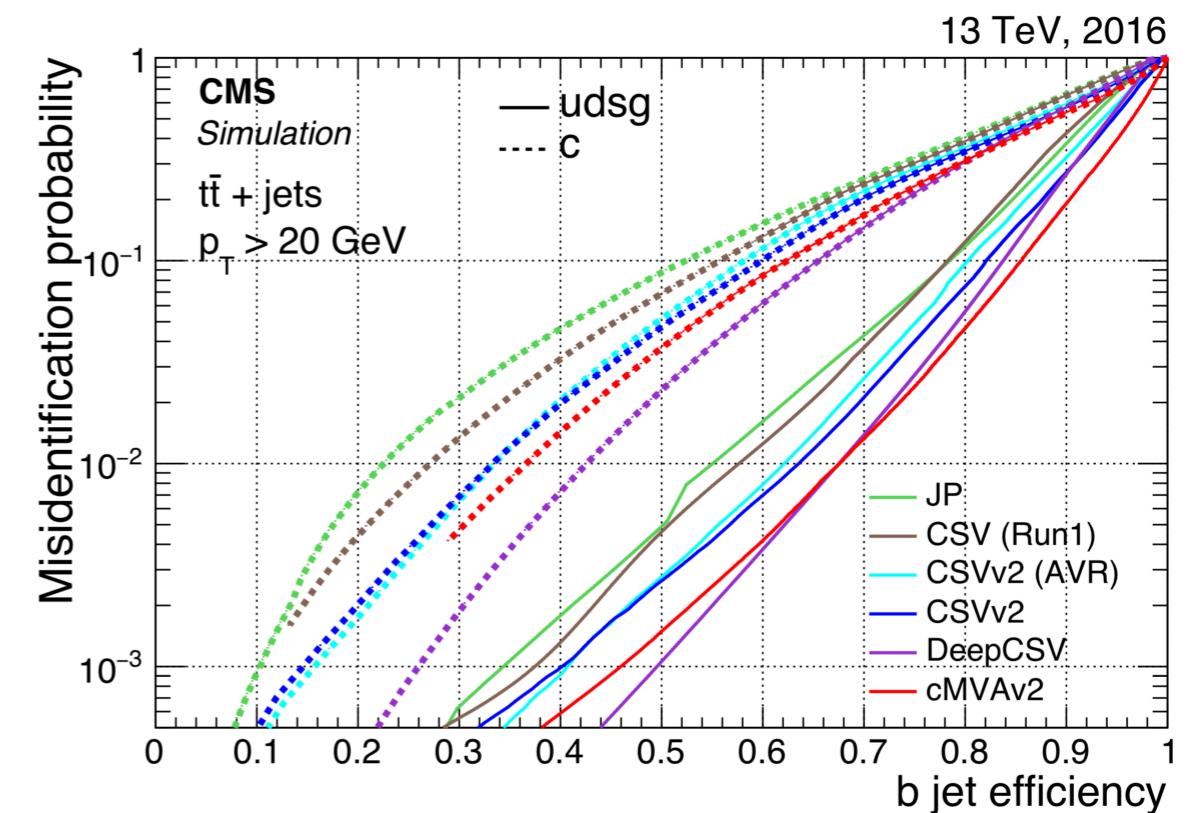
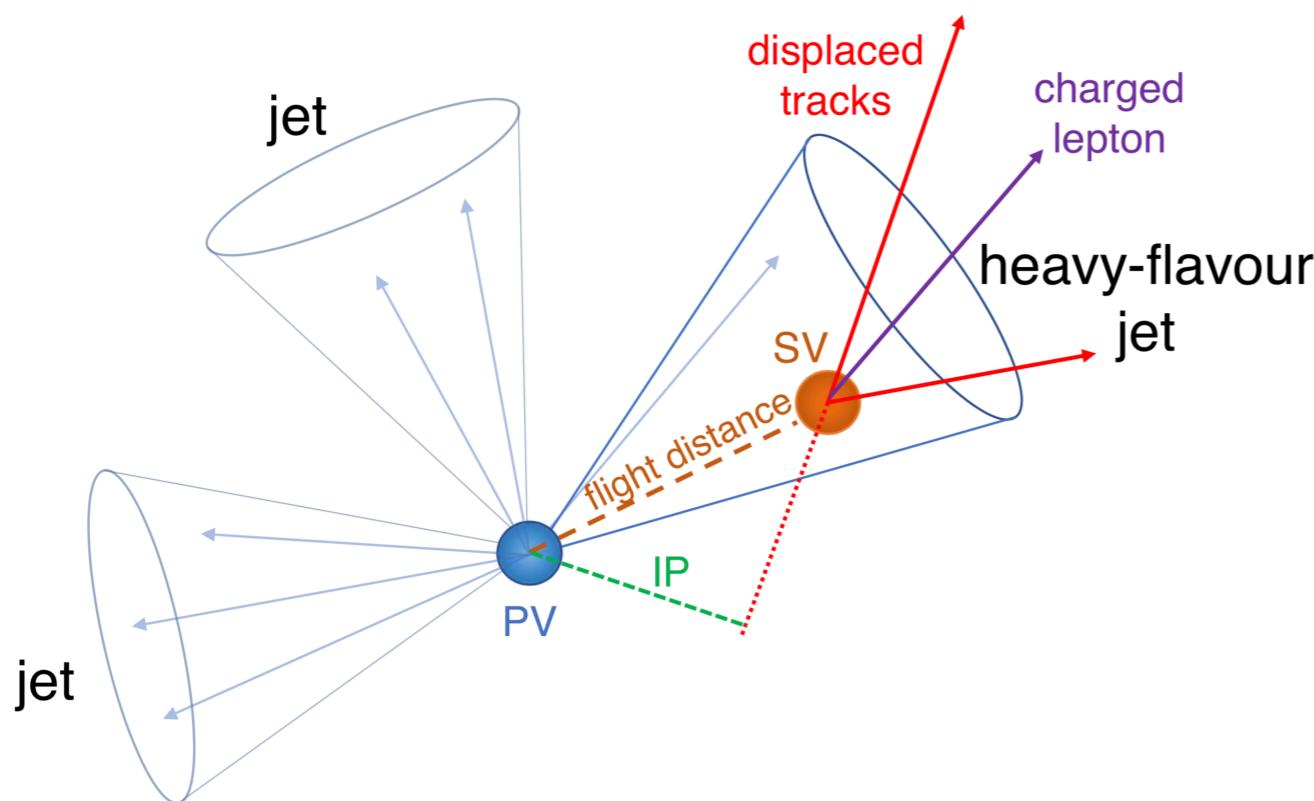


- After the energy calibration, the jet energy resolution is better in simulation than data.
- Stochastically “smear” the reconstructed jets in simulation.

$$c_{\text{JER}} = 1 + \mathcal{N}(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)}$$

b-Tagging

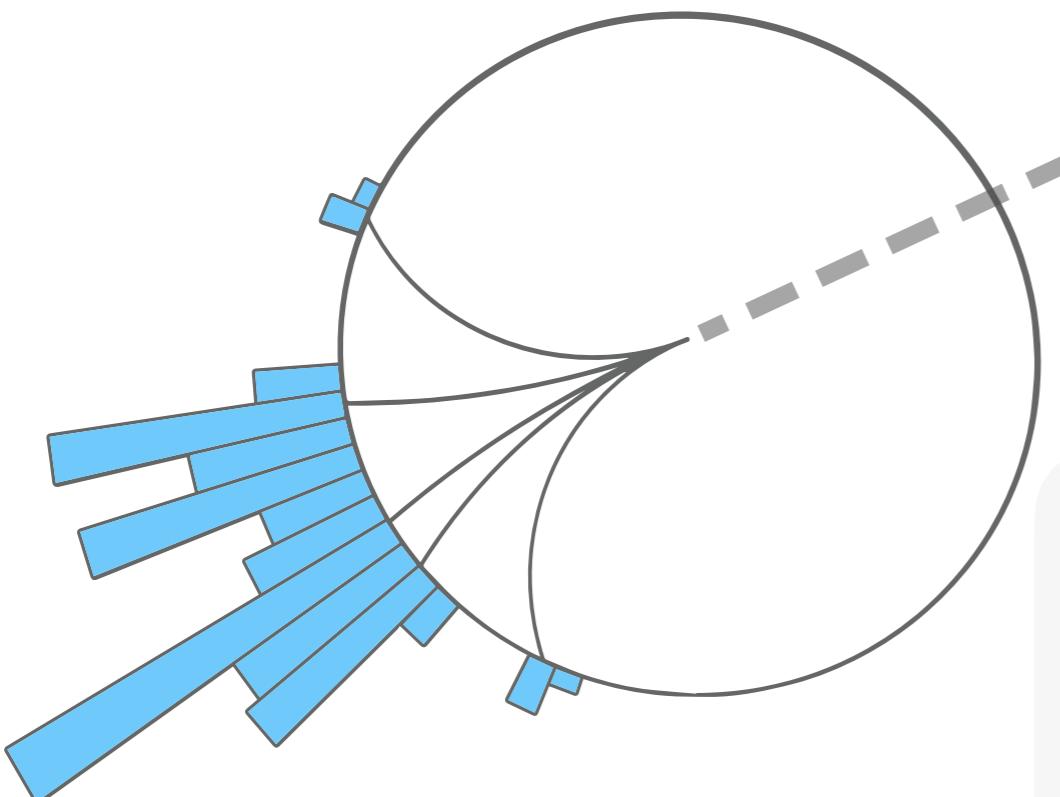
- The decay of long-lived *b*-hadrons are displaced from the primary vertex.



- Artificial neural network based DeepCSV tagger uses information from secondary vertices and their associated tracks to classify the flavor of jets.
- Achieves a *b*-jet identification efficiency of 84% with a misidentification rate of 11%.

Missing Transverse Energy

- Neutrinos, being electrically neutral and weakly interacting, escape detection.
- Infer their presence based on an imbalance in the transverse momentum sum.



$$\vec{p}_T^{\text{miss}} = - \sum \vec{p}_T,$$

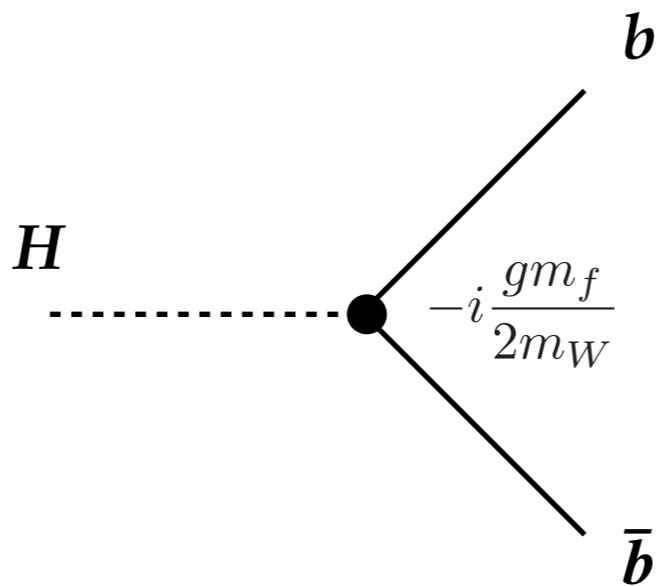
$$\vec{p}_T^{\text{miss}} = - \sum_{\text{jets}} \vec{p}_T^{\text{JEC}} - \sum_{\text{unclustered}} \vec{p}_T$$

Type-I Corrected

- Sources of spurious noise or mismeasurement are explicitly filtered based on recommendations from MET data quality monitoring experts.

Motivation for VH($b\bar{b}$) Searches

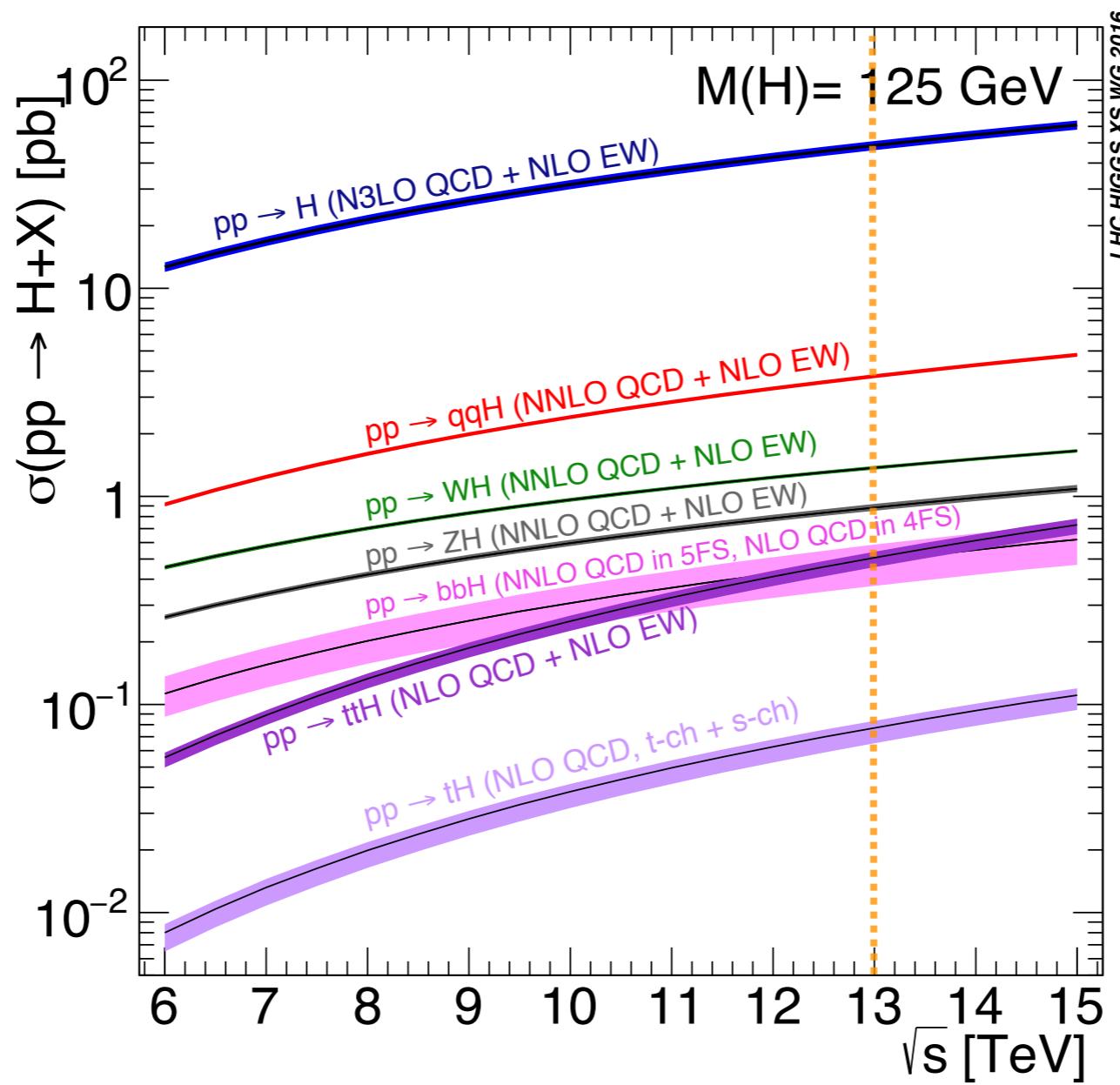
The $H \rightarrow b\bar{b}$ Decay



- Has the largest branching fraction at $\approx 58\%$.
- Opportunity to observe and measure the Higgs coupling to down-type quarks.
- Precise measurement of the branching fraction improves constraints on the total decay width.

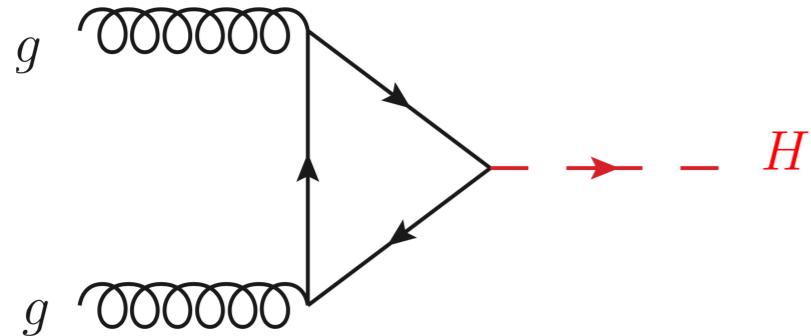
Higgs Boson Production

The Higgs boson is produced at the LHC through one of the following processes...

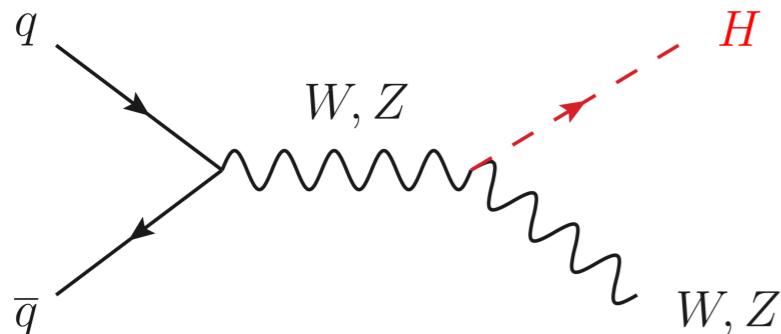


Motivation for VH($b\bar{b}$) Searches

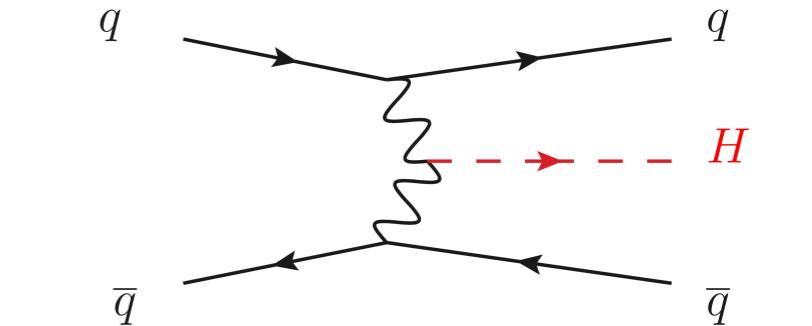
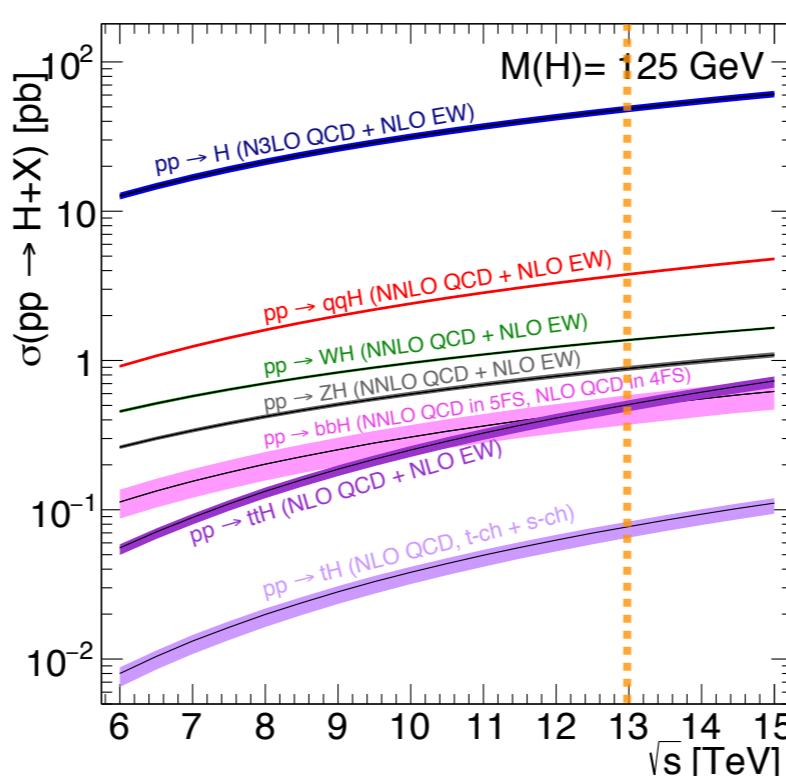
Higgs Boson Production



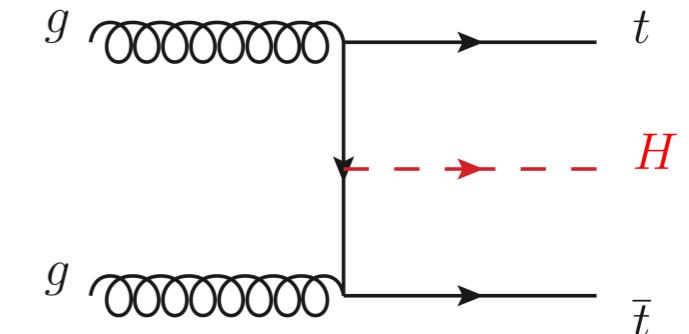
Gluon Fusion ($\approx 87\%$)



Vector Boson Associated Production ($\approx 5\%$)



Vector Boson Fusion ($\approx 7\%$)

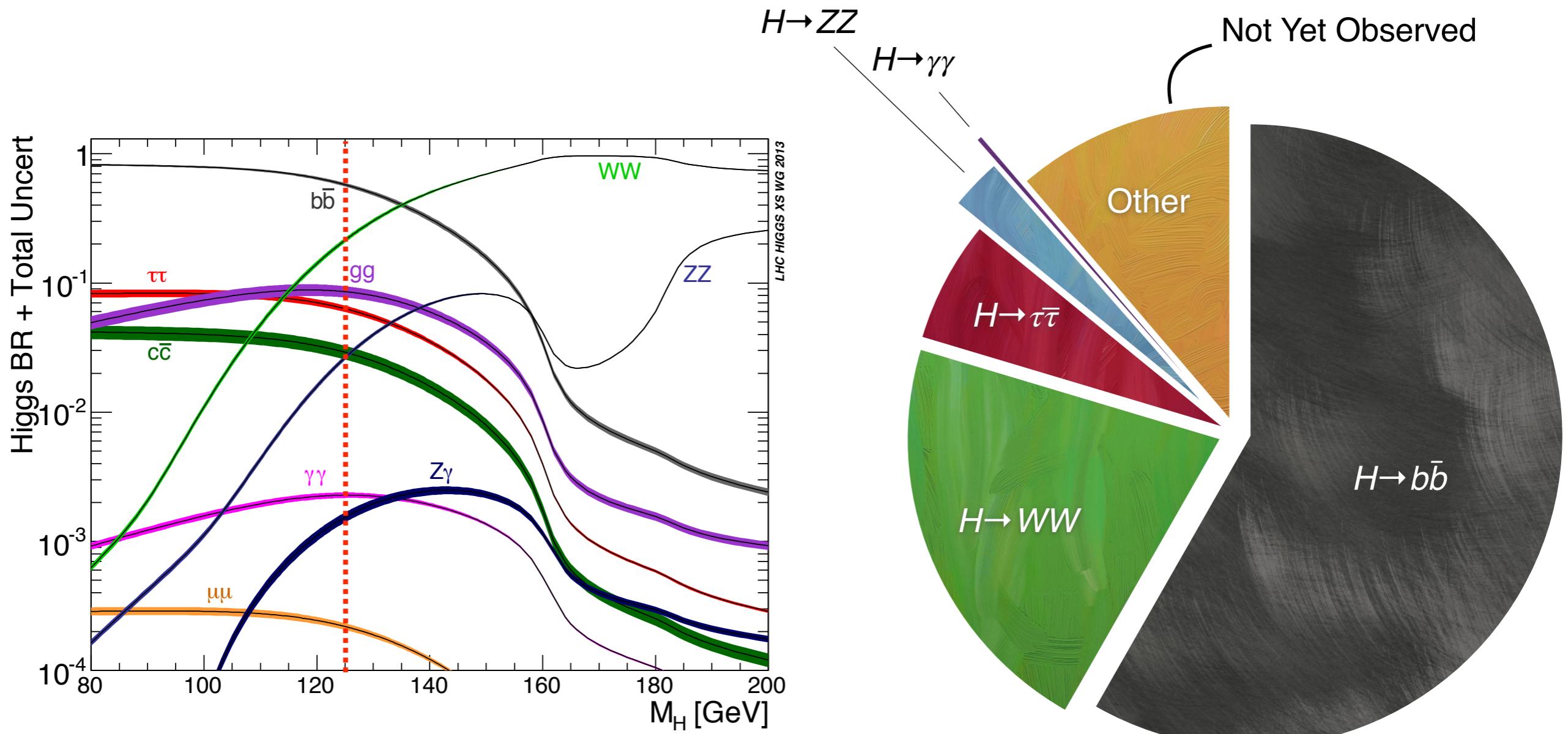


$t\bar{t}$ Associated Production ($\approx 1\%$)

Motivation for VH($b\bar{b}$) Searches

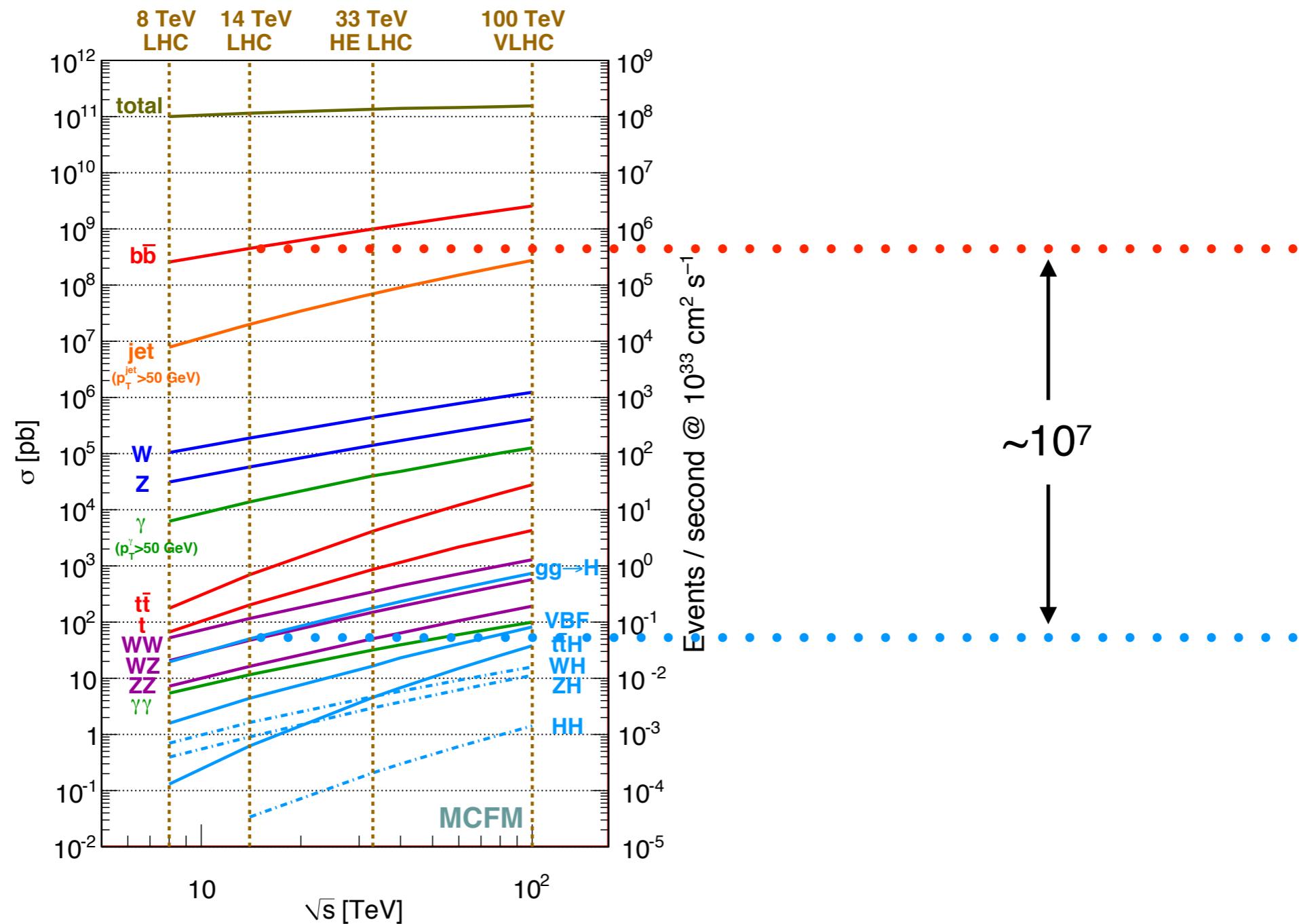
Higgs Boson Decay

... and, being short-lived, are only identified by their decay products...



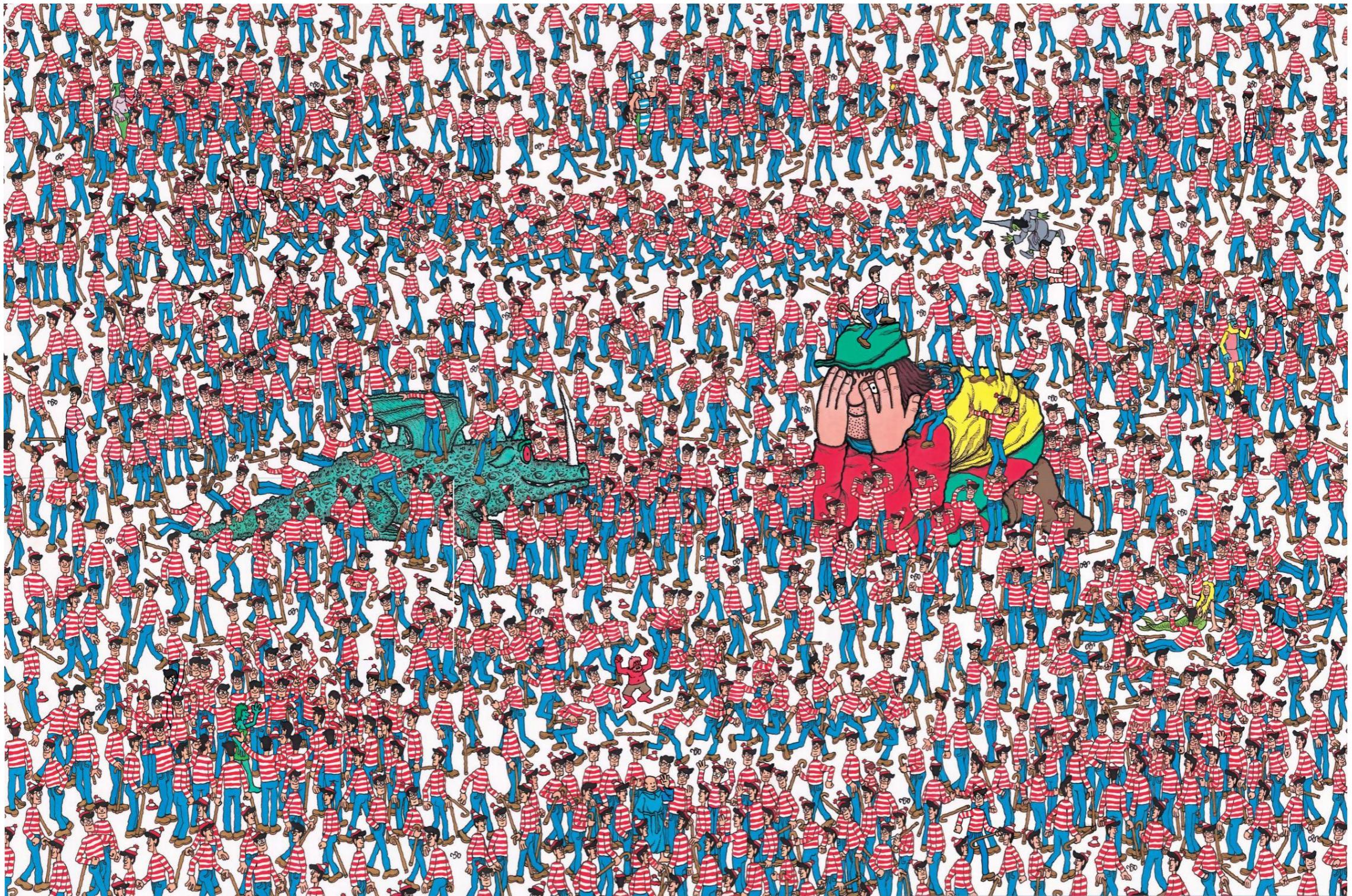
Higgs Boson Decay

... which can be difficult given the variety of background processes.



Motivation for $VH(b\bar{b})$ Searches

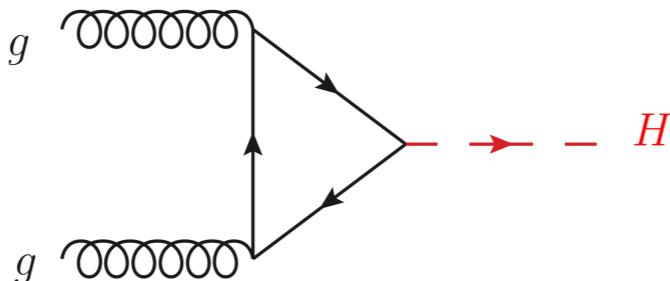
Higgs Boson Decay



Motivation for VH($b\bar{b}$) Searches

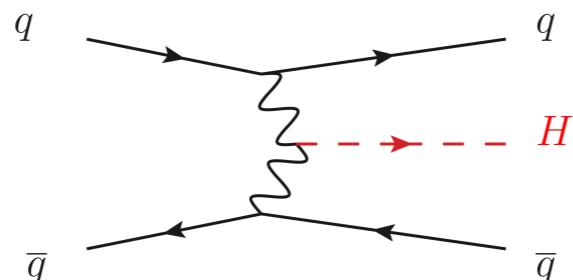
$H \rightarrow b\bar{b}$ Searches at the LHC

Gluon Fusion



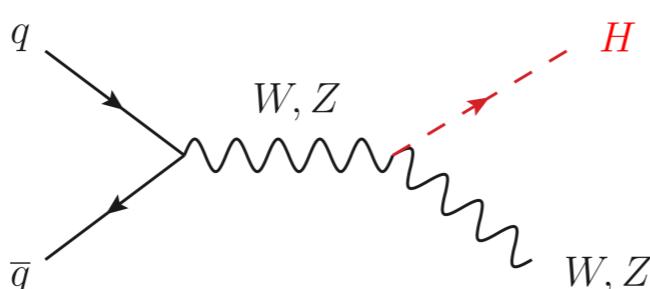
Suffers from the large multijet background, but advances in jet substructure are helping to overcome that.

VBF



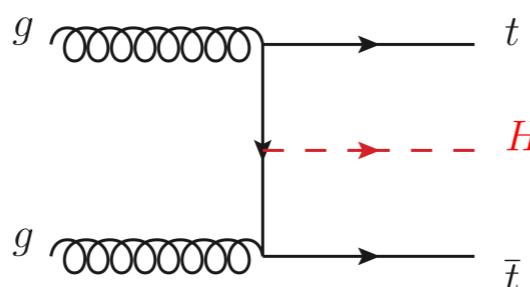
Also contends with the large multijet background, but its fully hadronic final state offers a challenging trigger topology.

VH



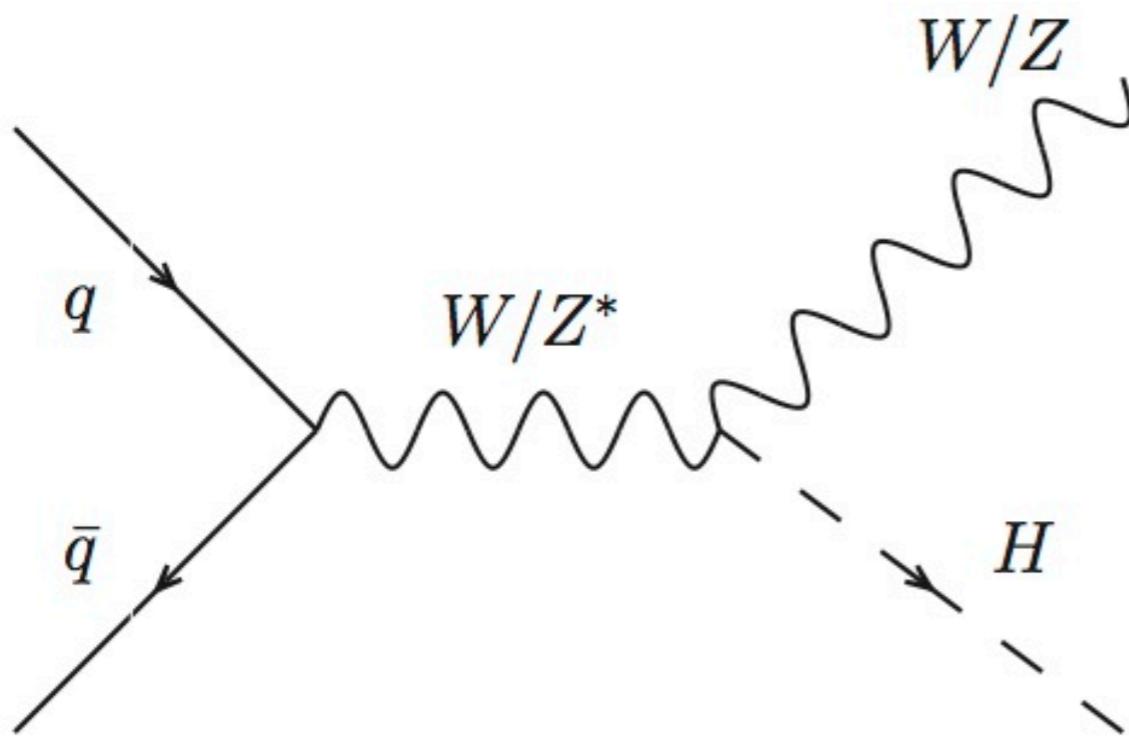
The vector boson's leptonic decays provide an easy trigger strategy and help suppress the multijet background.

ttH



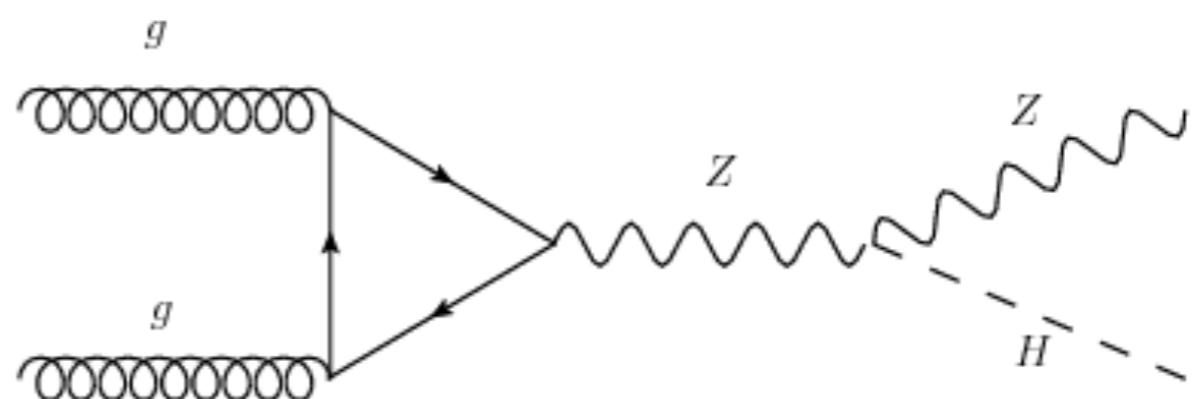
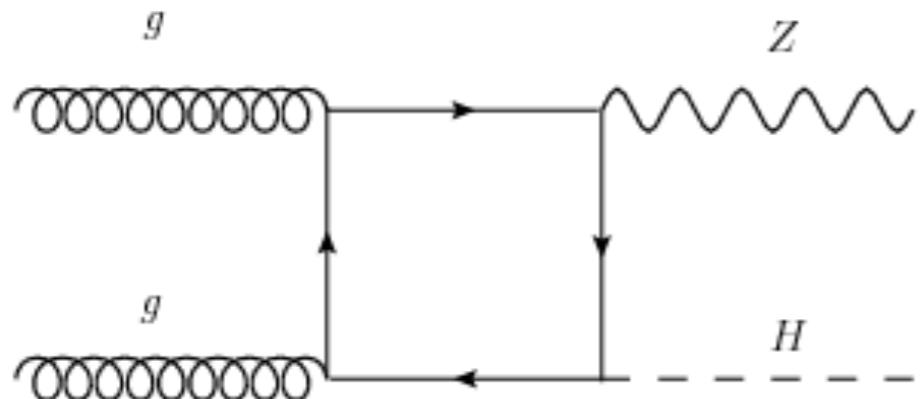
Its final state offers leptons and jets for triggering but offers a combinatorial challenge while dealing with $t\bar{t}+j$ ets.

The VH($b\bar{b}$) Decay



- The leptonic decays of the vector boson
 - suppress the multijet background
 - provide simple trigger topologies
- By far the most sensitive channel in the search for $H \rightarrow b\bar{b}$.

Higher order contributions from ggZH ($\approx 10\%$).



Previous LHC Results

- Feasibility of searching for VH($b\bar{b}$) established during LHC Run 1.

Run 1	Expected Significance	Observed Significance	Reference
ATLAS	2.6	1.4	JHEP01 (2015) 069
CMS	2.1	2.1	Phys. Rev. D 89, 012003
ATLAS + CMS	3.7	2.6	JHEP08 (2016) 045

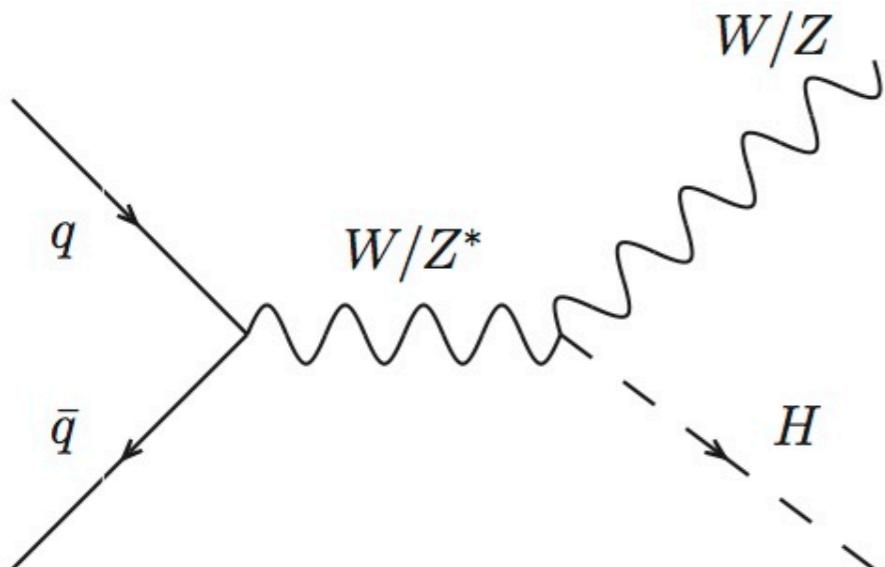
- Evidence of the VH($b\bar{b}$) decay established in 2016 during LHC Run 2.

Run 1 + 2 (up to 2016)	Expected Significance	Observed Significance	Reference
ATLAS	4.0	3.6	JHEP12 (2017) 024
CMS	3.8	3.8	Phys. Lett. B 780 (2018) 501

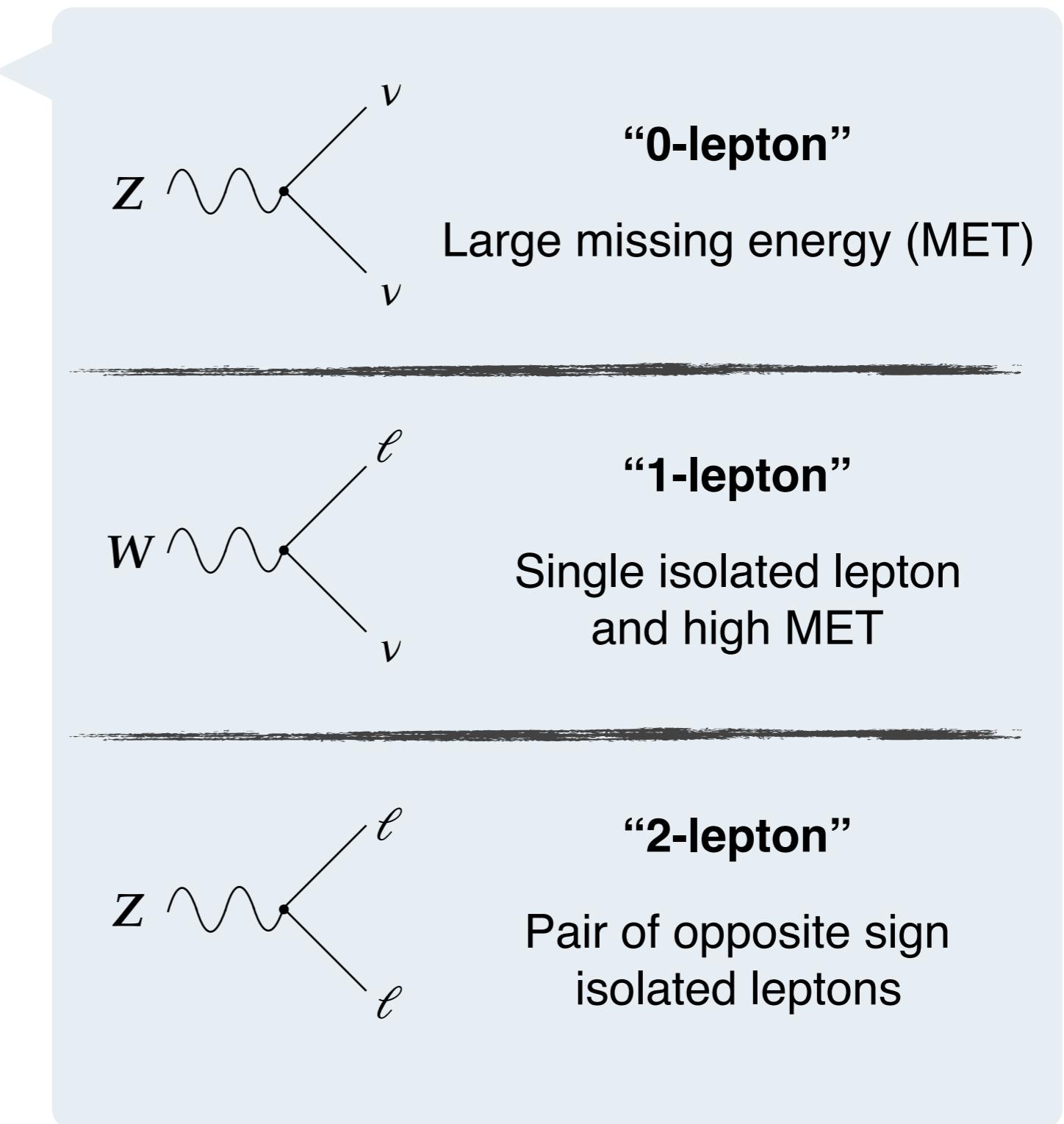
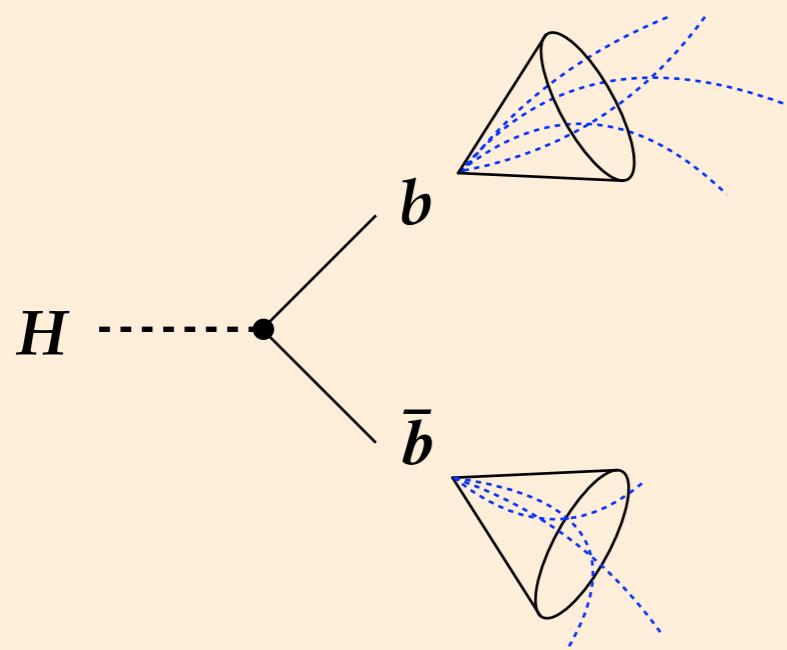
- With LHC Run 2 concluding in 2017, another year's worth of data may be enough to achieve an observation of the VH($b\bar{b}$) decay.

2017 VH($b\bar{b}$) Analysis

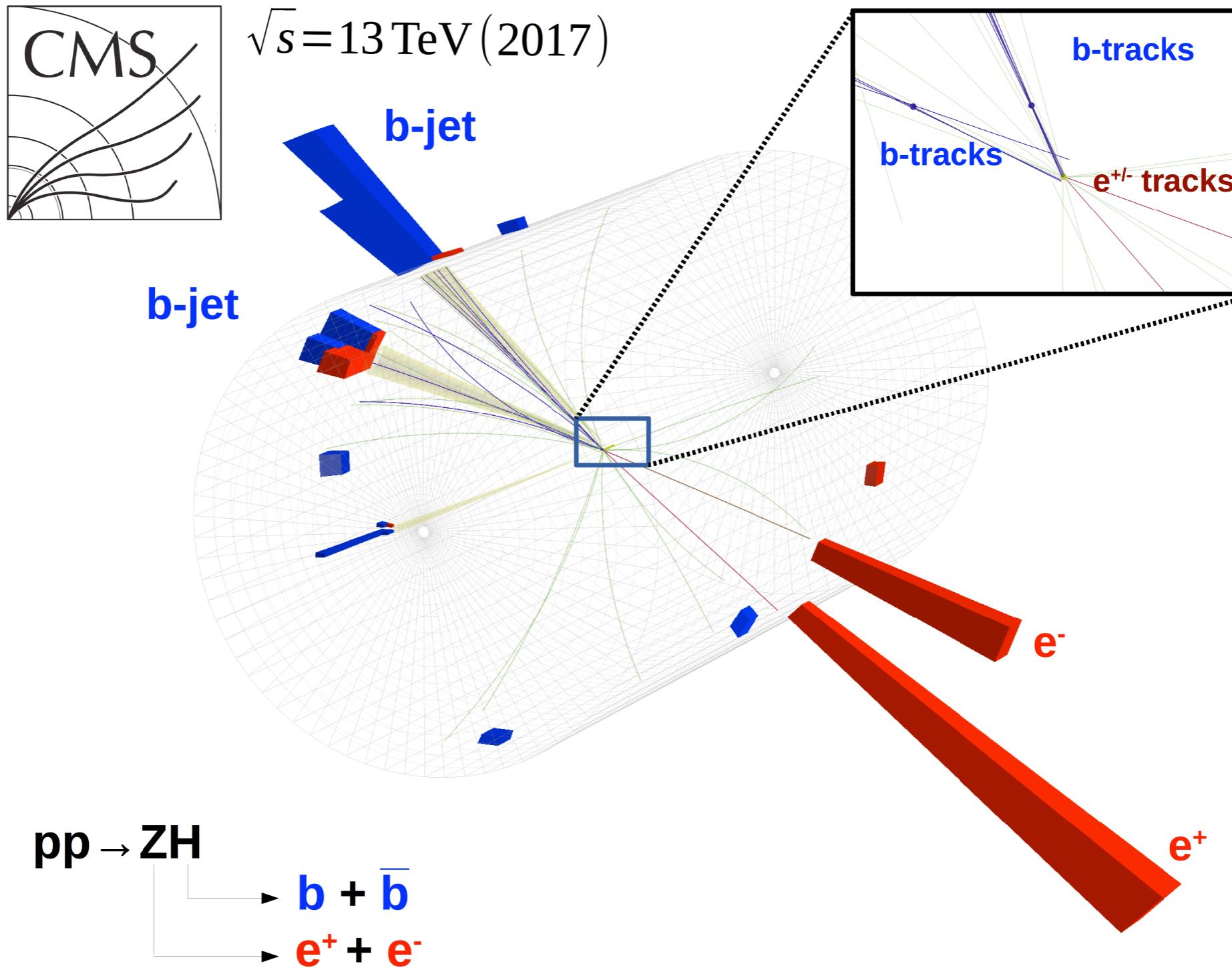
Final State Reconstruction



Pair of b -tagged anti- k_T jets ($R = 0.4$)

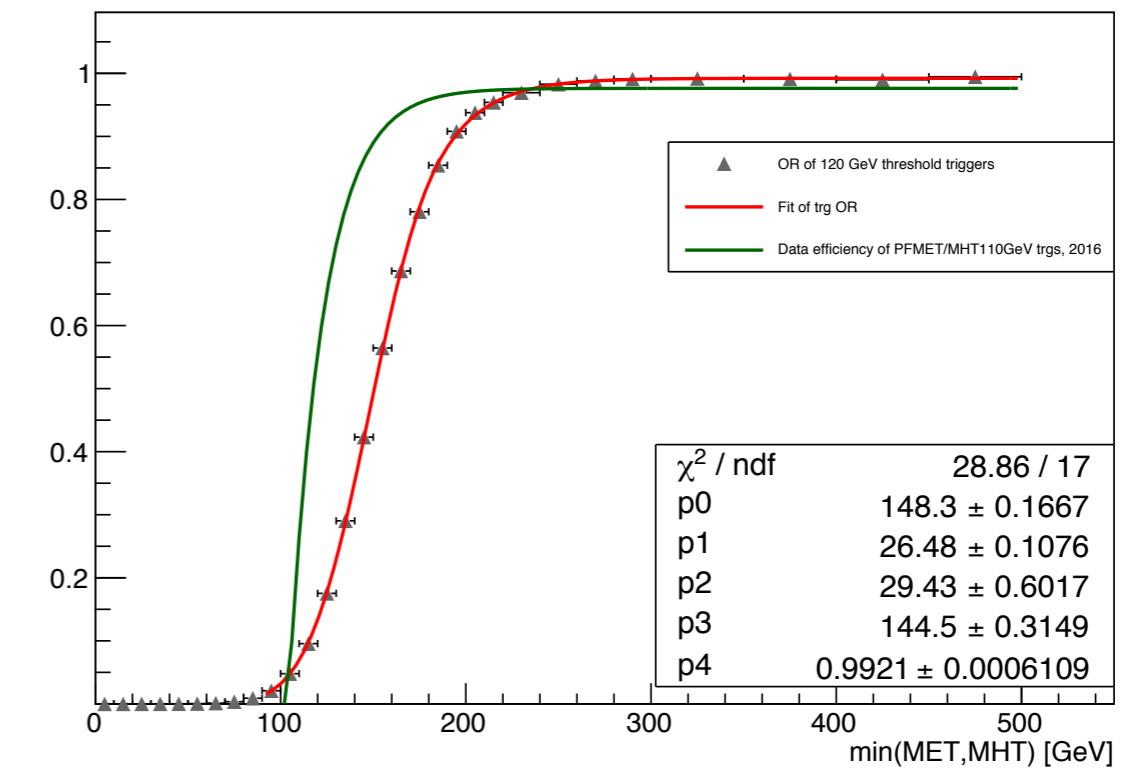
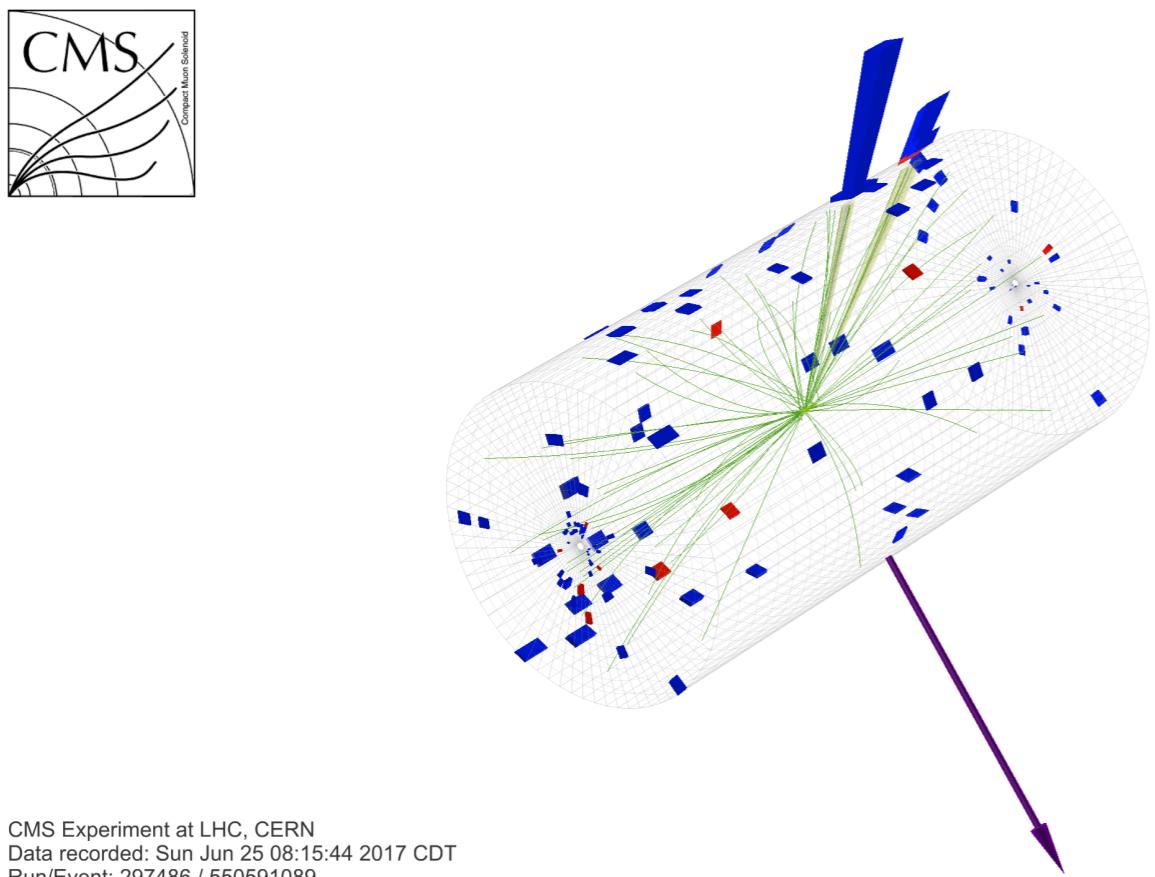


Final State Reconstruction



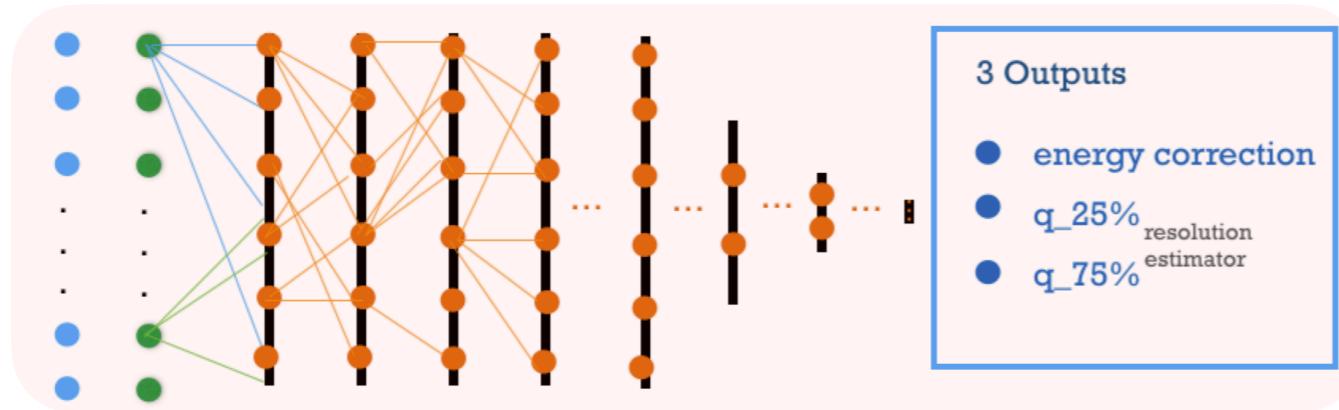
Final State Reconstruction: 0-lepton

- Data events selected from MET dataset using the logical *OR* of two triggers:
 - HLT_PFMET120_PFMHT120_IDTight
 - HLT_PFMET120_PFMHT120_IDTight_PFHT60
- min(MET,MHT) threshold raised to 170 GeV for analysis.

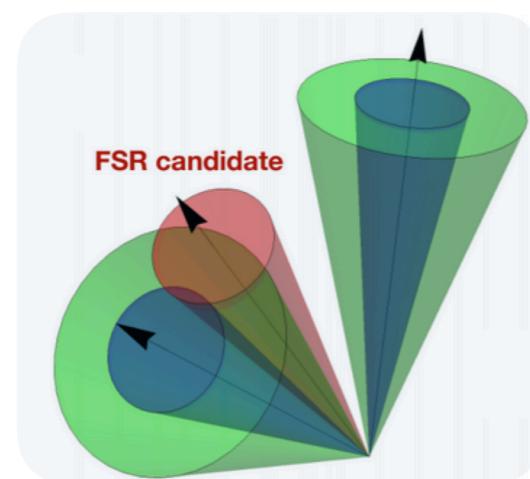


Final State Reconstruction

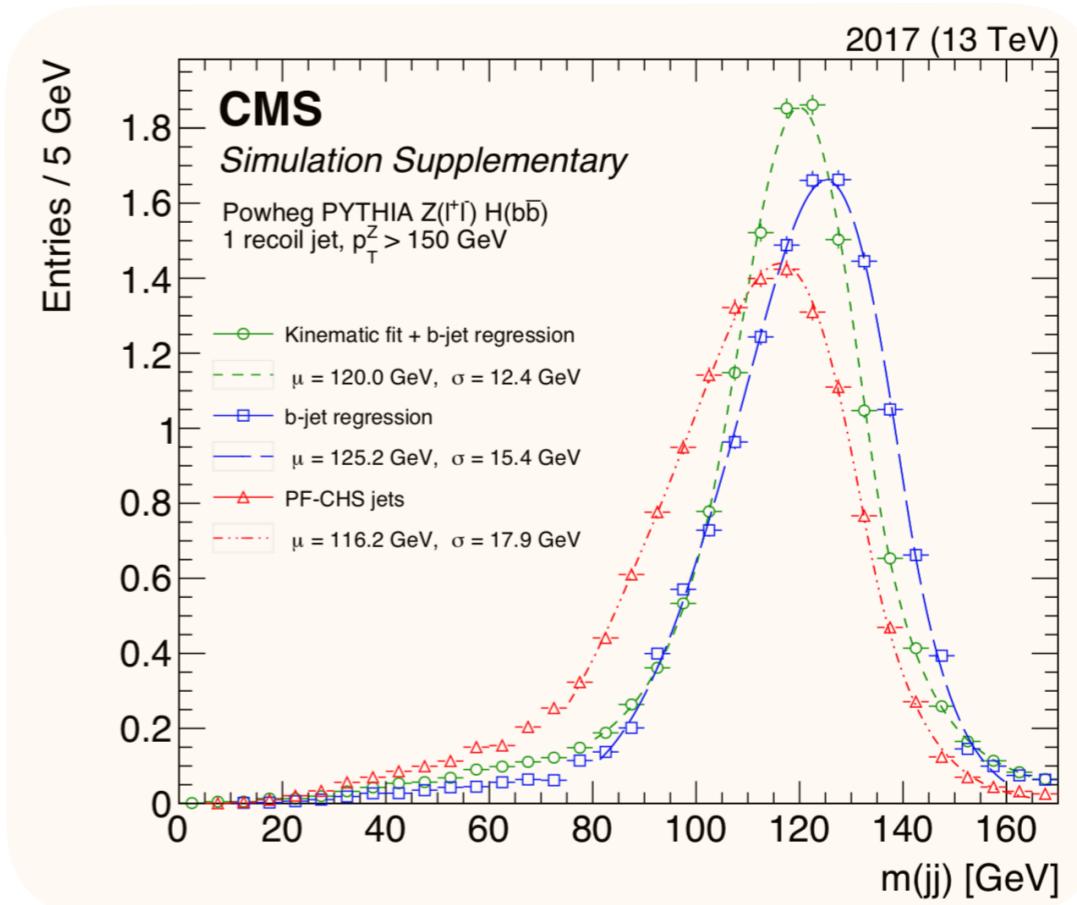
- Good dijet invariant mass resolution is crucial to the analysis.
- Correct b -jet energies using a **neural network regression** model.
- Kinematic fit** exclusive to the 2-lepton channel.



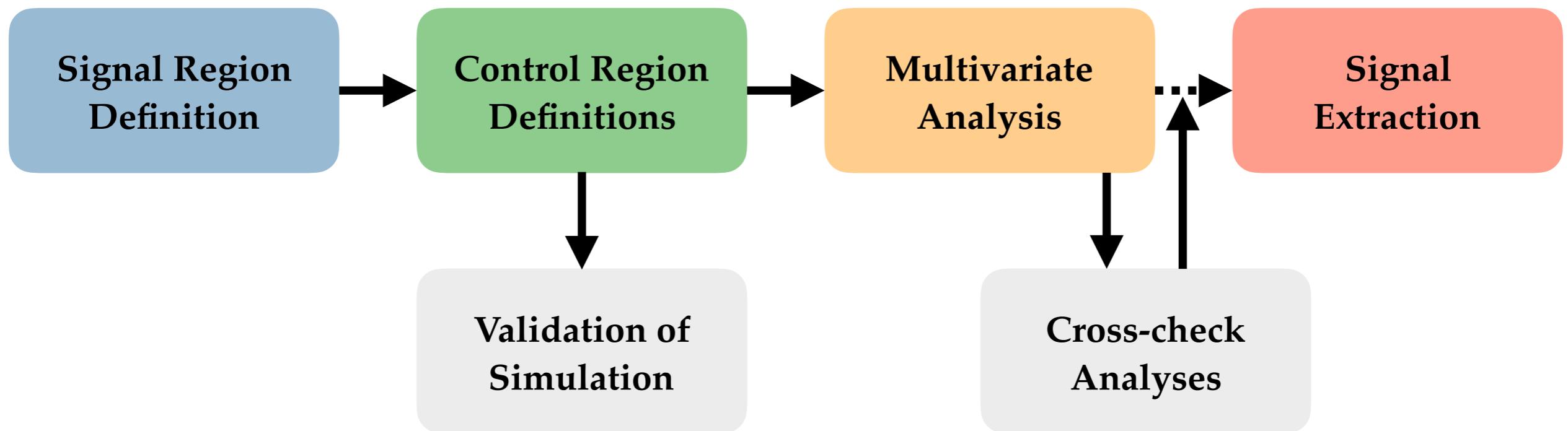
- Recovery of **final state radiation**.



- Overall improvement of $\approx 10\text{-}23\%$, depending on the decay channel.



Analysis Strategy



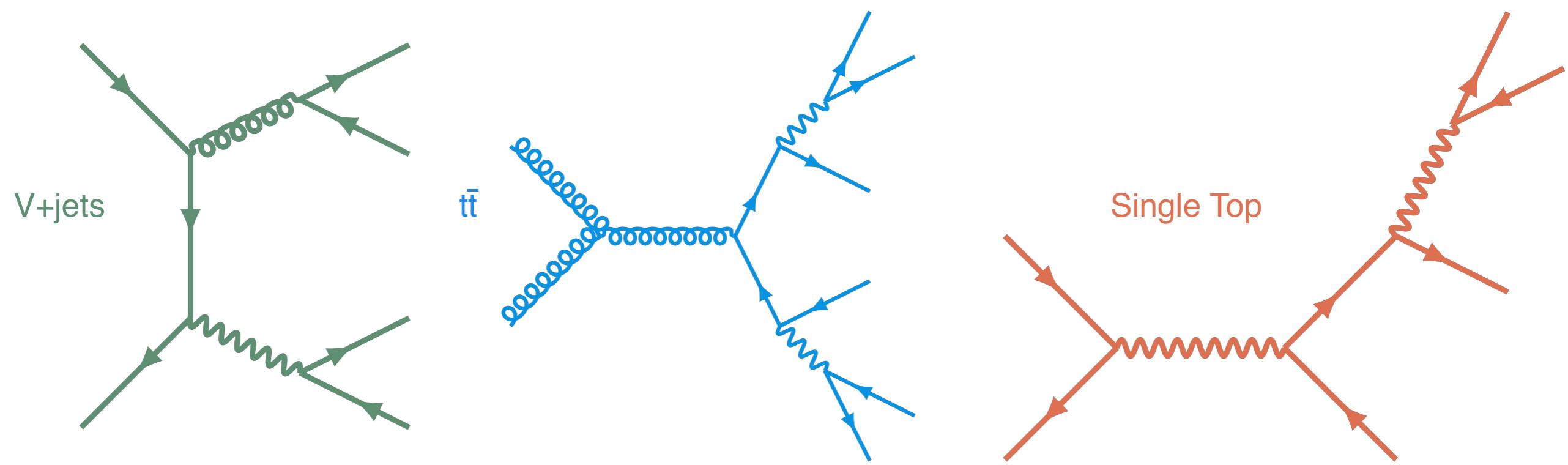
Signal Region

For each of the decay channels, a signal region is defined with cuts that reduce the background contamination.

	Variable	$Z(\nu\bar{\nu})H$	$W(\ell\nu)H$	$Z(\ell\ell)H$
Vector Boson	$p_T(V)$	> 170	> 150	$[50, 150], > 150$
	$m(\ell\ell)$	—	—	$[75, 105]$
	p_T^ℓ	—	$> 25, > 30$	> 20
Higgs Boson	$p_T(jj)$	> 120	> 100	—
	$m(jj)$	$[60, 160]$	$[90, 150]$	$[90, 150]$
	$p_T(j)_{\max}$	> 60	> 25	> 20
	$p_T(j)_{\min}$	> 35	> 25	> 20
	$b\text{-tag}_{\max}$	$>$ Tight	$>$ Tight	$>$ Loose
	$b\text{-tag}_{\min}$	$>$ Loose	$>$ Loose	$>$ Loose
Final State Topology	N_j^{add}	< 2	< 2	—
	N_ℓ^{add}	$= 0$	$= 0$	—
	p_T^{miss}	> 170	—	—
	Anti-QCD	Yes	—	—
	$ \Delta\phi(V, H) $	> 2.0	> 2.5	> 2.5
	$ \Delta\phi(p_T^{\text{miss}}, p_T^{\text{miss,track}}) $	< 0.5	—	—
	$ \Delta\phi(p_T^{\text{miss}}, \ell) $	—	< 2.0	—
	Lepton Isolation	—	$0.06, 0.06$	—

Control Regions

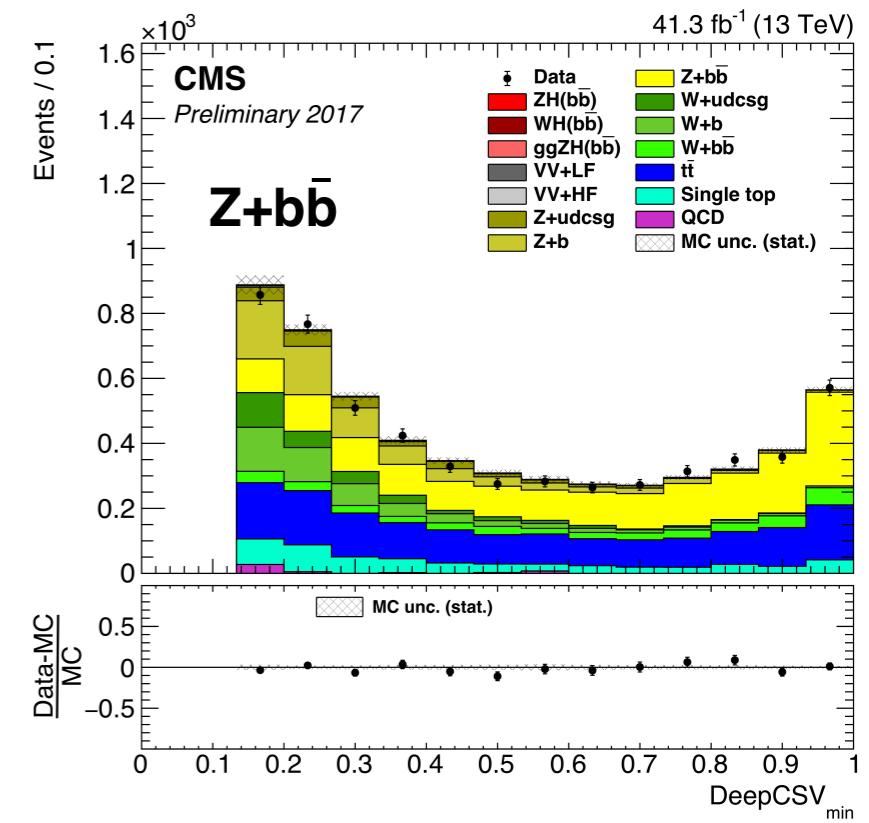
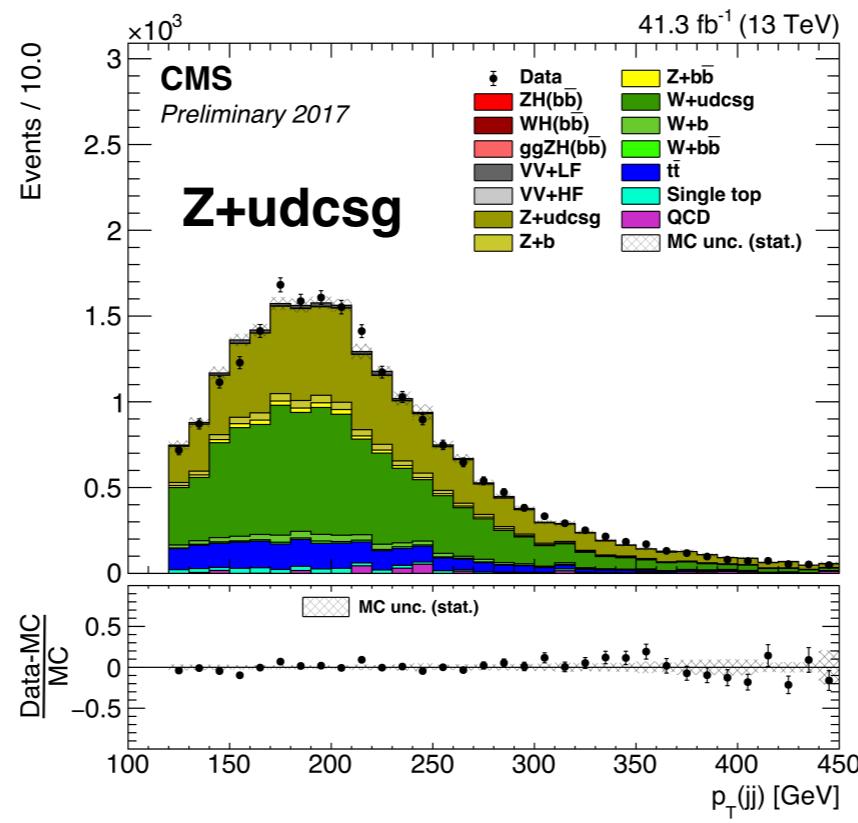
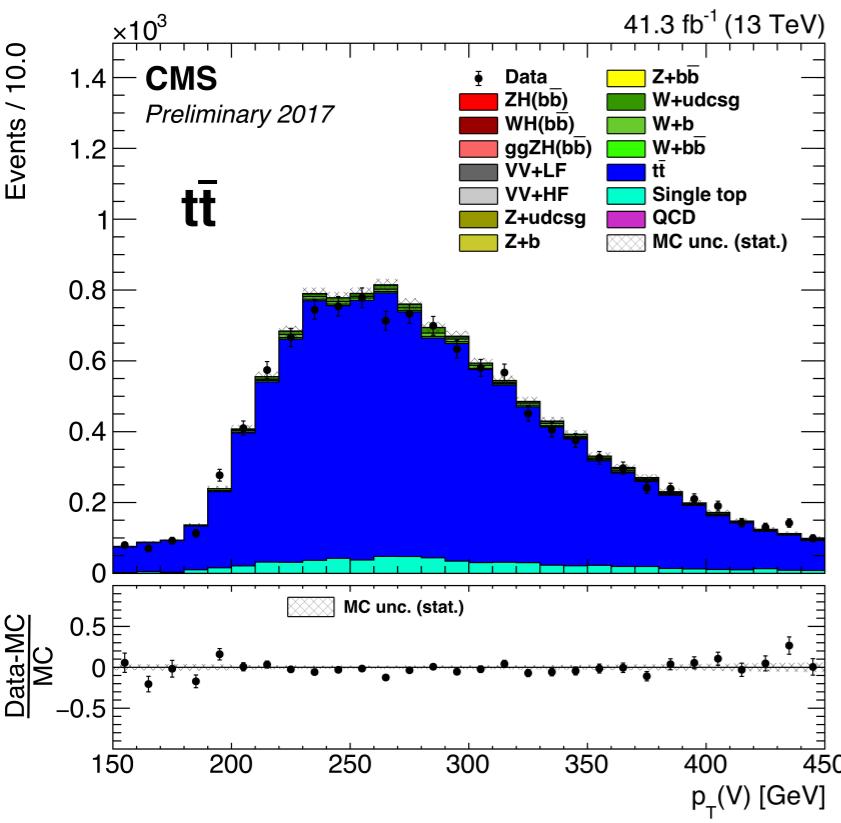
- The final state contends with large and irreducible backgrounds.



- Define control regions that are enriched in specific backgrounds and orthogonal to the signal region and each other.
 - Necessary to validate key distributions used by the multivariate analysis.
 - Helps to constrain the background normalizations during signal extraction.

Control Regions: 0-lepton

	Variable	$t\bar{t}$	$Z + \text{light}$	$Z + \text{heavy}$
Vector Boson	V Decay Type	$W(\mu\nu)$	$Z(\nu\bar{\nu})$	$Z(\nu\bar{\nu})$
	p_T^{miss}	> 170	> 170	> 170
Higgs Boson	$m(jj)$	—	—	$\notin [60, 160]$
	$p_T(jj)$	> 120	> 120	> 120
	$b\text{-tag}_{\max}$	$>$ Medium	$<$ Medium	$>$ Tight
	$b\text{-tag}_{\min}$	$>$ Loose	$>$ Loose	$>$ Loose
	$p_T(j)_{\max}$	> 60	> 60	> 60
	$p_T(j)_{\min}$	> 35	> 35	> 35
Final State Topology	$ \Delta\phi(V, H) $	> 2.0	> 2.0	> 2.0
	N_{ℓ}^{add}	≥ 1	$= 0$	$= 0$
	N_j^{add}	≥ 2	≤ 1	≤ 1
	$ \Delta\phi(p_T^{\text{miss}}, p_T^{\text{miss,track}}) $	—	< 0.5	< 0.5
	Anti-QCD	Modified	Yes	Yes

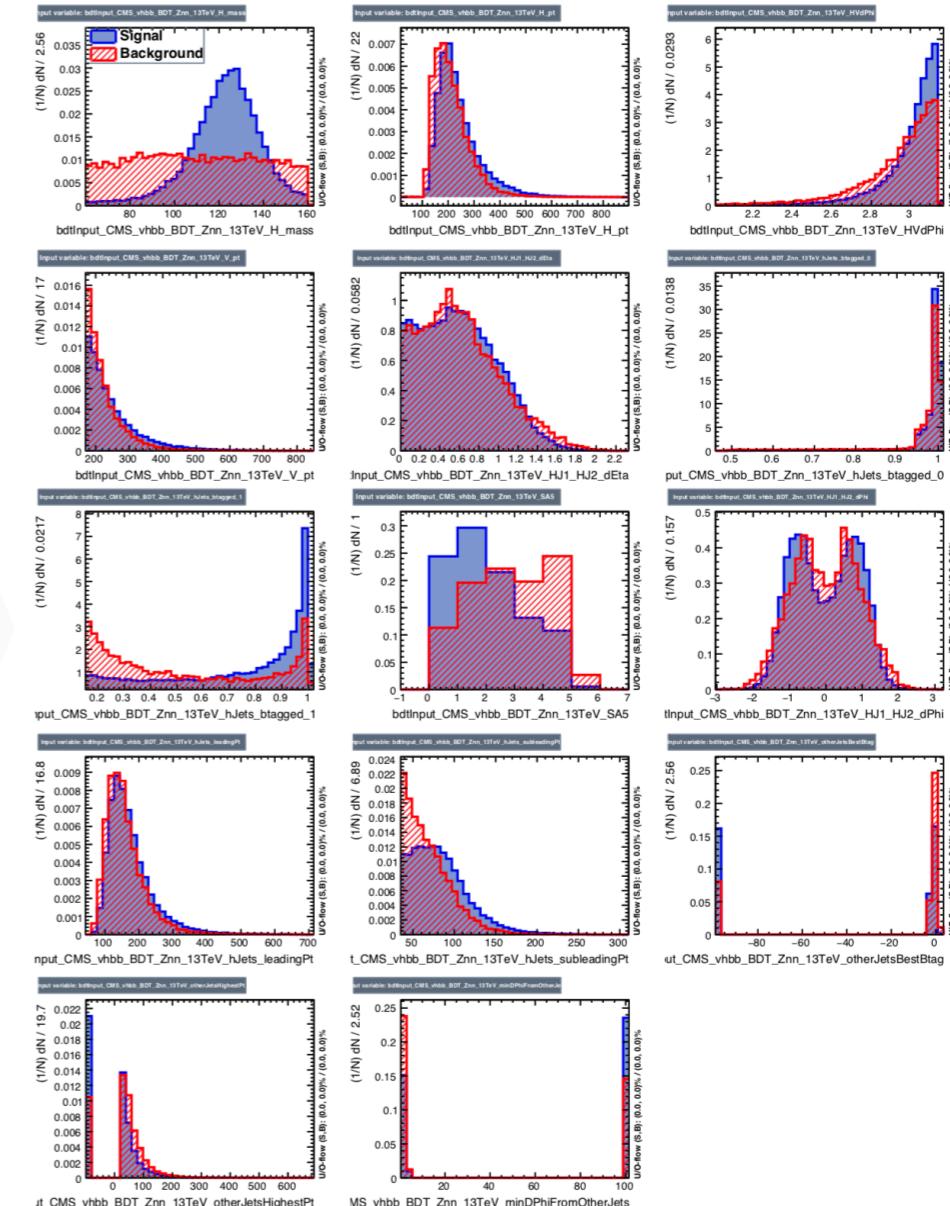


2017 VH($b\bar{b}$) Analysis

Multivariate Analysis

- The dijet invariant mass distribution is the most discriminating individual variable.
- The signal-to-background ratio may be further enhanced by considering information from other features of the event.

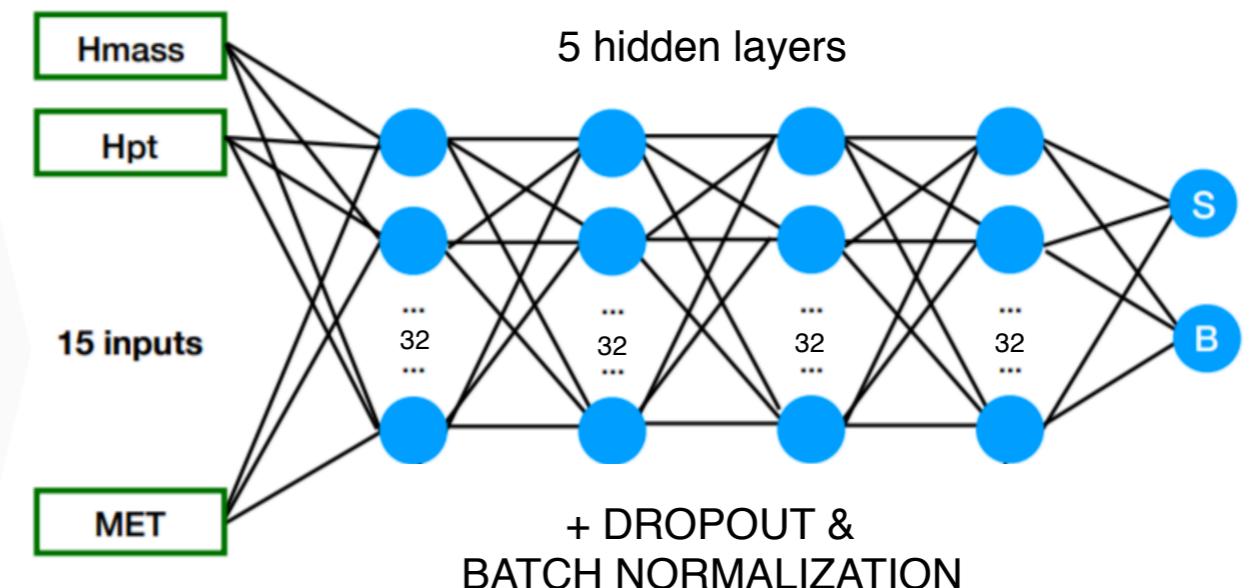
Variable	Description	$Z(\nu\bar{\nu})H$	$W(\ell\nu)H$	$Z(\ell\bar{\ell})H$
$m(jj)$	dijet invariant mass	✓	✓	✓
$p_T(jj)$	dijet transverse momentum	✓	✓	✓
$p_T(j)_\text{max}, p_T(j)_\text{min}$	transverse momentum of the two b -jets	✓		✓
$\Delta R(jj)$	distance between the two b -jets in η - ϕ space			✓
$\Delta\eta(jj)$	difference in η between the two b -jets	✓		✓
$\Delta\phi(jj)$	difference in ϕ between the two b -jets	✓		✓
$p_T(V)$	vector boson transverse momentum		✓	✓
$\Delta\phi(V, H)$	difference in ϕ between the dijet system and vector boson	✓	✓	✓
$p_T(jj)/p_T(V)$	p_T ratio of the dijet system to the vector boson			✓
$m(Z)$	reconstructed mass of the Z boson			✓
$b\text{-tag}_\text{max}$	leading DeepCSV score of the two b -jets	✓		✓
$b\text{-tag}_\text{min}$	subleading DeepCSV score of the two b -jets	✓	✓	✓
$b\text{-tag}_\text{add}$	leading DeepCSV score of all additional jets	✓		
p_T^miss	missing transverse energy	✓	✓	✓
$ \Delta\phi(p_T^\text{miss}, j) $	difference in ϕ between p_T^miss and the nearest jet with $p_T > 30$ GeV	✓		
$ \Delta\phi(p_T^\text{miss}, \ell) $	difference in ϕ between p_T^miss and the isolated lepton		✓	
m_T	mass of the isolated lepton $\vec{p}_T + \vec{p}_T^\text{miss}$	✓		
m_t	reconstructed mass of the top quark	✓		
N_j^add	number of additional jets	✓		✓
$p_T(j)_\text{add}$	highest p_T of all additional jets	✓		
SA5	number of soft-track jets with $p_T > 5$ GeV	✓	✓	✓



Multivariate Analysis

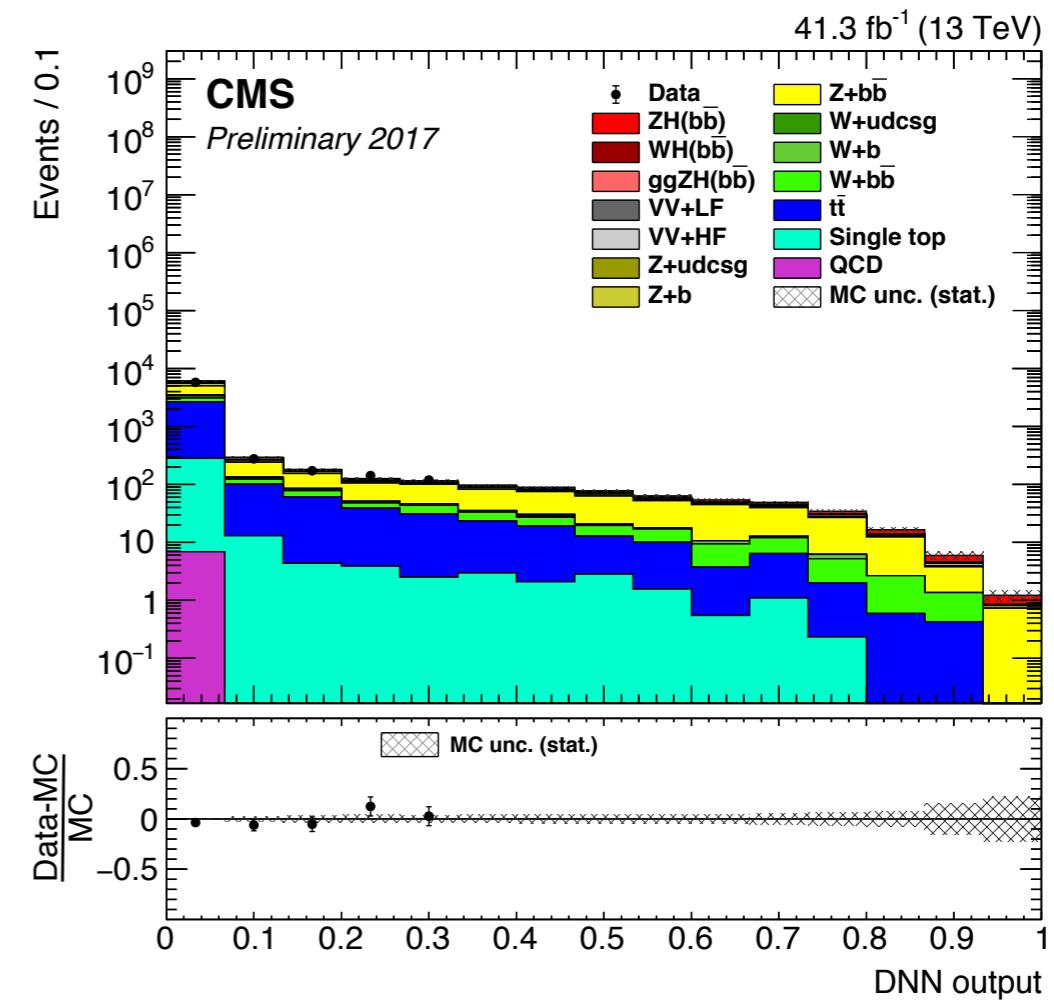
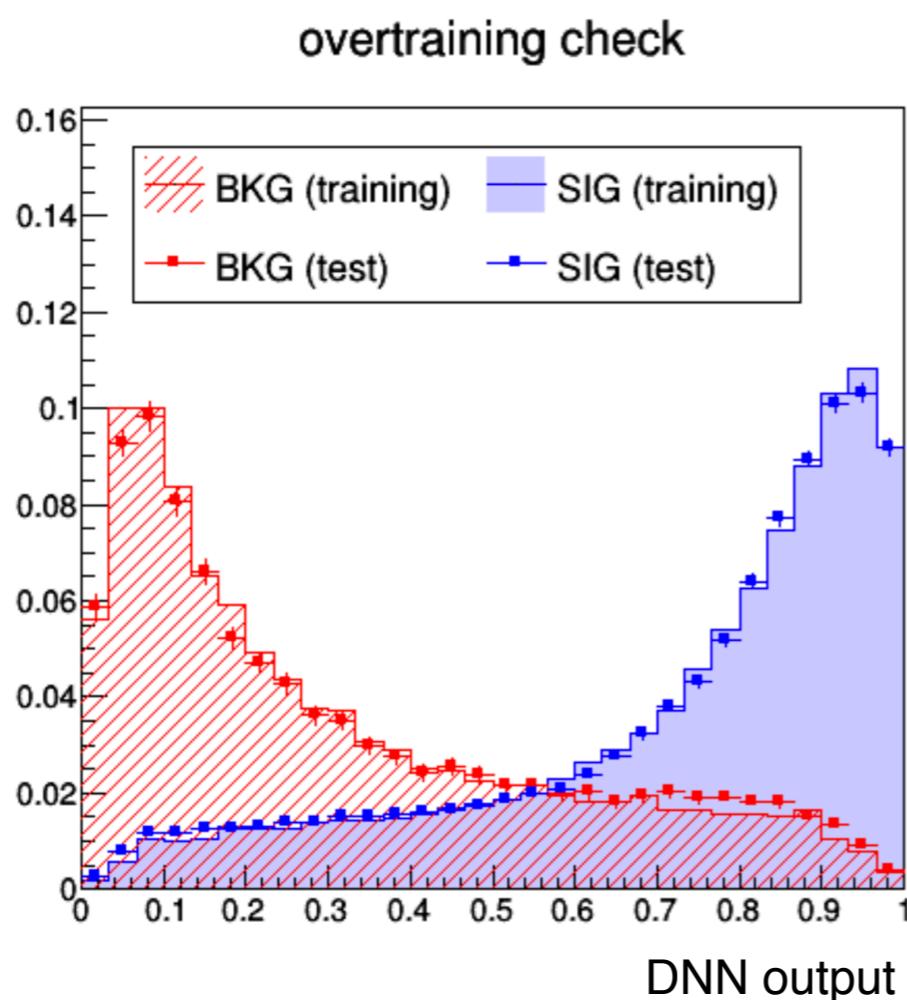
- Apply machine learning towards this supervised learning task.
- Train a “deep” neural network to classify signal and background events for each decay channel.
- Training and validation sets of ~200,000 events.

Variable	Description	$Z(\nu\bar{\nu})H$	$W(\ell\nu)H$	$Z(\ell\ell)H$
$m(jj)$	dijet invariant mass	✓	✓	✓
$p_T(jj)$	dijet transverse momentum	✓	✓	✓
$p_T(j)_\text{max}, p_T(j)_\text{min}$	transverse momentum of the two b -jets	✓		✓
$\Delta R(jj)$	distance between the two b -jets in $\eta\phi$ space			✓
$\Delta\eta(jj)$	difference in η between the two b -jets	✓		✓
$\Delta\phi(jj)$	difference in ϕ between the two b -jets	✓		✓
$p_T(V)$	vector boson transverse momentum		✓	✓
$\Delta\phi(V, H)$	difference in ϕ between the dijet system and vector boson	✓	✓	✓
$p_T(jj)/p_T(V)$	p_T ratio of the dijet system to the vector boson			✓
$m(Z)$	reconstructed mass of the Z boson			✓
$b\text{-tag}_\text{max}$	leading DeepCSV score of the two b -jets	✓		✓
$b\text{-tag}_\text{min}$	subleading DeepCSV score of the two b -jets	✓	✓	✓
$b\text{-tag}_\text{add}$	leading DeepCSV score of all additional jets	✓		
p_T^miss	missing transverse energy	✓	✓	✓
$ \Delta\phi(p_T^\text{miss}, j) $	difference in ϕ between p_T^miss and the nearest jet with $p_T > 30$ GeV	✓		
$ \Delta\phi(p_T^\text{miss}, \ell) $	difference in ϕ between p_T^miss and the isolated lepton		✓	
m_T	mass of the isolated lepton $\vec{p}_T + \vec{p}_T^\text{miss}$		✓	
m_t	reconstructed mass of the top quark		✓	
N_j^add	number of additional jets		✓	✓
$p_T(j)_\text{add}$	highest p_T of all additional jets	✓		
SA5	number of soft-track jets with $p_T > 5$ GeV	✓	✓	✓



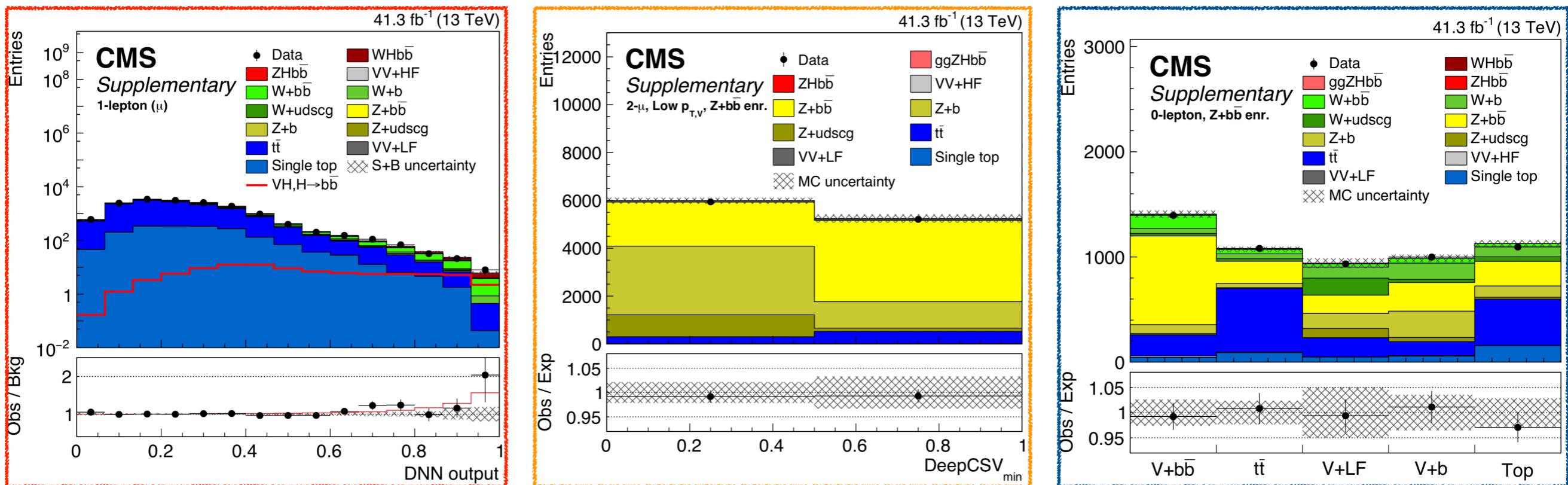
Multivariate Analysis

- Validation of the classifier inputs and outputs develop confidence in the model.
- The classifier's output scores events from 0 (background-like) to 1 (signal-like).
- Serves as the multivariate discriminant in the signal extraction fit.



Binned Shape Analysis

- Final signal extraction performed using a simultaneous binned maximum likelihood fit over the signal regions and background control regions of every decay channel.
 - DNN output distributions for the signal regions (**7 total**).
 - DeepCSV_{min} distributions for the control regions (**18 total**).
 - Multi-background DNN distributions for the heavy flavor control regions of the 0- and 1-lepton channels (**3 total**).



- Allow normalizations of V+jets and t t background processes to float independently.

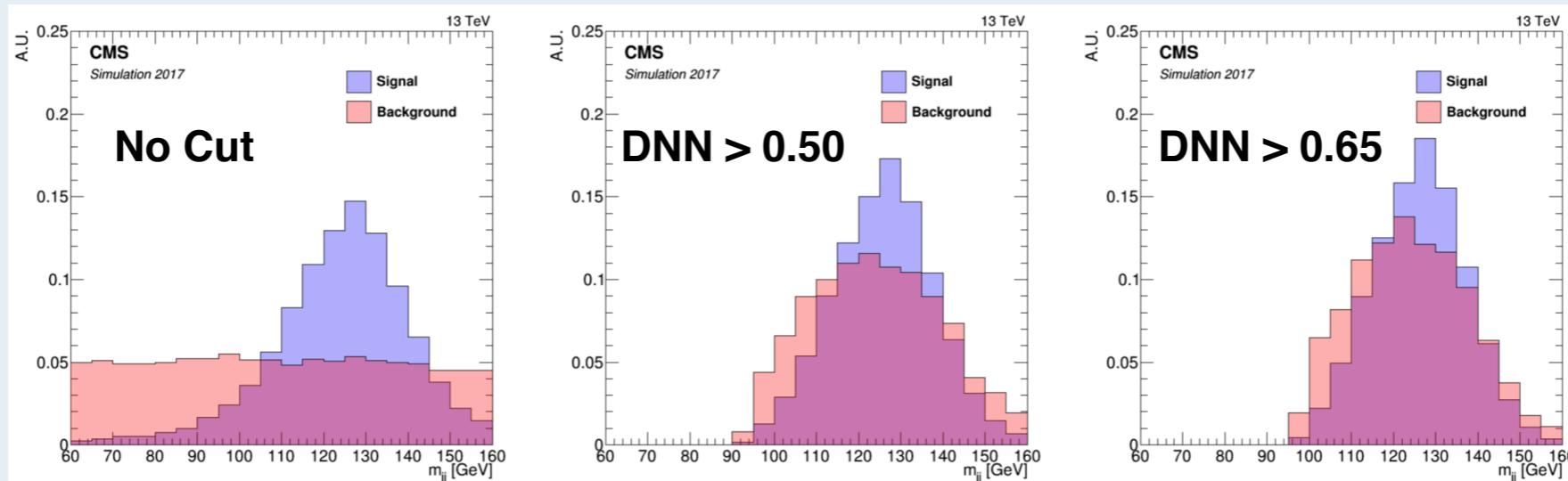
Cross-check Analyses

1. Use the known VZ($b\bar{b}$) decay to validate the analysis strategy.

- Similar final state and kinematics to VH($b\bar{b}$).
- Minimal modifications needed: merely adjust the dijet invariant mass window.
- Retrain classifiers with VZ($b\bar{b}$) as signal and VH($b\bar{b}$) as background.

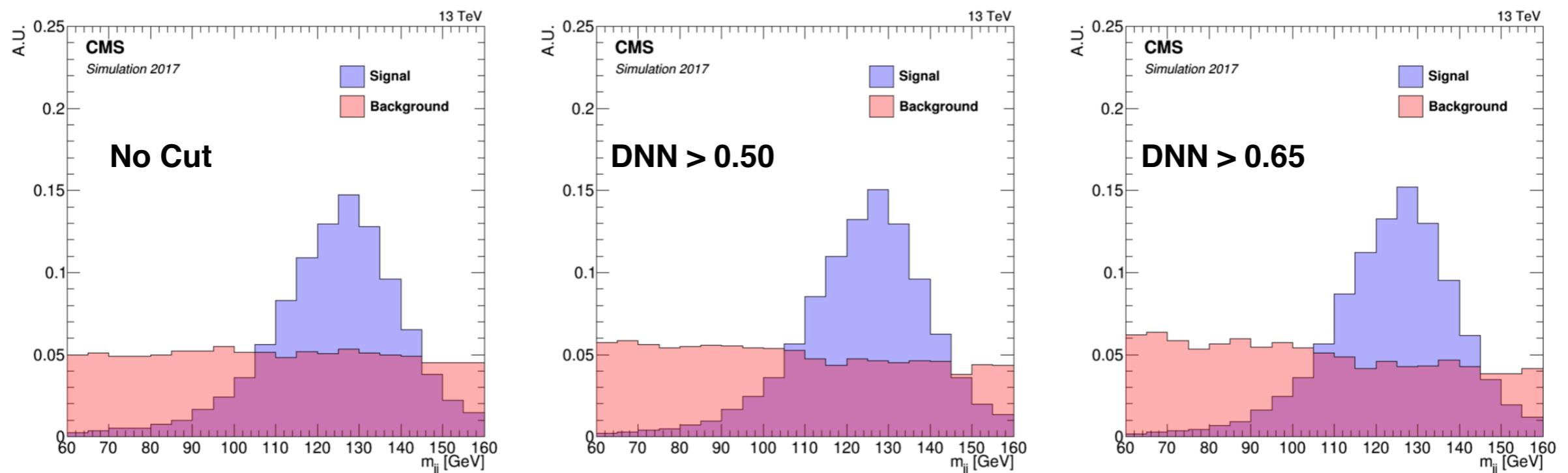
2. Perform a dijet invariant mass analysis anyway.

- Expected to be up to 20% less sensitive than the multivariate analysis, but the dijet invariant mass is the most powerful physical observable.
- Provides a visually interpretable fit result: VZ($b\bar{b}$) and VH($b\bar{b}$) mass peaks.
- How to design a signal region without distorting the background shapes?



Massless Multivariate Discriminant

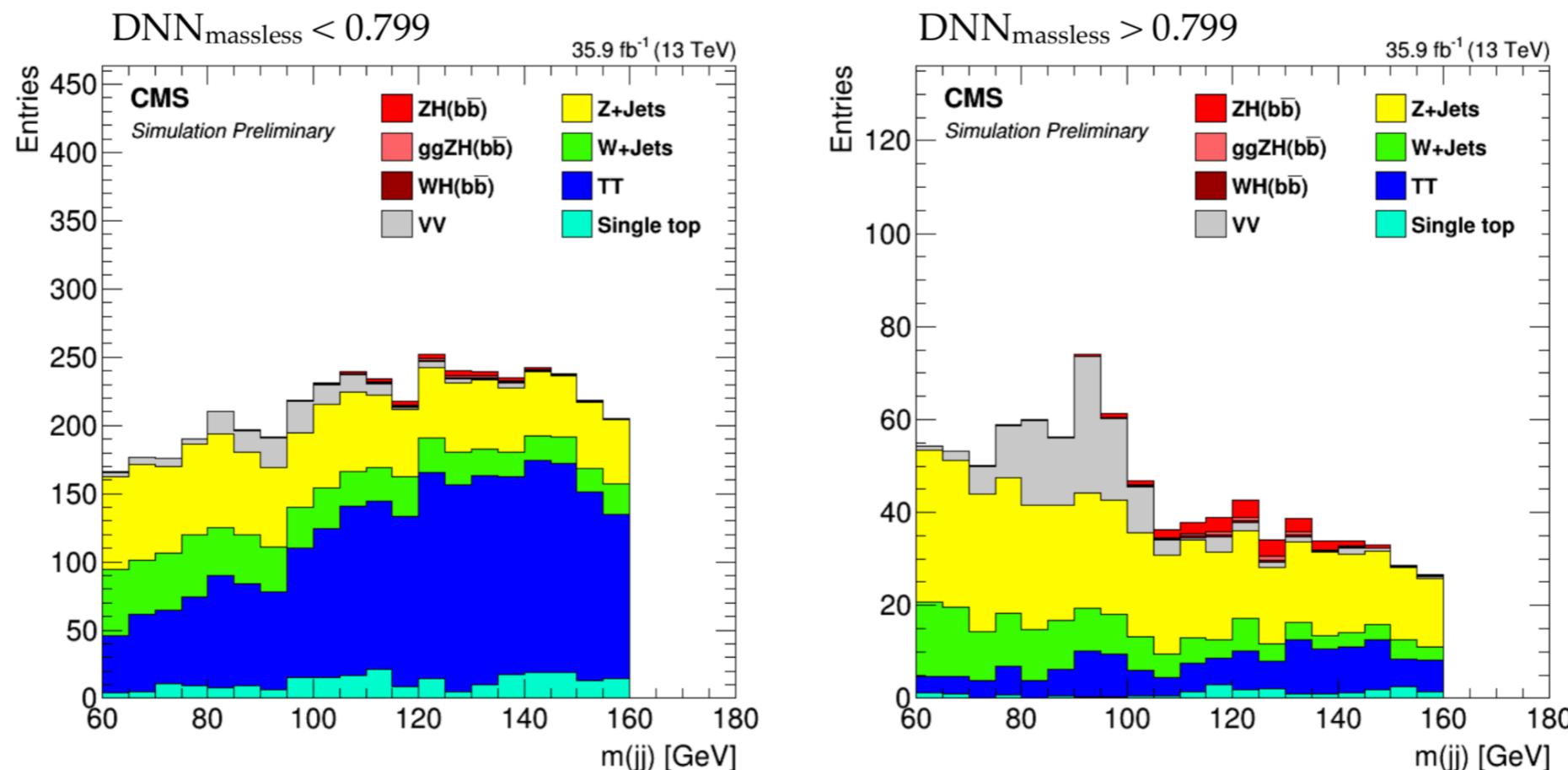
- Marginalize over the mass information contained in the classifier inputs.
 - For each decay channel, identify the features most correlated with the mass.
 - Fix those features to the mean values of their distribution over the background training events and reevaluate the classifier to obtain a “massless” DNN score.



- Successfully mitigates background distortion at the cost of classifier performance.

Automated Event Categorization

- Use the AUTOCATEGORIZER package developed by Andrew Carnes (UF).
- Implements a greedy tree-based optimization algorithm that maximizes the expected discovery significance for a PDF of interest by generating distinct event categories.
 - Dijet invariant mass distribution as the PDF of interest.
 - Massless DNN score as the categorization variable.



Systematic Uncertainties

Theoretical and experimental systematic uncertainties affect the normalizations and shapes of the signal and background processes. Each uncertainty source is treated as a nuisance parameter.

- **Luminosity (2.3%)**
- **Trigger and Lepton Efficiencies (1-2%)**
- **Jet Energy Scale**
27 independent sources based on JME recommendations, decorrelated in p_T and η .
- **Jet Energy Resolution**
10% uncertainty on b -jets after energy regression, JME standard otherwise.
- **b -Tagging**
6% per b -jet, 12% per c -jet, and 15% per light-jet, with shape variations decorrelated across p_T and η .
- **Background Normalizations**
Derived from final fit for floating processes, 15% for VV and single top.
- **Monte-Carlo Statistics**
- **Signal Cross Section (4%)**
- **Monte-Carlo Corrections**
- **Parton Distribution Functions**
- **QCD Renormalization and Factorization**

Results

Statistical Treatment

σ , Discovery Significance

- Computed using a profile likelihood asymptotic approximation.

$$q_0 = -2 \ln \frac{\mathcal{L}(\text{data}|b, \hat{\theta}_0)}{\mathcal{L}(\text{data}|\hat{\mu} \times s + b, \hat{\theta})} \text{ for } \hat{\mu} > 0$$

- Local p-value measures the probability that an upward fluctuation of background is responsible for an observed excess of events in the absence of signal.

$$p_0 = P(q_0 \geq q_0^{\text{data}}|b) = \frac{1}{\sqrt{2\pi}} \int_z^{+\infty} e^{-\frac{1}{2}x^2} dx$$

- Quoted in units of standard deviations σ , where a threshold of **3 σ** is required to establish **evidence** and **5 σ** is required to claim an **observation**.

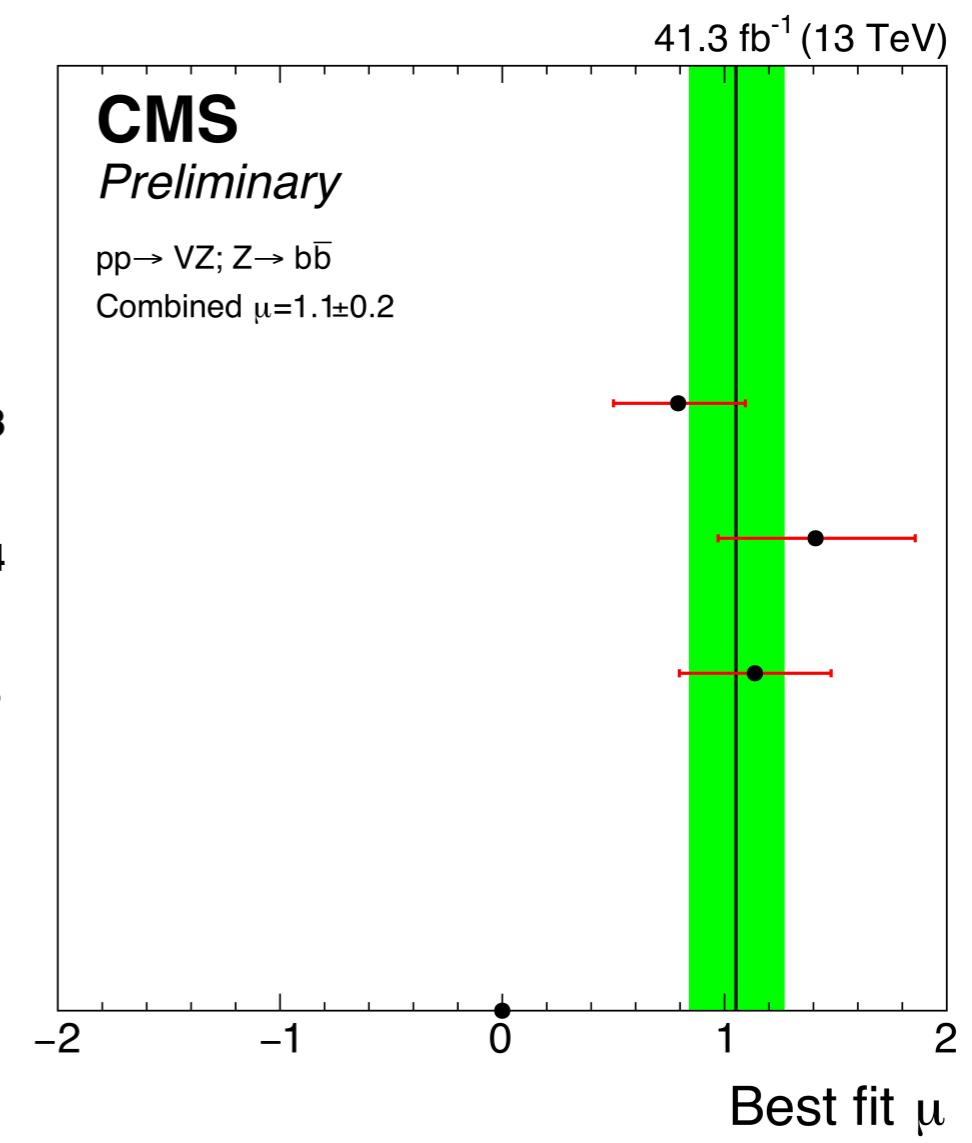
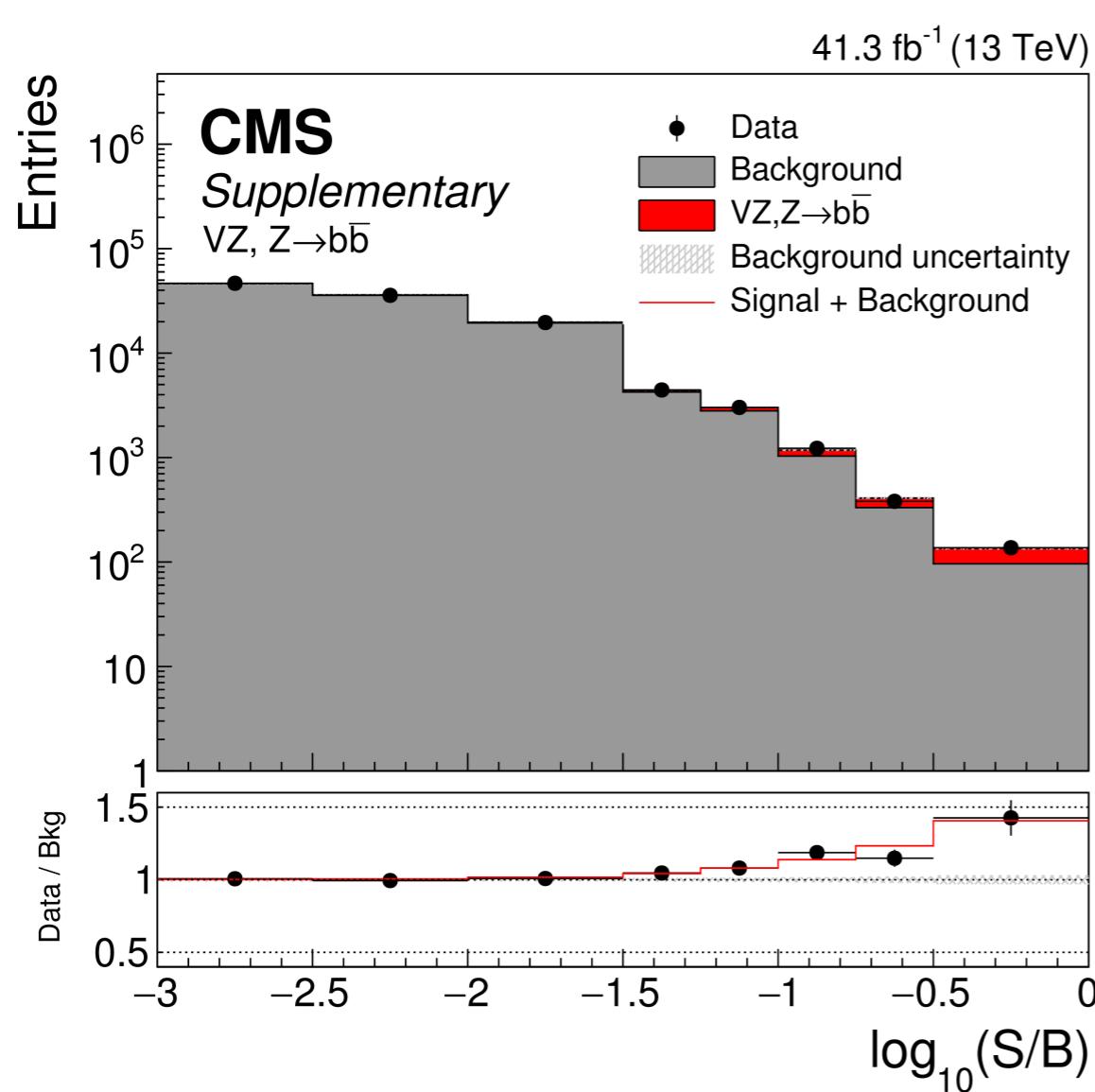
$\hat{\mu}$, Signal Strength

- Defined as the measured production cross section times the $H \rightarrow b\bar{b}$ branching fraction divided by the expected SM value.
- Value of $\hat{\mu}$ is allowed to vary freely and the nuisance parameters $\hat{\theta}$ vary within their uncertainties during the final fit.
- A best fit value compatible with unity indicates agreement with the Standard Model.

Results

VZ($b\bar{b}$) Analysis

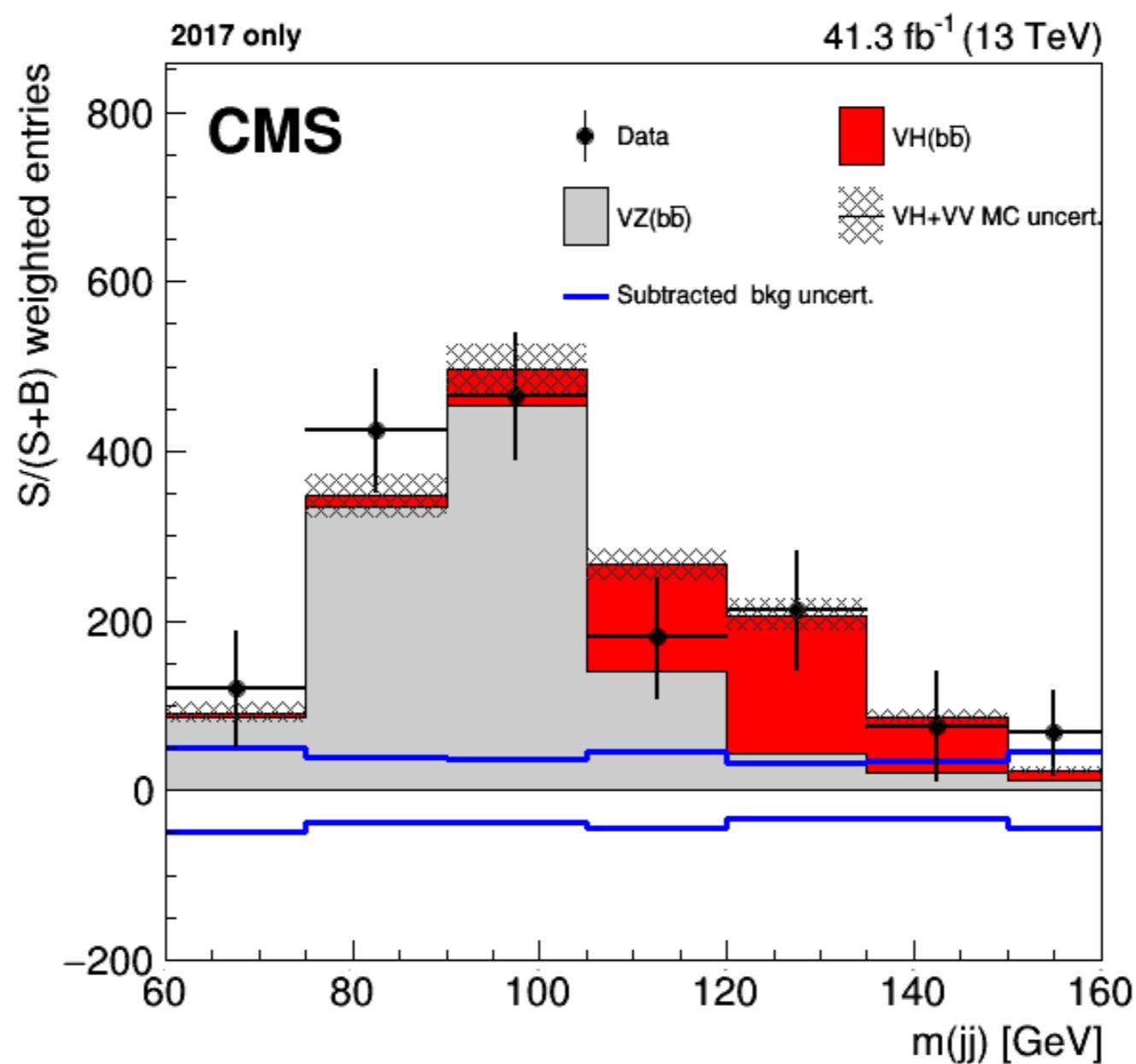
Expected Significance	Observed Significance	Signal Strength
4.99	5.19	1.05 ± 0.22



Results

Dijet Invariant Mass Analysis

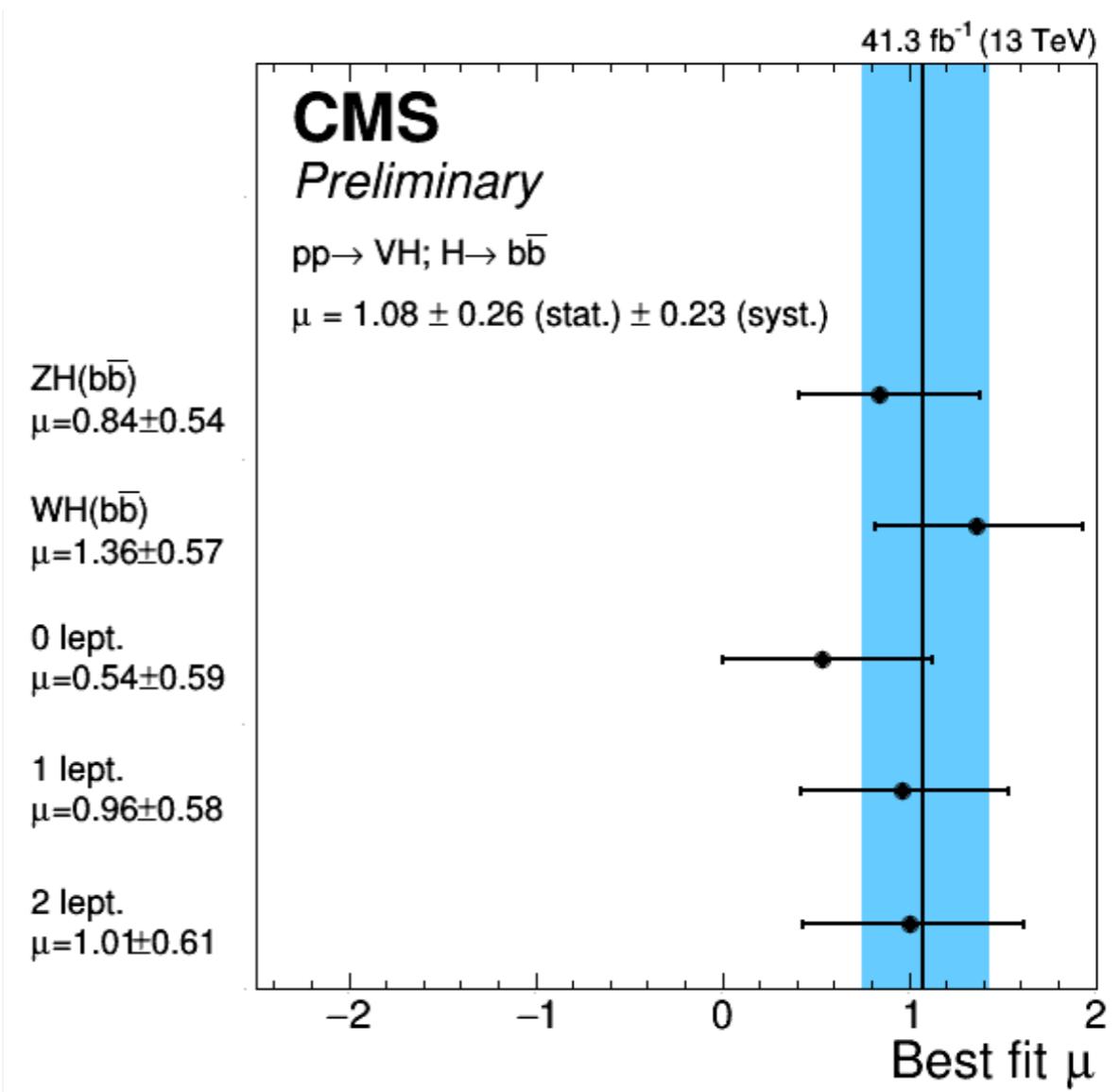
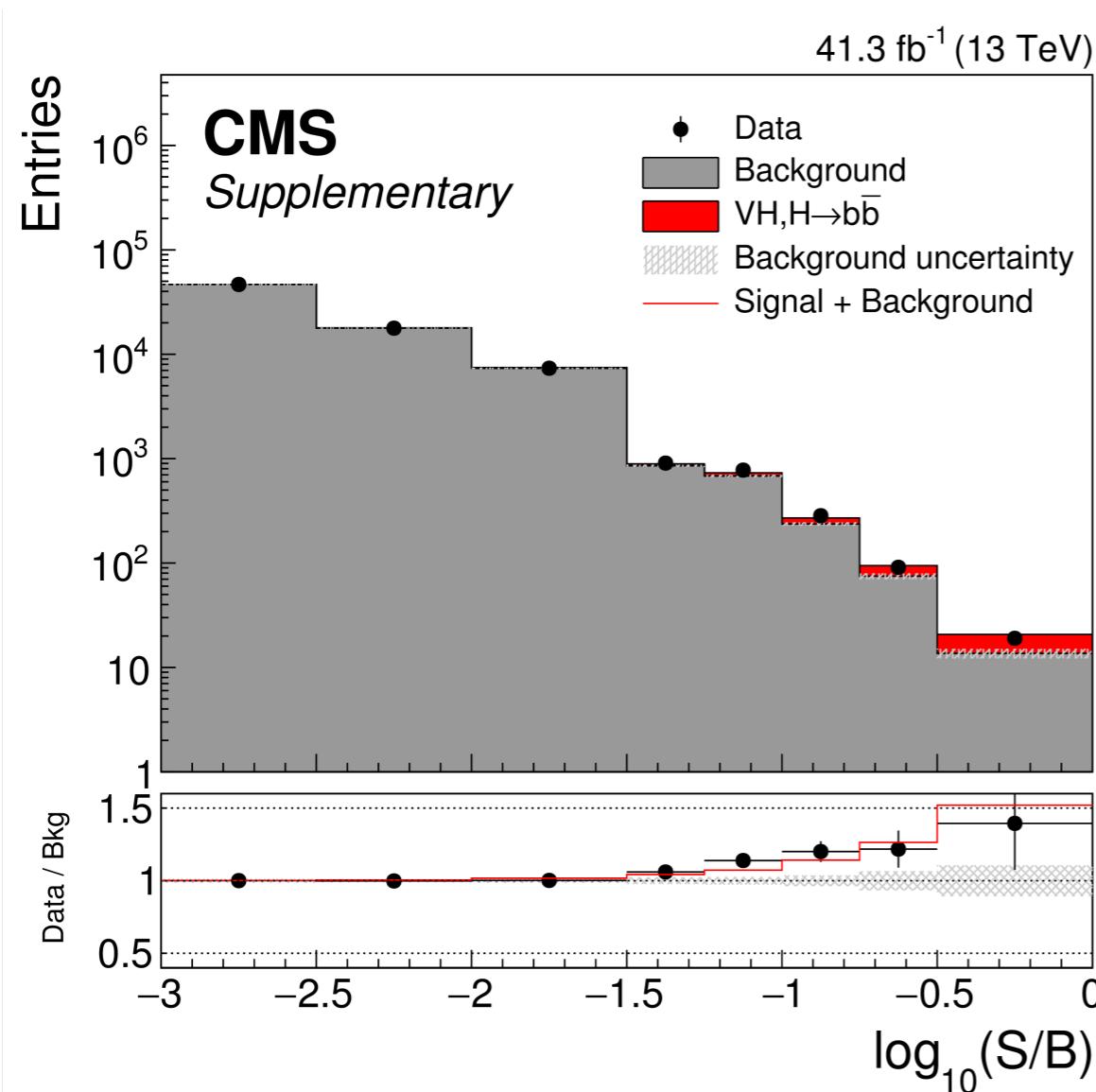
Expected Significance	Observed Significance	Signal Strength
2.19	1.33	0.59 ± 0.45



Results

2017 VH($b\bar{b}$) Analysis

Expected Significance	Observed Significance	Signal Strength
3.14	3.35	1.08 ± 0.34



Results

2017 VH($b\bar{b}$) Analysis

Uncertainty source	$\Delta\mu$	
Statistical	+0.26	-0.26
Normalization of backgrounds	+0.12	-0.12
Experimental	+0.16	-0.15
b-tagging efficiency and misid	+0.09	-0.08
V+jets modeling	+0.08	-0.07
Jet energy scale and resolution	+0.05	-0.05
Lepton identification	+0.02	-0.01
Luminosity	+0.03	-0.03
Other experimental uncertainties	+0.06	-0.05
MC sample size	+0.12	-0.12
Theory	+0.11	-0.09
Background modeling	+0.08	-0.08
Signal modeling	+0.07	-0.04
Total	+0.35	-0.33

Dominant sources of systematic uncertainty due to normalization and modeling of backgrounds, b -tagging, and available Monte-Carlo statistics.

Results

Observation of $H \rightarrow b\bar{b}$

A global fit is performed, combining all available CMS analyses of different $H \rightarrow b\bar{b}$ production processes, including the 2016 and 2017 VH($b\bar{b}$) analyses, gluon fusion, vector boson fusion, and $t\bar{t}$ associated production.

Expected Significance	Observed Significance	Signal Strength
5.5	5.6	1.04 ± 0.20

The $H \rightarrow b\bar{b}$ decay has been observed by CMS!

Long-sought decay of Higgs boson observed

Six years after its discovery, the Higgs boson has at last been observed decaying to fundamental particles known as bottom quarks.

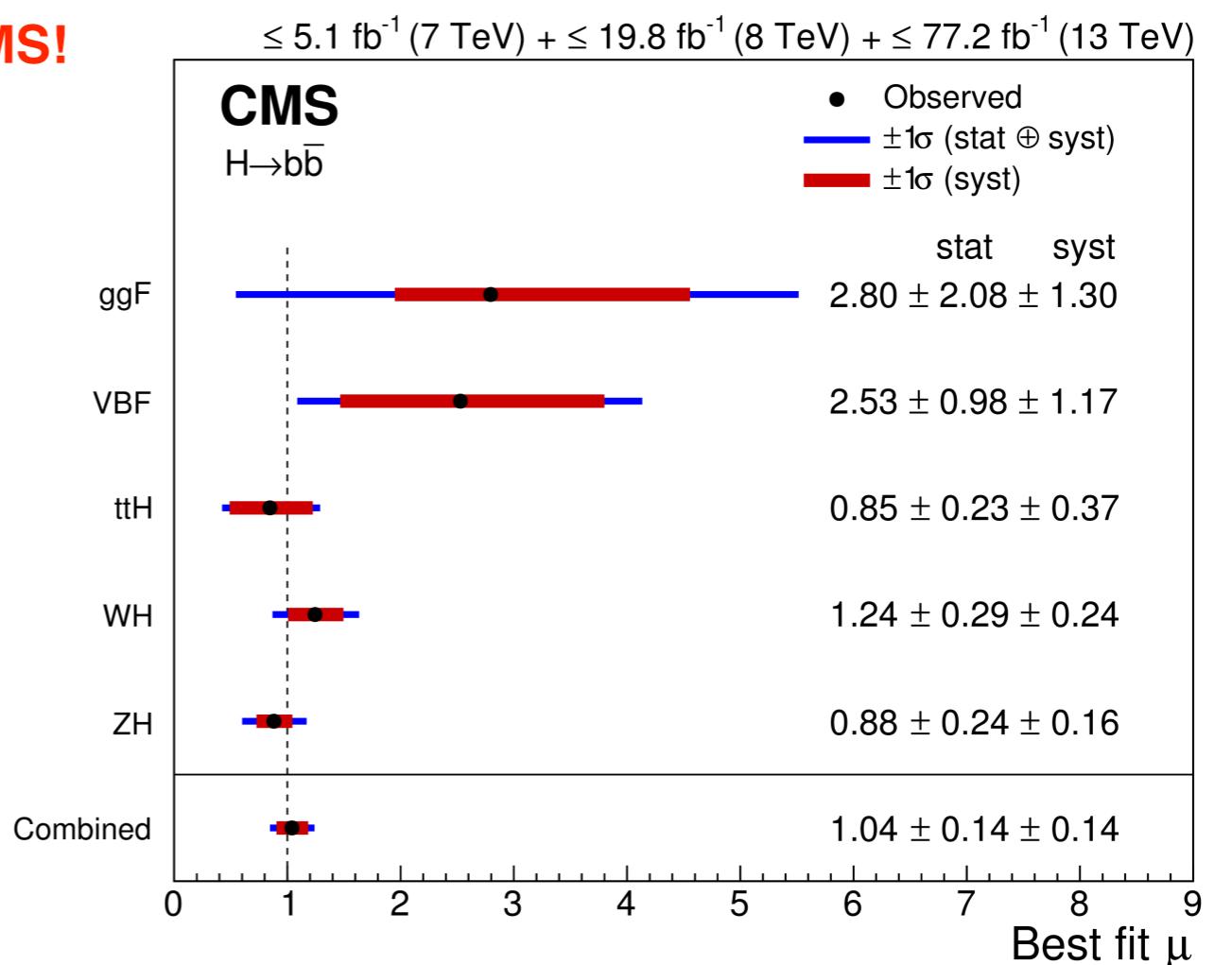
28 AUGUST, 2018

Viewpoint: Higgs Decay into Bottom Quarks Seen at Last

Howard E. Haber, Santa Cruz Institute for Particle Physics, University of California, Santa Cruz, CA, USA

September 17, 2018 • Physics 11, 91

Two CERN experiments have observed the most probable decay channel of the Higgs boson—a milestone in the pursuit to confirm whether this remarkable particle behaves as physicists expect.



Conclusions

Conclusions

- As of the Summer of 2018, the $H \rightarrow b\bar{b}$ decay has finally been observed by CMS!
- This is a remarkable scientific achievement that marks the end of an endeavor spanning over 30 years and multiple colliders and experiments worldwide.

Started from a bottom in a pair...



...started from the bottom, now we're here!

Conclusions

Remarks

- How to improve on the 20% precision for the signal strength?
 - The dijet invariant mass analysis is a promising avenue, as the signal peak is only affected by the systematic variations of *flat* background shapes.
 - The $H \rightarrow \mu\bar{\mu}$ analysis rescales the classifier input features before training to achieve a model that is insensitive to the mass.
 - Adversarial techniques developed for jet tagging algorithms have demonstrated the ability to decorrelate a model's output from a target variable.
- Future analyses will face higher luminosities and energies and comb through more data than ever before.
 - Better pileup removal algorithms could be required to achieve good dijet invariant mass resolution and *b*-tagging performance.
 - Dedicated jet substructure analyses may be needed to recover highly boosted signal events.
 - Novel machine learning approaches have the potential to play key roles.
- The $VH(b\bar{b})$ analysis can be adapted towards a search for the $H \rightarrow c\bar{c}$ decay.

