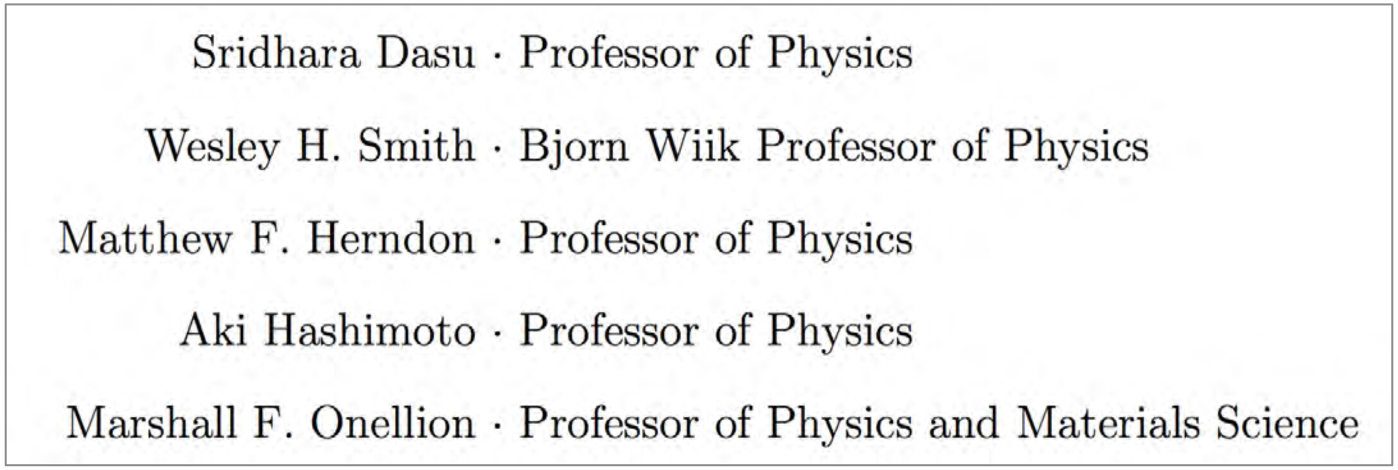
A Study of the Standard Model Higgs Boson Decaying to a Pair of Tau Leptons with the CMS Detector at the LHC

PhD Defense - 10 May 2018

Thesis Committee:



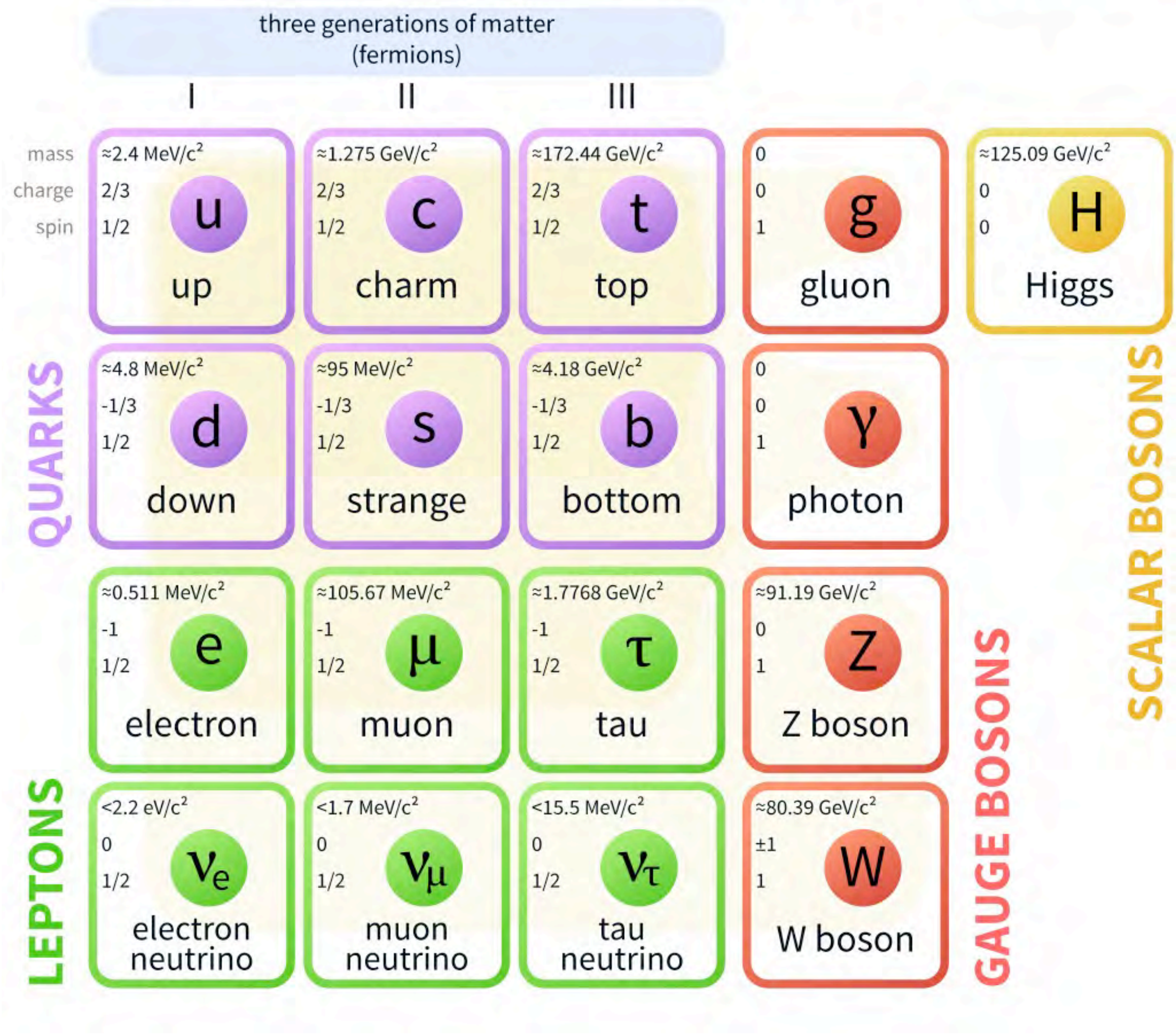
Sridhara Dasu · Professor of Physics
Wesley H. Smith · Bjorn Wiik Professor of Physics
Matthew F. Herndon · Professor of Physics
Aki Hashimoto · Professor of Physics
Marshall F. Onellion · Professor of Physics and Materials Science

Tyler Ruggles

Outline

- The Standard Model
- Previous Experimental Results
- The LHC & CMS
- Event Capture and Reconstruction
- Event Simulation
- 2016 Run-II Higgs → ττ Analyses
  - Gluon Fusion and VBF
  - Associated Production
- Results
- Conclusion

The Standard Model Particles

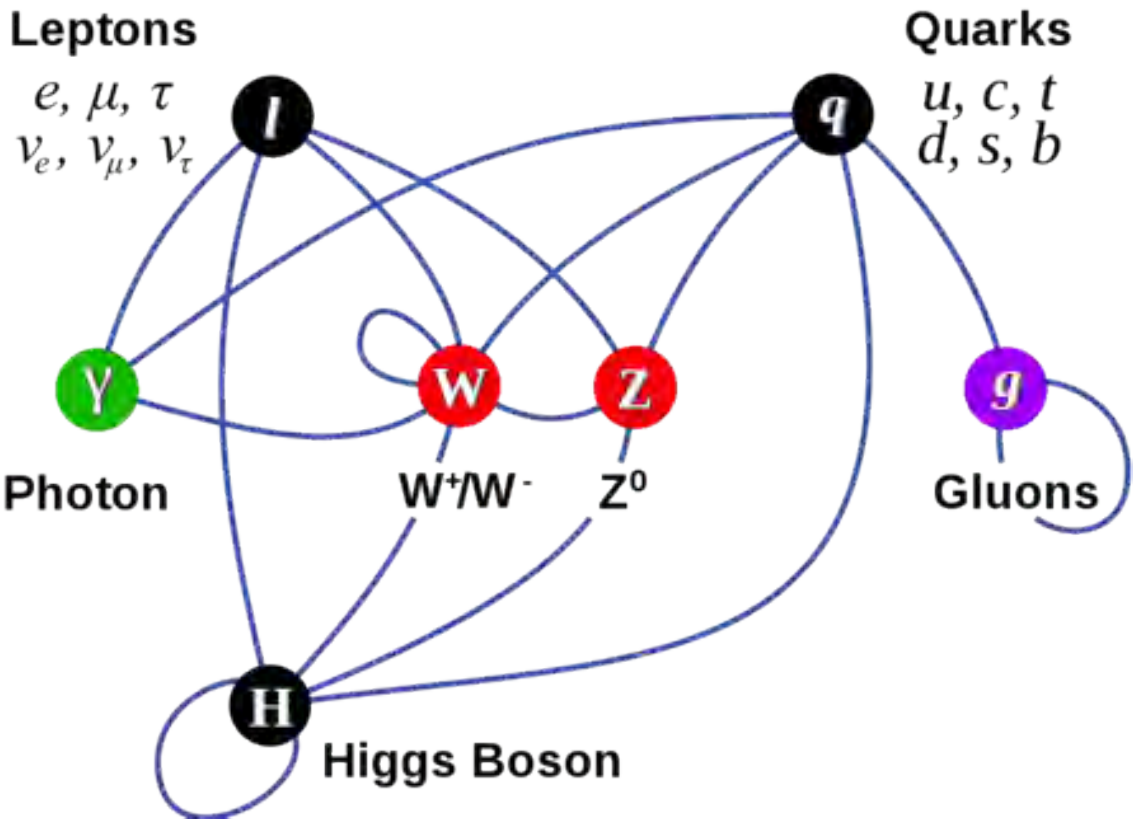


three generations of matter (fermions)

	I	II	III		
mass charge spin	≈2.4 MeV/c^2 2/3 1/2 u up	≈1.275 GeV/c^2 2/3 1/2 c charm	≈172.44 GeV/c^2 2/3 1/2 t top	0 0 1 g gluon	≈125.09 GeV/c^2 0 0 H Higgs
QUARKS	≈4.8 MeV/c^2 -1/3 1/2 d down	≈95 MeV/c^2 -1/3 1/2 s strange	≈4.18 GeV/c^2 -1/3 1/2 b bottom	0 0 1 γ photon	SCALAR BOSONS
	≈0.511 MeV/c^2 -1 1/2 e electron	≈105.67 MeV/c^2 -1 1/2 μ muon	≈1.7768 GeV/c^2 -1 1/2 τ tau	≈91.19 GeV/c^2 0 1 Z Z boson	
LEPTONS	<2.2 eV/c^2 0 1/2 ν_e electron neutrino	<1.7 MeV/c^2 0 1/2 ν_μ muon neutrino	<15.5 MeV/c^2 0 1/2 ν_τ tau neutrino	≈80.39 GeV/c^2 ±1 1 W W boson	GAUGE BOSONS

The Force Carriers

- Gluons:
  - Mediate the Strong Force
  - Hold quarks in bound states of mesons and baryons
- Photons:
  - Mediate Electromagnetic Force
  - Bind electrons to atoms
- $W^{+/-}$ and Z:
  - Mediate the Weak Force
  - Responsible for some nuclear decays

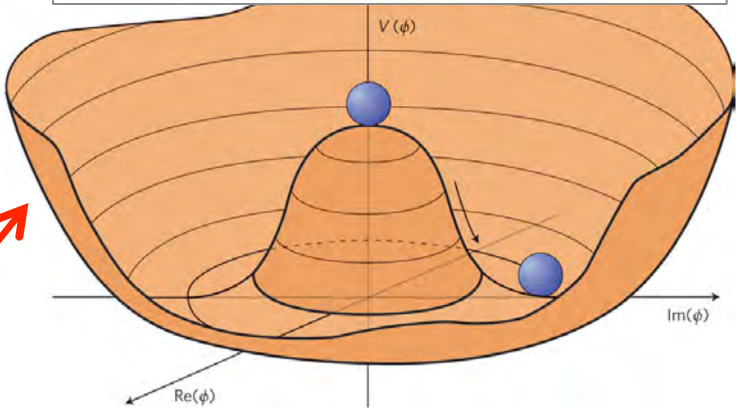


Electroweak Symmetry Breaking

- In $SU(2)_L \times U(1)_Y$ symmetry gauge bosons are massless
- Electroweak Symmetry Breaking → Higgs potential:

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

$$v^2 = Re(\Phi)^2 + Im(\Phi)^2$$



- $\mu^2 < 0$ & $\lambda > 0$ leads to a potential depicted above, v = 246 GeV
- Using a complex doublet scalar field

$$\phi = \begin{pmatrix} \phi_\alpha \\ \phi_\beta \end{pmatrix} = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$$

- Three components of the scalar field give mass to the $W^{+/-}$ and Z bosons

$$M_Z = \frac{1}{2} v \sqrt{g^2 + g'^2}$$

$$M_W = \frac{1}{2} v g$$

- The fourth component is a new fundamental scalar particle with mass

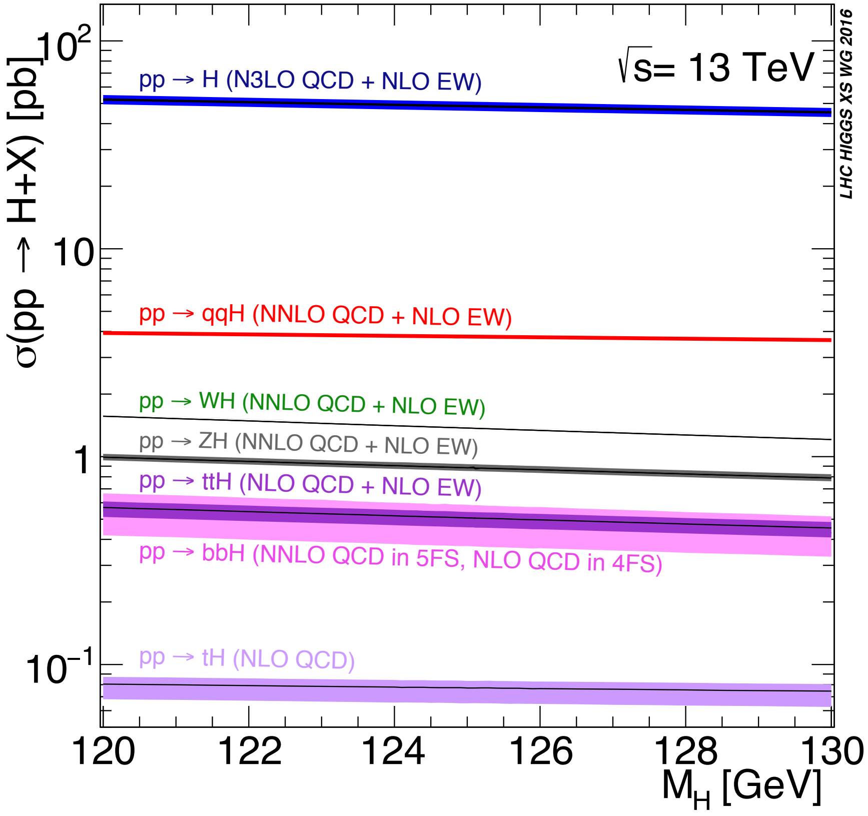
$$m_H^2 = 2\lambda v^2$$

- The scalar field interacts with the fermions through a Yukawa coupling relating the masses of the fermions to their Higgs couplings

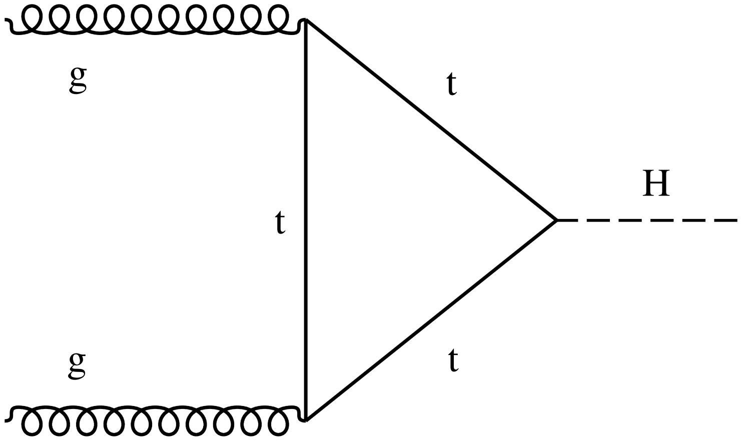
$$m_f = \frac{h_f v}{\sqrt{2}}$$

Higgs Boson Production

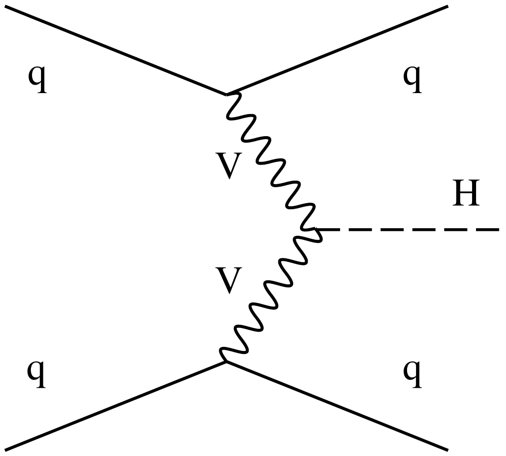
- This thesis targets the four leading production processes



gluon fusion: $\sigma \approx 49$ pb

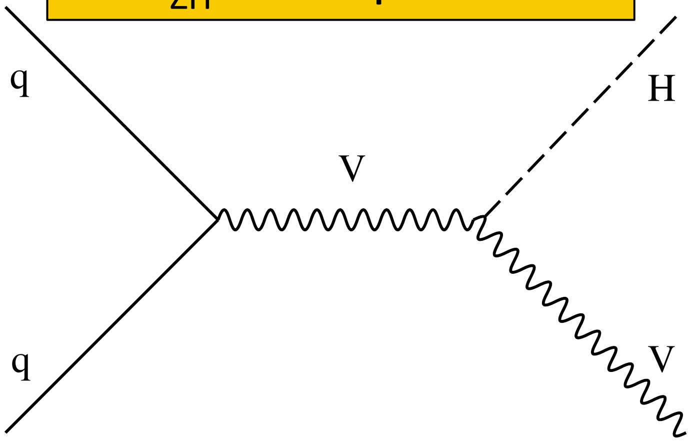


vector boson fusion: $\sigma \approx 3.8$ pb



associated production:

- $\sigma_{WH} \approx 1.4$ pb
- $\sigma_{ZH} \approx 0.88$ pb

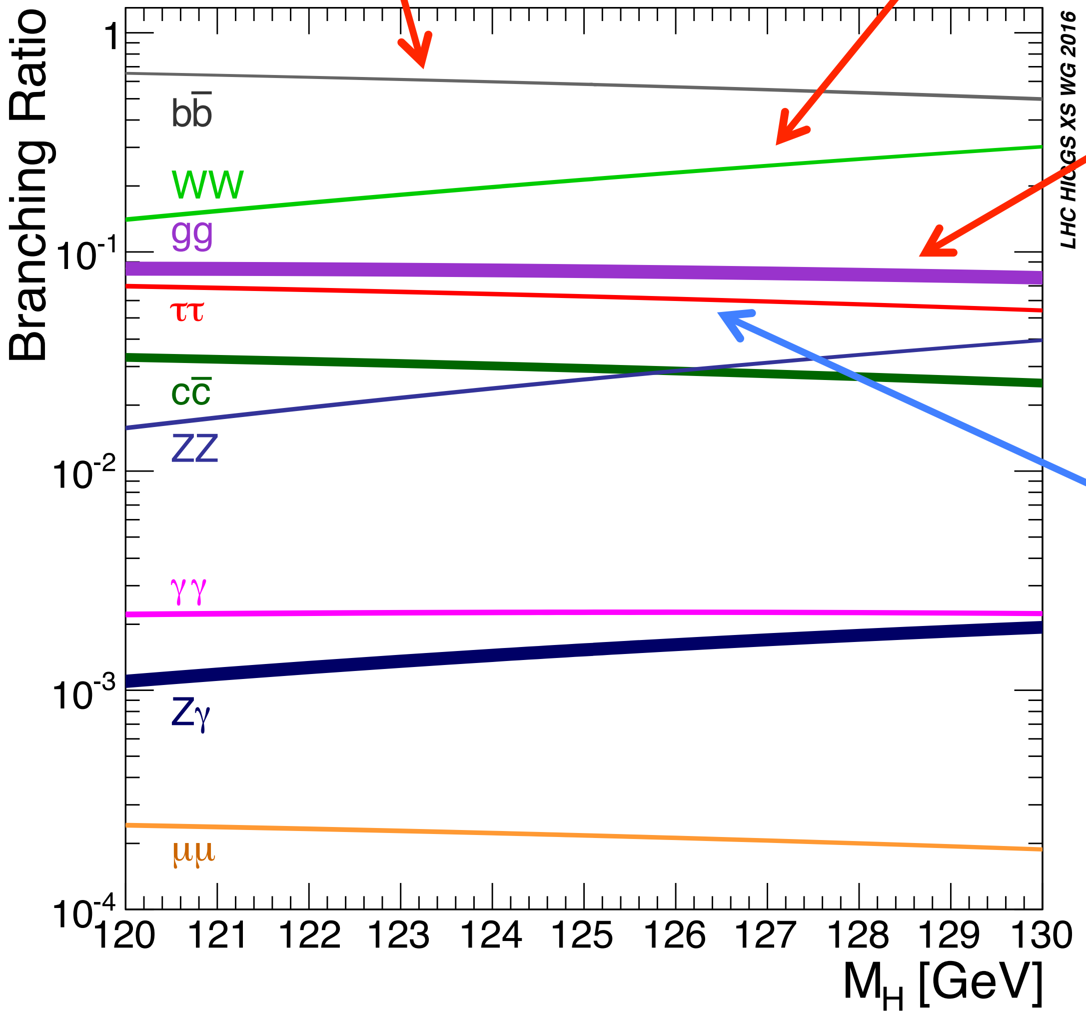


Higgs Boson Decays

H → bb: 58.1%

H → WW: 21.5%

H → gg: 8.2%



H → ττ: 6.3%
- Largest leptonic decay
- Much reduced background compared to H → bb

- H → cc: 2.9%
- H → ZZ: 2.6%
- H → Others: < 1%

Higgs Measurements

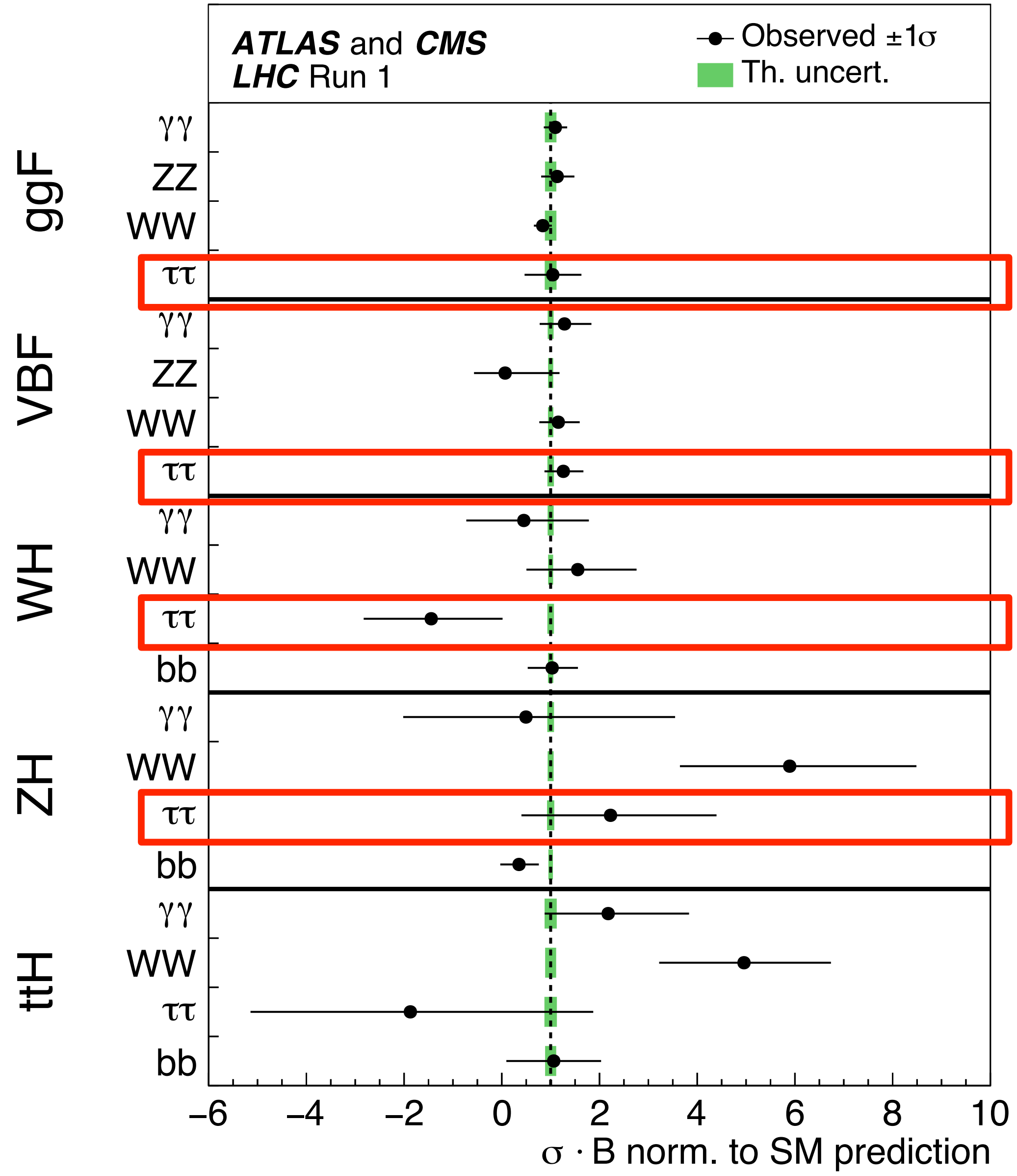
- Higgs boson discovered in 2012

- Mass measured very precisely:
  - $m_H = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.})$ GeV

- CMS and ATLAS have moved from “Higgs discovery mode” to “Higgs measurement mode”

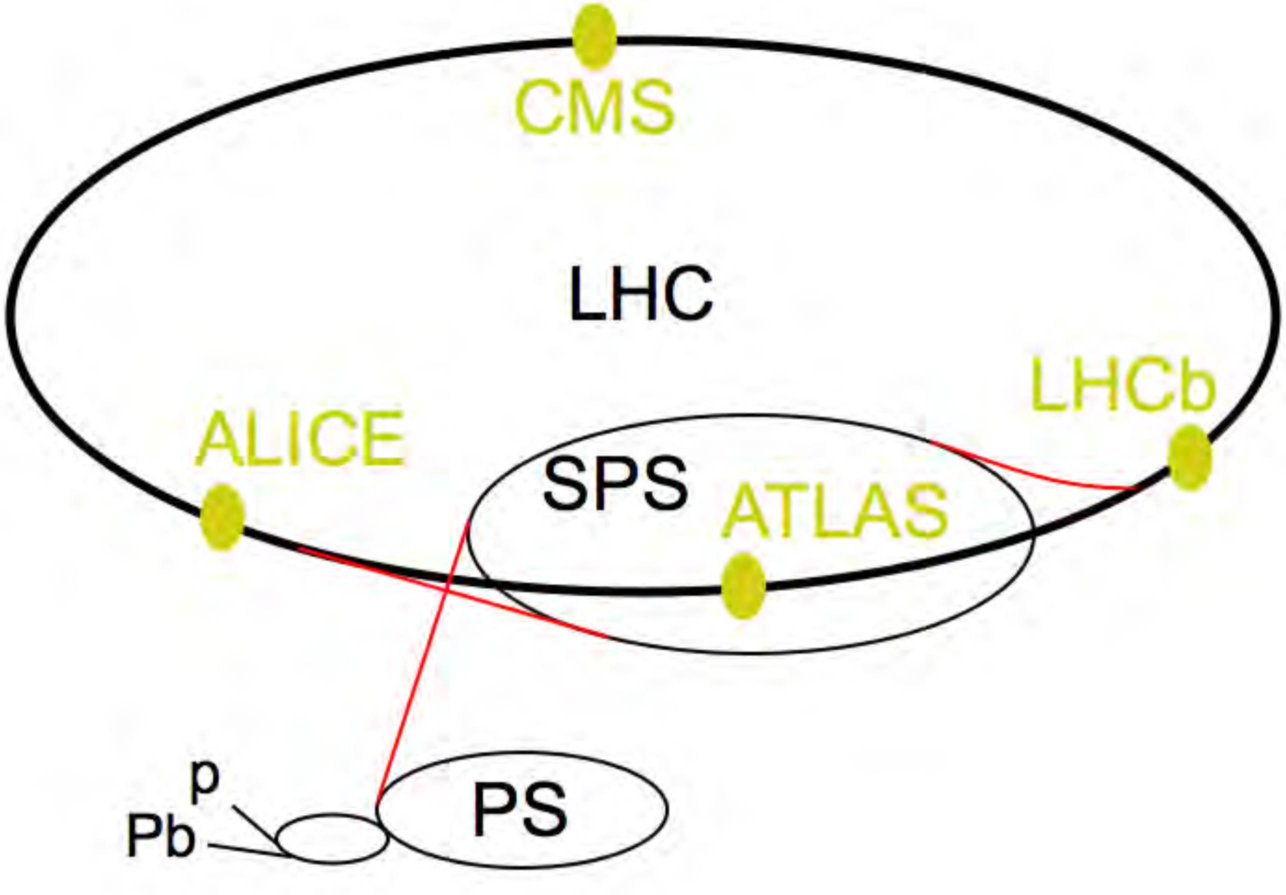
- Want to measure the different Higgs boson decay processes, production mechanisms, and couplings for deviations from the standard model predictions

- Specifically target the four highlighted processes



The Large Hadron Collider

2010-2012	Max Energy per Beam
Linac 2	50 MeV
Proton Synchrotron Booster	1.4 GeV
Proton Synchrotron	25 GeV
Super Proton Synchrotron	450 GeV
LHC	7 TeV



- 27km circumference
- General purpose experiments:
  - CMS, ATLAS
- Heavy Ions: ALICE
- b physics: LHC-b

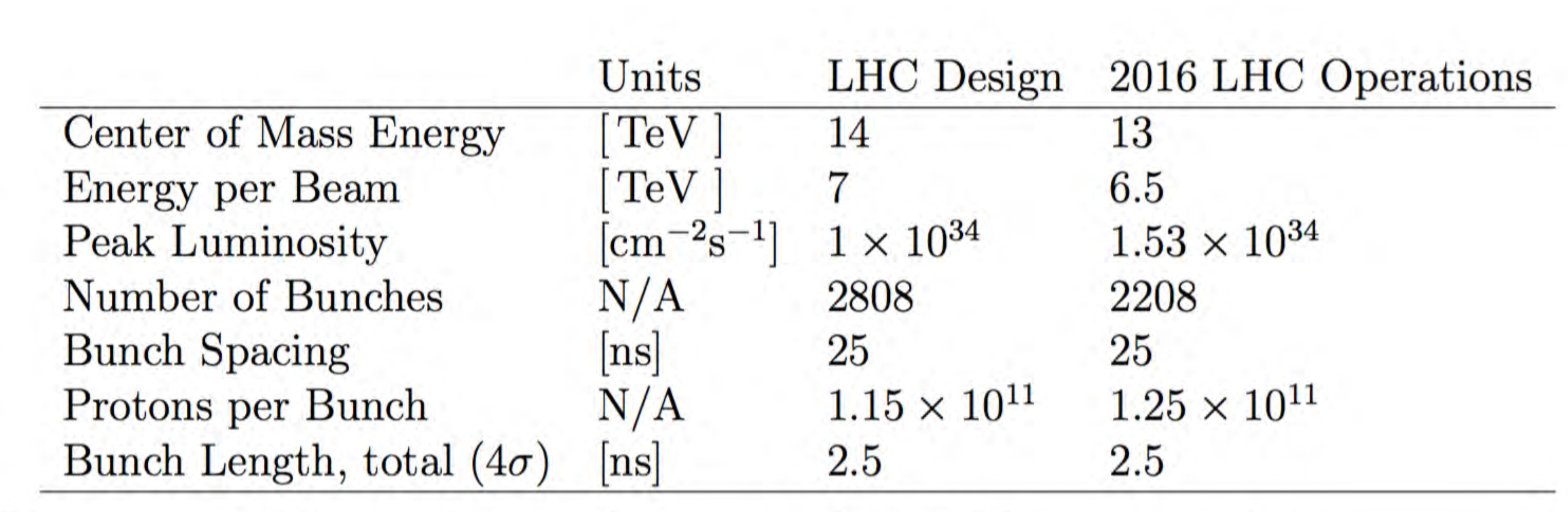
Proton Beam

- Number of events for a given process:

$$N = \sigma \int \mathcal{L}dt$$

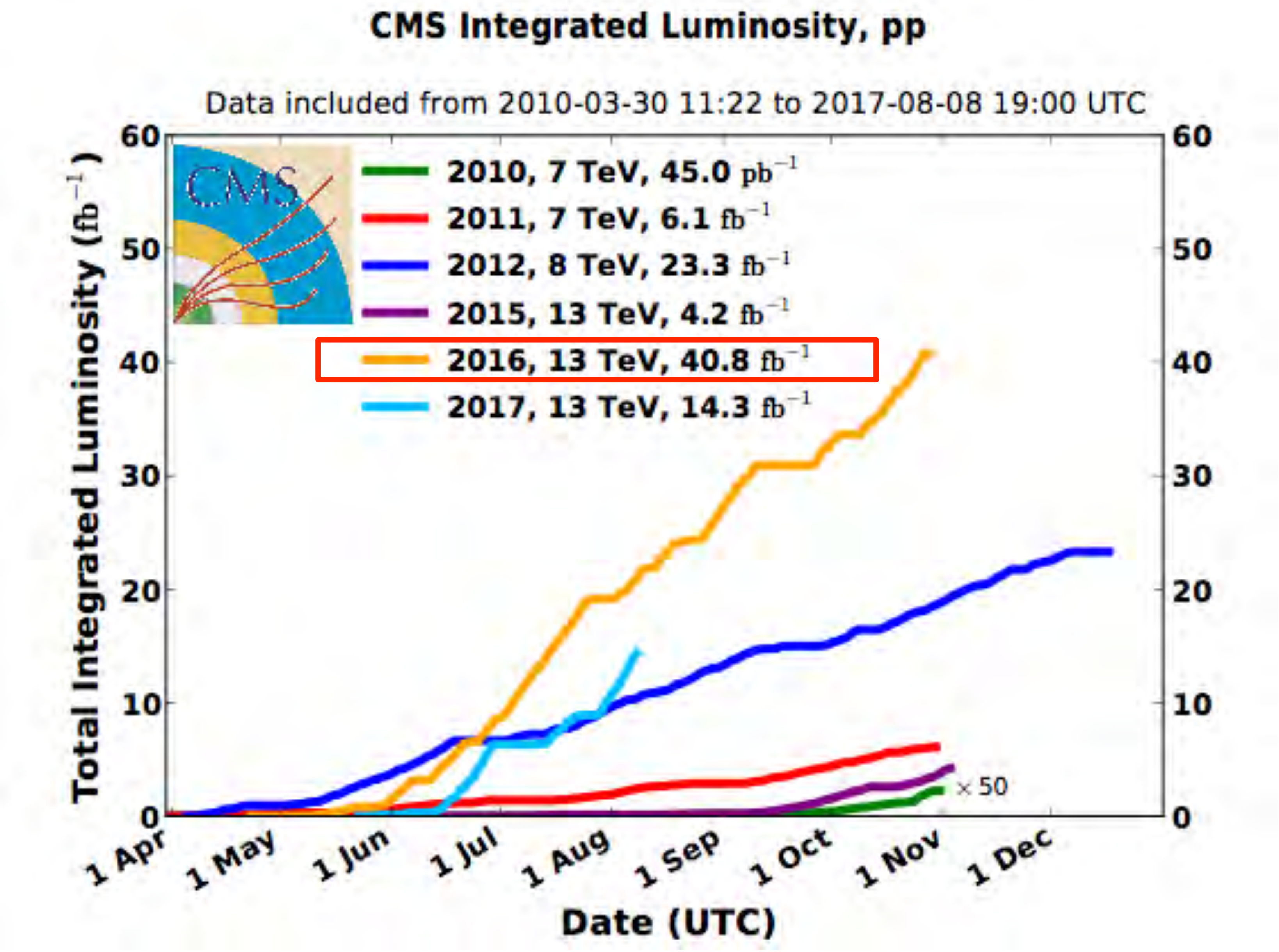
  - σ = cross section of process
  - L = Instantaneous luminosity of collider

- The LHC is surpassing some of its design targets allowing it to deliver increased peak luminosity to the experiments



	Units	LHC Design	2016 LHC Operations
Center of Mass Energy	[TeV]	14	13
Energy per Beam	[TeV]	7	6.5
Peak Luminosity	$[\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	1×10^{34}	1.53×10^{34}
Number of Bunches	N/A	2808	2208
Bunch Spacing	[ns]	25	25
Protons per Bunch	N/A	1.15×10^{11}	1.25×10^{11}
Bunch Length, total (4σ)	[ns]	2.5	2.5

CMS Integrated Luminosity



The Compact Muon Solenoid

Image: http://cds.cern.ch/record/1344500/files/image_0029.jpg?subformat=icon-1440, “March 24, 2015

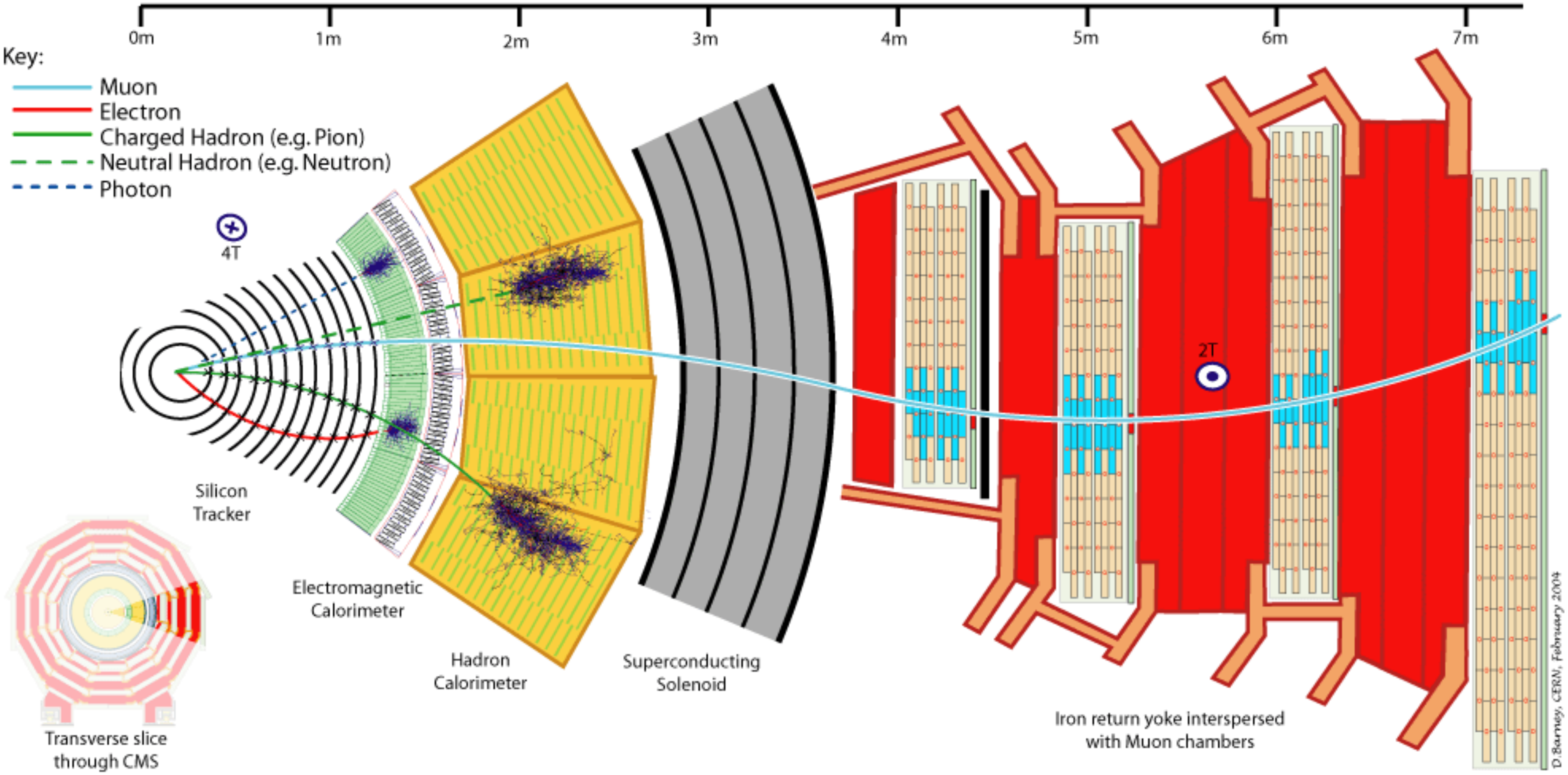
CMS Detector

Basic Dimensions / Stats

- Length: 21.6 m
- Diameter: 15m
- Weight: Approx. 14,000,000 kg

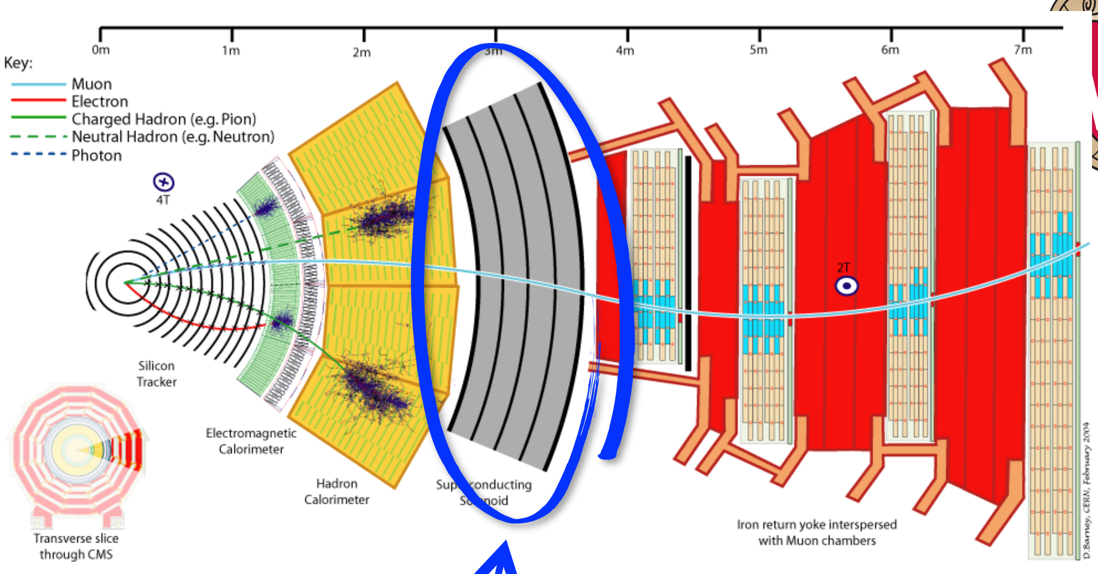
Beam line outwards

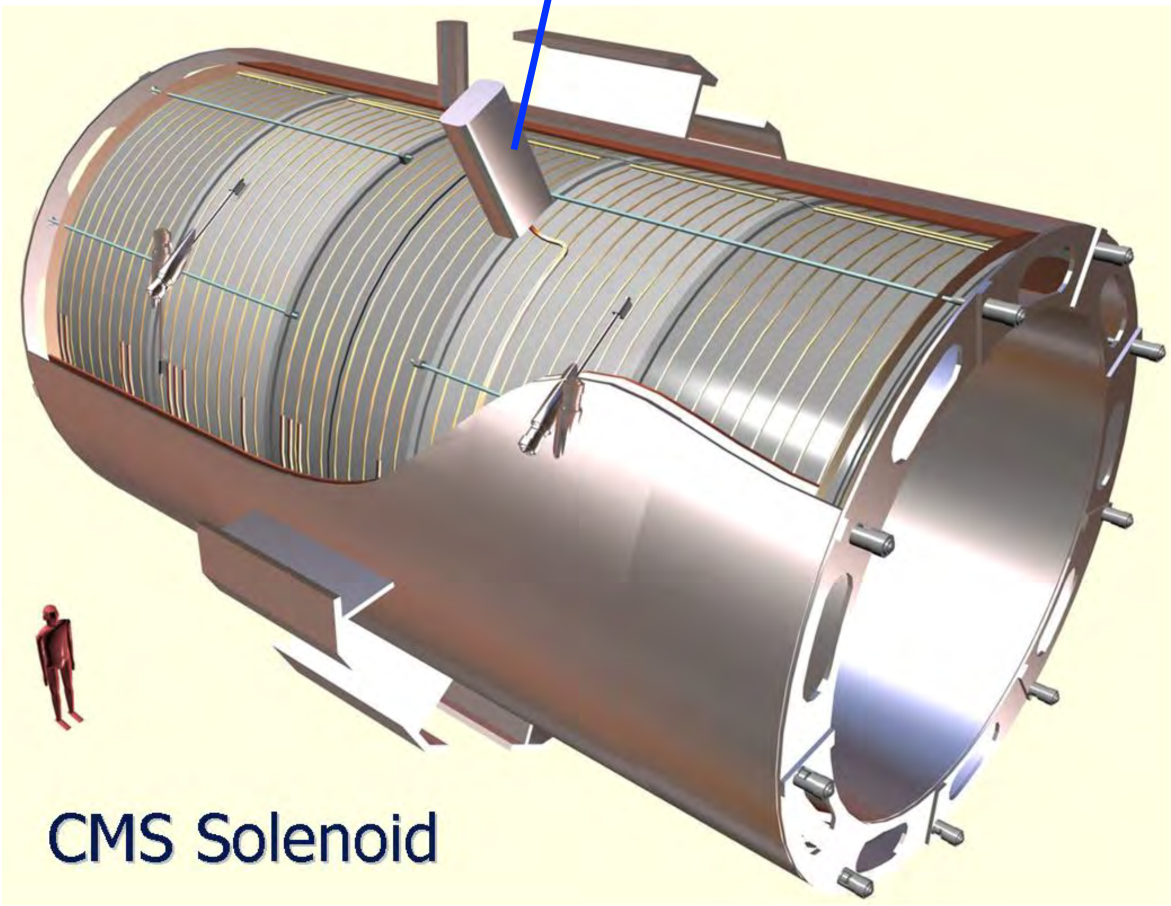
- tracker
- electromagnetic calorimeter
- hadronic calorimeter
- Solenoid
- muon tracker & iron yoke



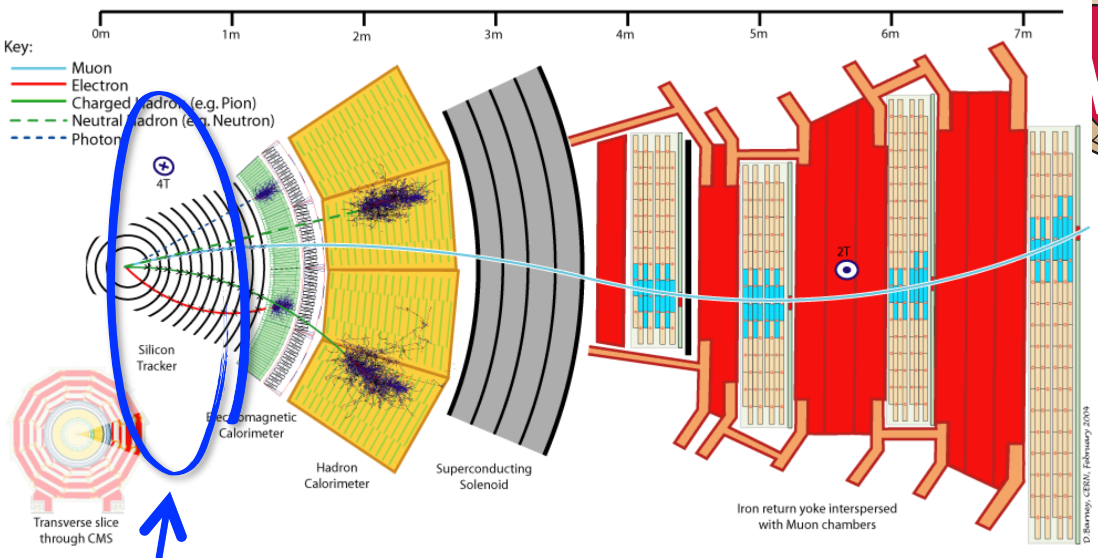
Solenoid

- **Purpose:** Strong magnetic field to bend path of charged particles
  - Allows momentum calculation
- 12.5 m length x 6.3 m diameter, cooled to 4.7 K
- 3.8 T field inside central barrel
- Iron return yoke provides ~2T field outside solenoid





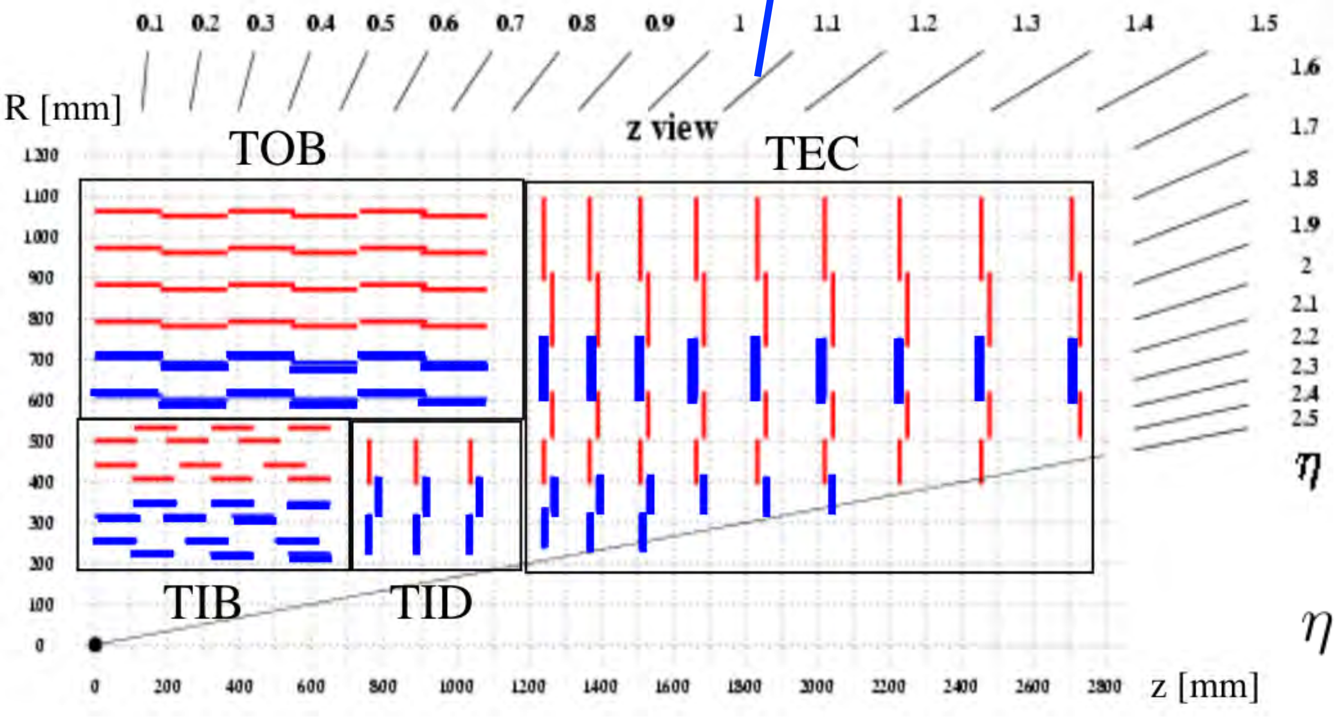
Silicon Tracker



- **Purpose:** High resolution tracks for p_T and vertex measurement & matching
- Inner Pixel Detector: 66M channels, fine grain resolution
- Outer Strip Detector: 9.6M channels, coarser granularity

Barrel resolution:

$$\frac{\delta p_T}{p_T} = (\frac{p_T}{1TeV}15\%) \oplus 0.5\%$$

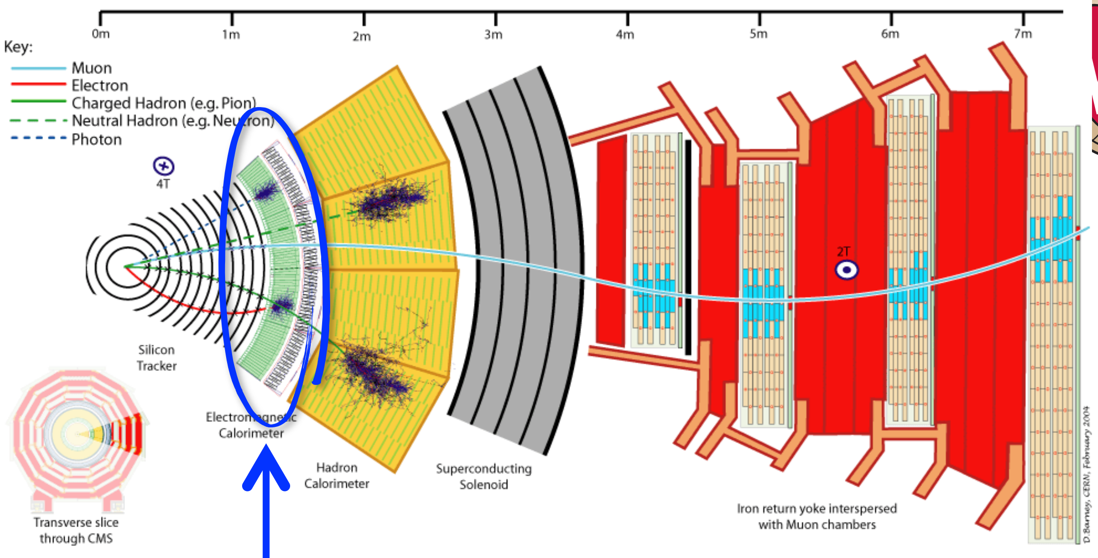


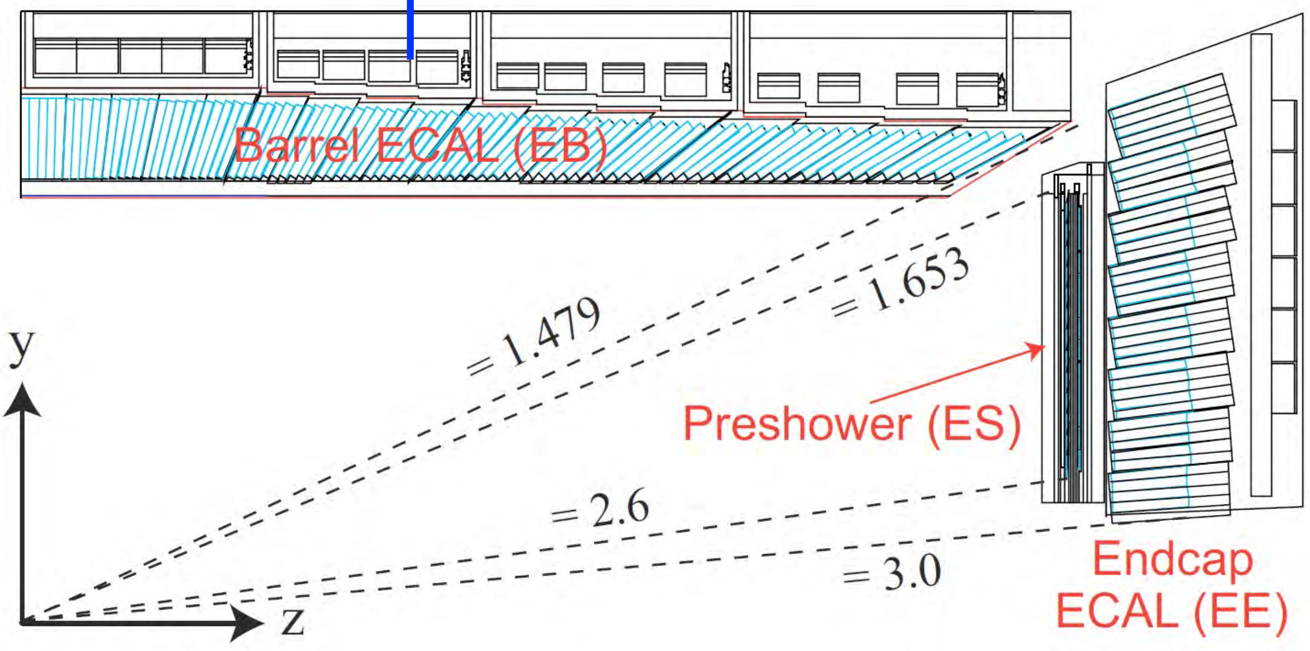
pseudo-rapidity

$$\eta = -ln\left[tan\left(\frac{\theta}{2}\right)\right]$$

Electromagnetic Calorimeter

- **Purpose:** measure position and energy of $e^{+/-}$ and photons
- $PbWO_4$ crystal
- Overall resolution
  - Electrons: 0.4% (0.8%) in the barrel (endcaps)
  - Photons: roughly ~2%(~4%) in the barrel (endcaps)



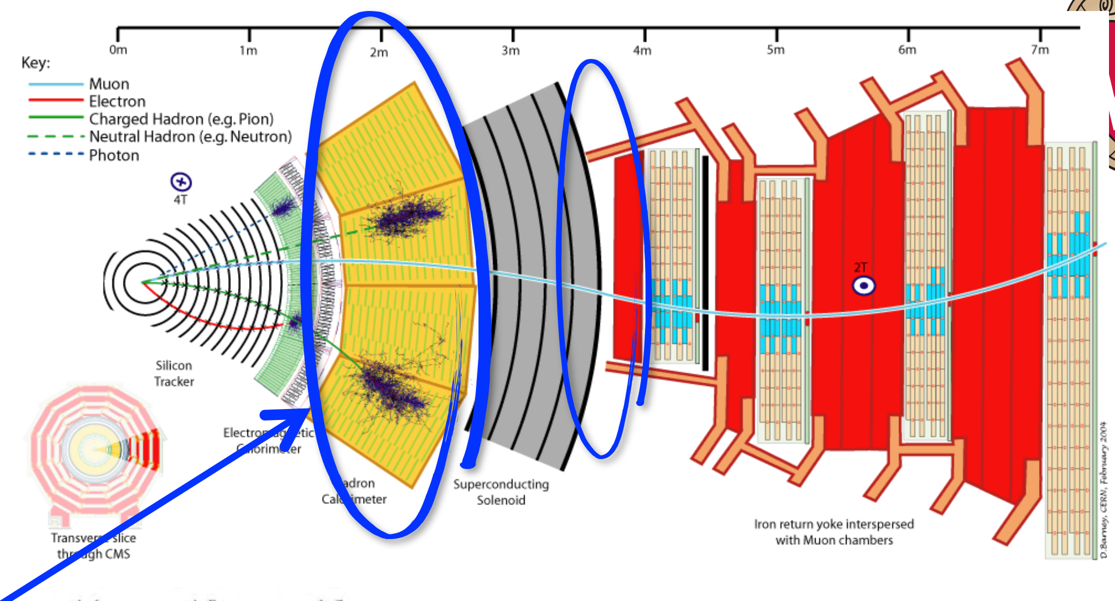


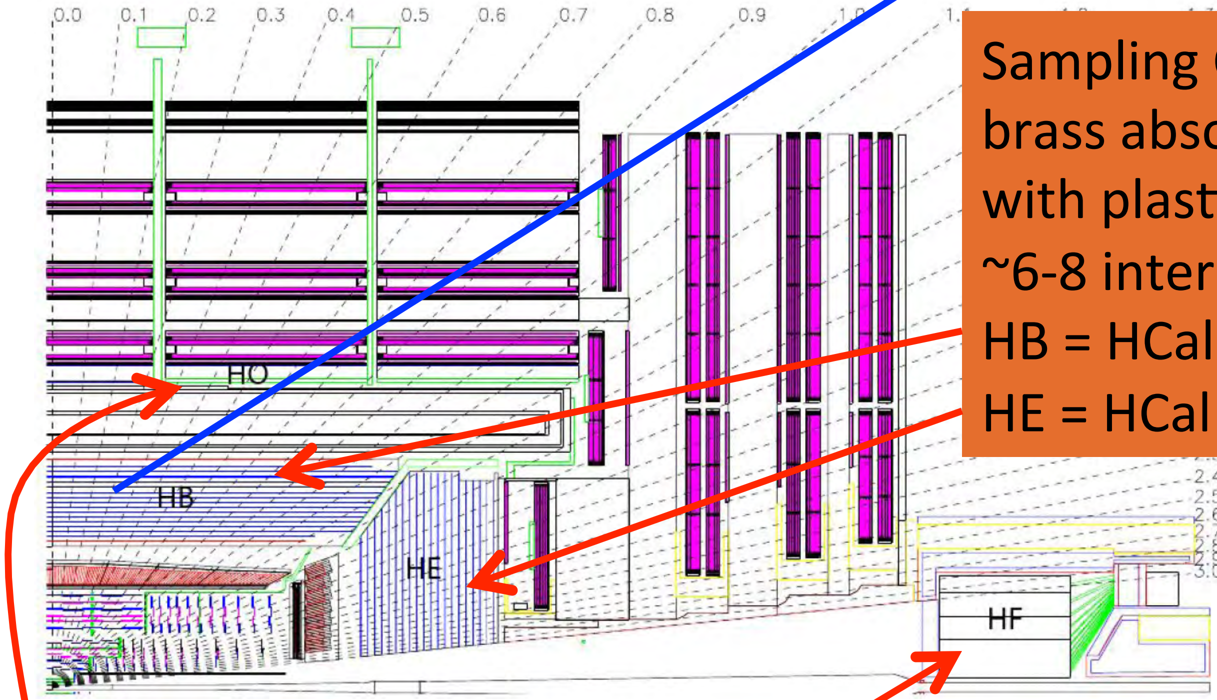
ECAL	
Density	$8.28\ g/cm^3$
$PbWO_4$ Rad. Len.	8.9 mm
Barrel Len.	230 mm
End Cap Len.	220 mm

Resolution: $$\frac{\sigma}{E} = \frac{2.8\%}{\sqrt{E/\mathrm{GeV}}} \oplus \frac{12\%}{E/\mathrm{GeV}} \oplus 0.3\%$$

Hadronic Calorimeter

- **Purpose:** Measure energy of neutral and charged long lived hadrons → Tau energy & missing transverse energy





Sampling Calo. → layers of brass absorber interspersed with plastic scintillator

~6-8 interaction lengths

HB = HCal Barrel

HE = HCal End Cap

Forward HCal made of steel with embedded quartz fibers for radiation hardness

- Measures EM rich jets beyond ECal range

HCal Outer (HO) uses solenoid as preshower

Resolution

- Barrel and End Cap

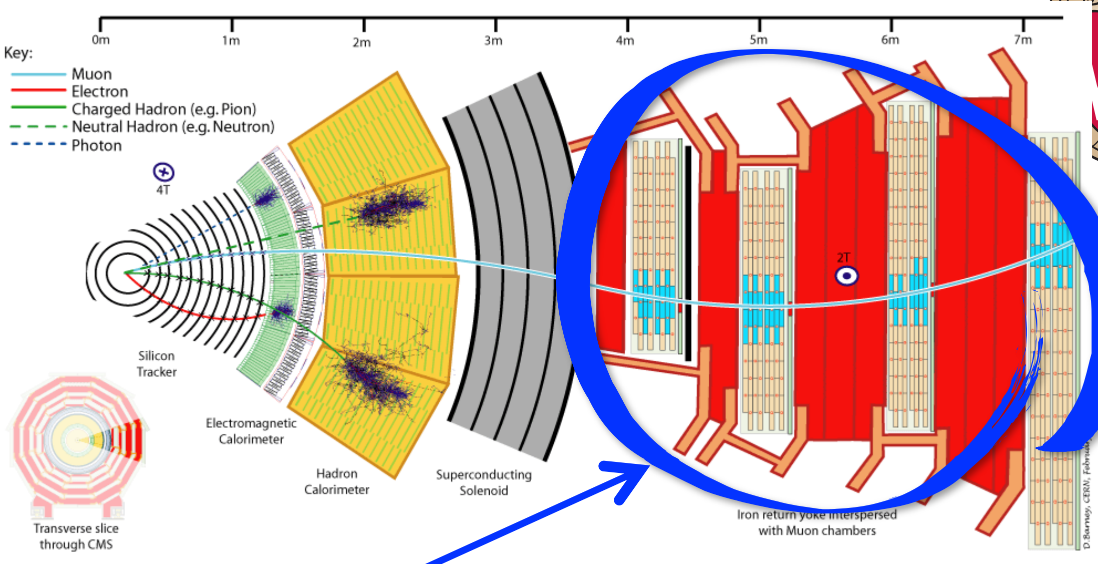
$$\frac{\sigma}{E} = \frac{115\%}{\sqrt{E}} \oplus 5.5\%$$

- Forward

$$\frac{\sigma}{E} = \frac{280\%}{\sqrt{E}} \oplus 11\%$$

Muon System

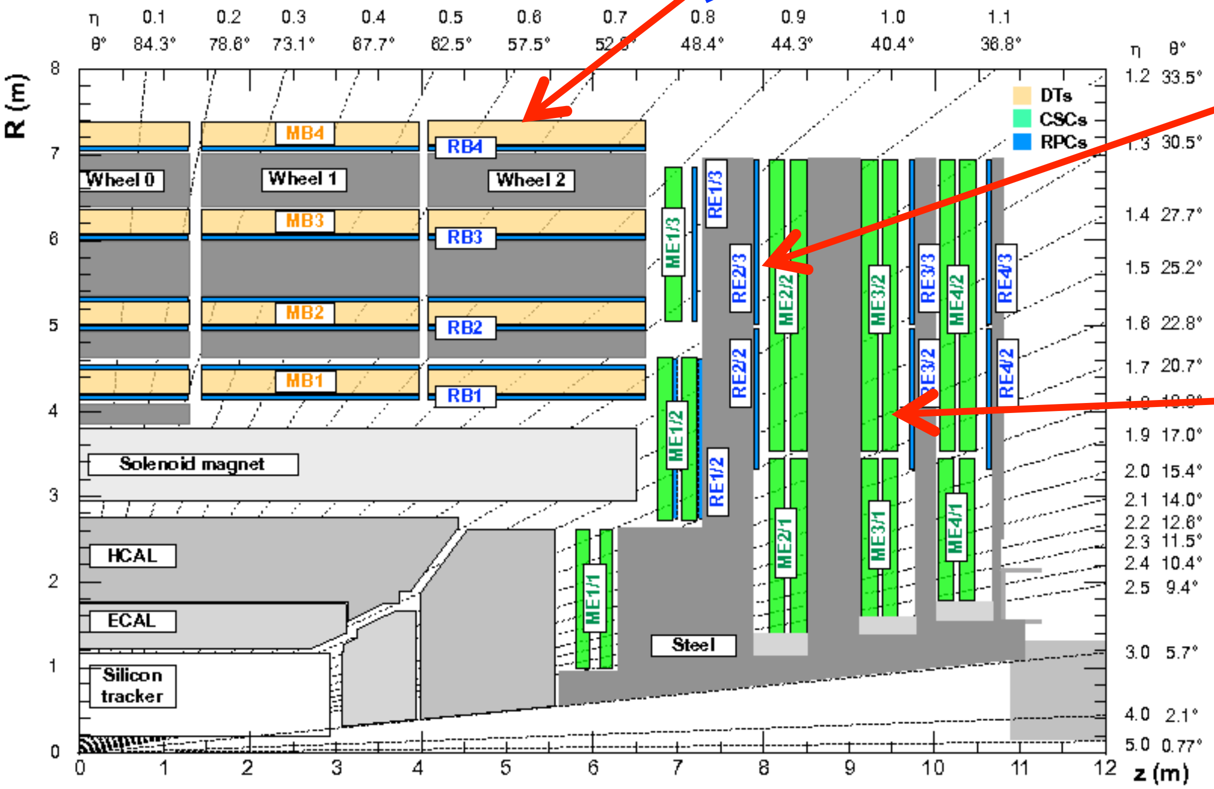
- **Purpose:** ID & measure p_T of muons, provide quick trigger info
- Low p_T muons (< 100 GeV):
  - Silicon tracker → p_T
  - Res. 6-20%
- High p_T muons (> 100 GeV):
  - Tracker + muon system → p_T
  - Res. 15-35%



Drift Tubes - Res. 80-120 µm & 3 ns

Resistive Plate Chambers - Res. 0.8-1.2 cm & 1 ns; timing and x-y correlation (pad geometry)

Cathode Strip Chambers - Res. 40-150 µm & 3 ns

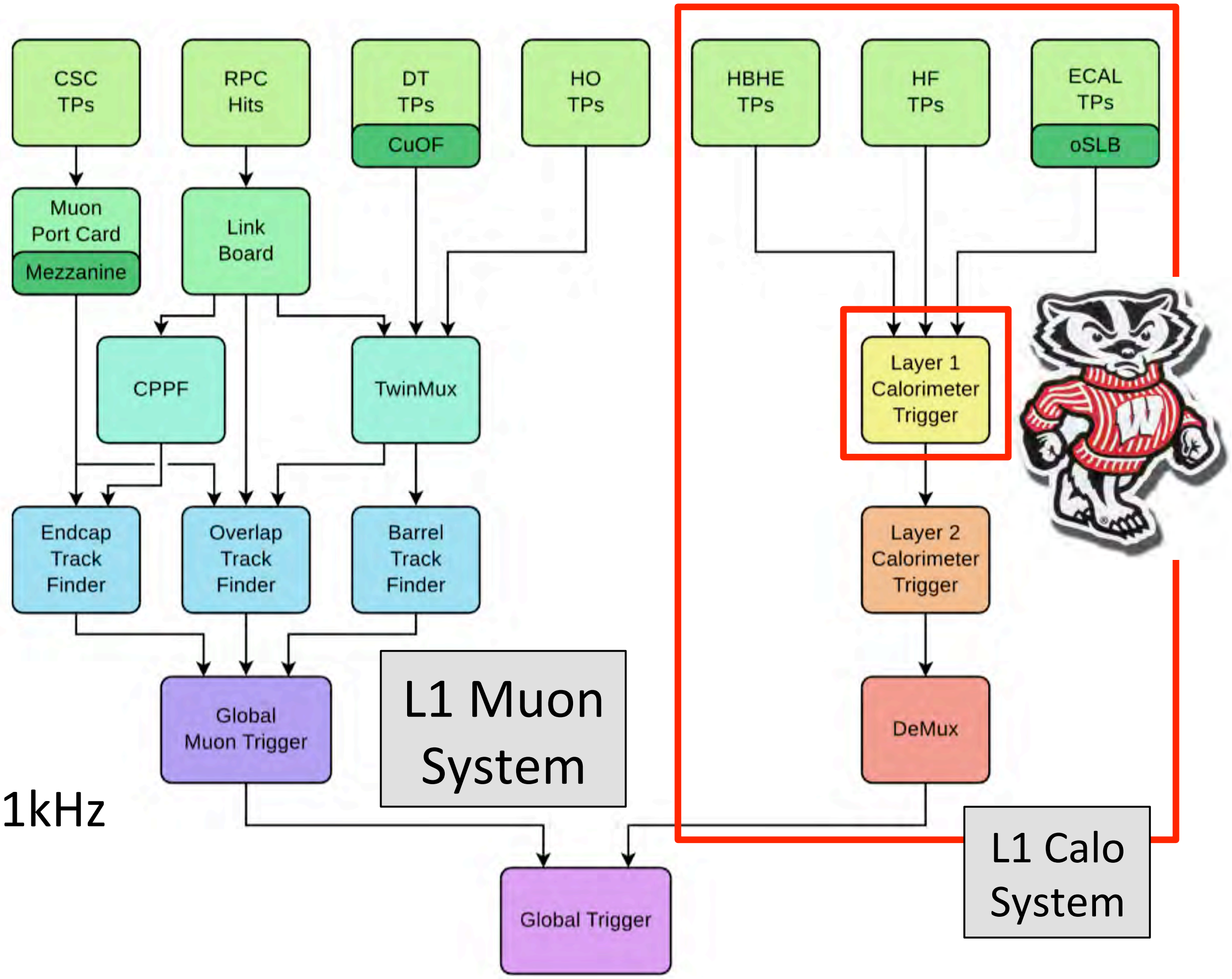


All muon chambers are used in the trigger system

Trigger

- Initial 40 MHz crossing rate, ~27 collisions/crossing → 1 GHz event rate
- Each event ~ 1 Mb of disk space
- **Purpose:** Keep the most interesting events so data is manageable to process, analyze, & store

- **Level-1 Trigger:**
  - High-speed custom hardware
  - Partial data from Calo and muon systems
  - Specialized algorithms
  - **4 μs** available to made decision



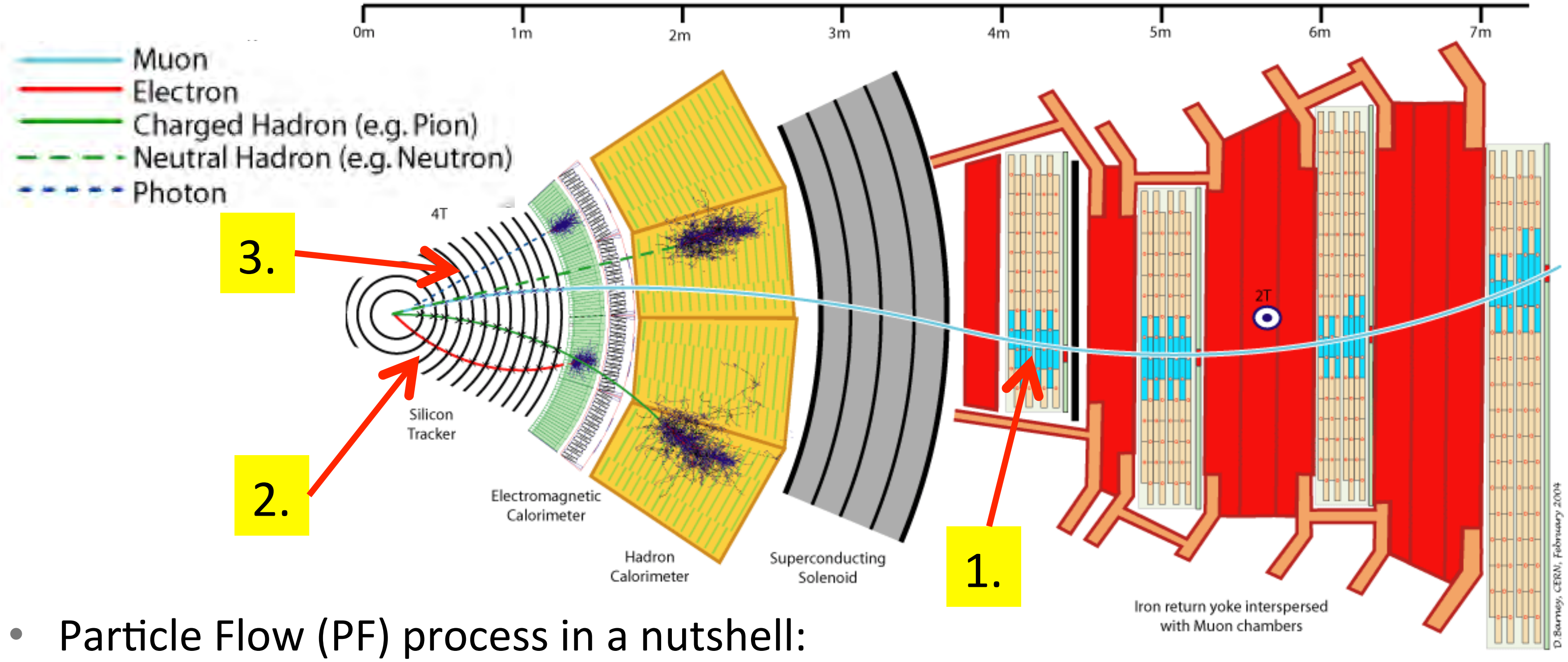
L1 Muon System

L1 Calo System

- **High Level Trigger (HLT):**
  - Reduces from ~100 kHz to ~1kHz
  - General purpose CPU farm
  - Software based system
  - Algos have access to entire event
  - On average event decisions must be made in **200 μs**

Particle Flow Reconstruction

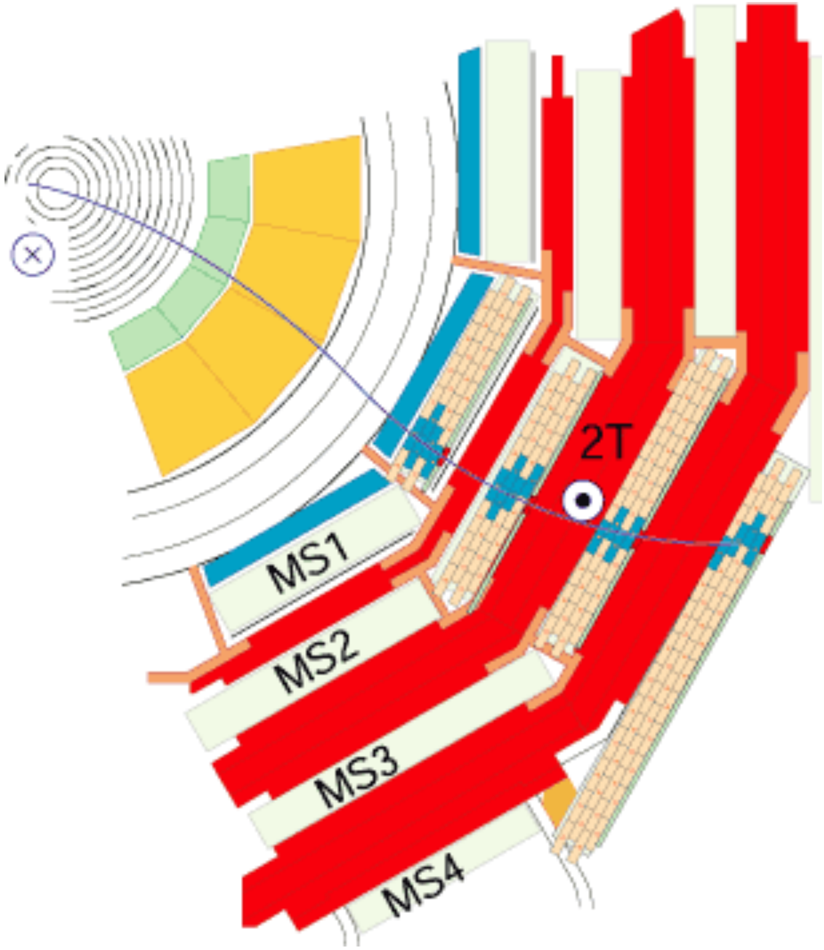
- Information from detector systems utilized simultaneously to provide best reconstruction of select physics objects
  - Improves track & p_T resolution & particle identification
  - Primary objects: muons, electrons, charged & neutral hadrons & photons



- Particle Flow (PF) process in a nutshell:
  1. Muon system tracks are matched to tracks in inner tracker
  2. Remaining inner tracker tracks associated with ECal energy deposits (electrons) and HCal energy deposits (charged hadrons)
  3. Remaining energy deposits are likely from neutral sources, ECal (photons) & HCal (neutral hadrons)

Muons

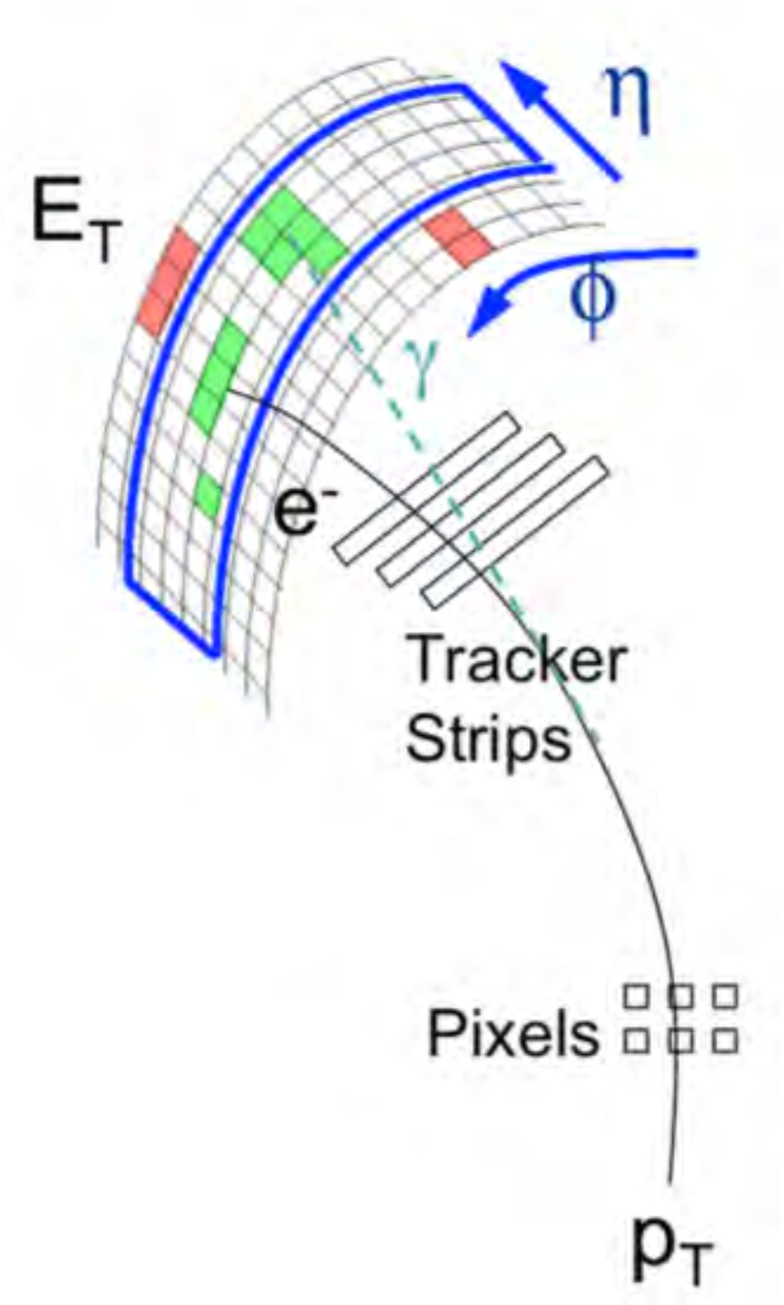
- Particle Flow muons:
  - Tracker muon – found and reconstructed by silicon tracker, tagged by muon system
  - Standalone muon – found and reconstructed solely by muon system
  - Global muon – found and reconstructed in muon and tracker systems



- Analyses use 2 muon definitions
  - “loose muon” = tracker || global
  - “tight muon” = “loose muon” + additional track quality cuts
- Additional muon requirements
  - Originate close to primary vertex
  - Small X^2 of track to eliminate punch through to the muon system

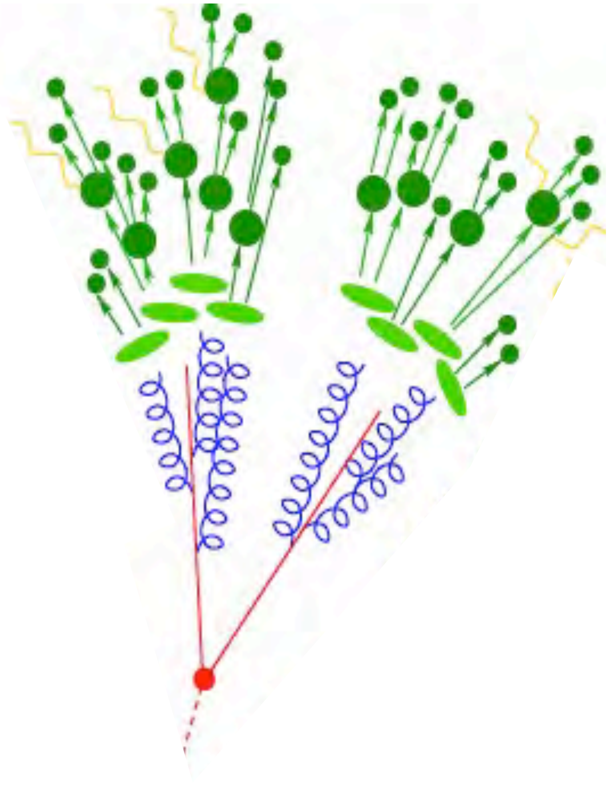
Electrons

- Constructed from ECal energy deposits matching nearby tracks
- ECal supercluster elongated in φ to contain bremsstrahlung photons
- Track must originate close to primary vertex
- Multivariate based discriminants used to select electrons in analysis
    - Inputs: track criteria, H/E, ECAL energy deposit shape information
    - Two working points used, 90% efficient and 80% efficient



Jets & Missing Transverse Energy

- Jets are collections of energy deposits and tracks grouped together within a defined radius, ΔR = 0.4

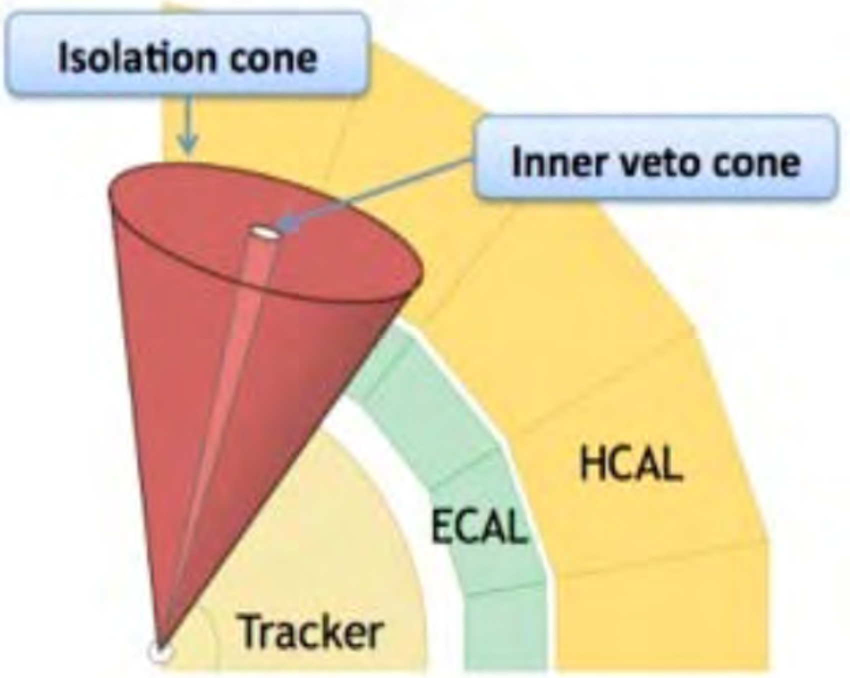


- Missing Transverse Energy (MET) is calculated as the negative of the vector sum of the transverse momentum of all particles in the event
    - Only particles associated with the primary vertex
- Usually results from neutrinos or other particles escaping detection

- In this thesis MET is intended to capture the contributions of the neutrinos from the tau lepton decays

Lepton Isolation

*Artists interpretation on cone size

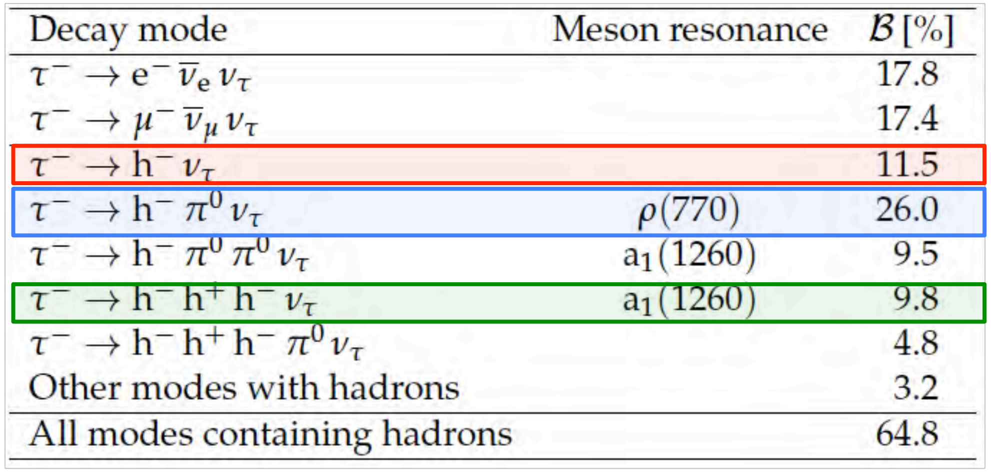


- Requiring isolated leptons strongly rejects QCD backgrounds
- Sum the energy deposited within ΔR < 0.3 (0.4) electrons (muons)
  - Pile-up (PU) corrections are applied
  - Charged hadron PU already removed with vertex cuts
  - Neutral hadron pile-up approximated as 0.5 x charged hadron PU and is called Δβ

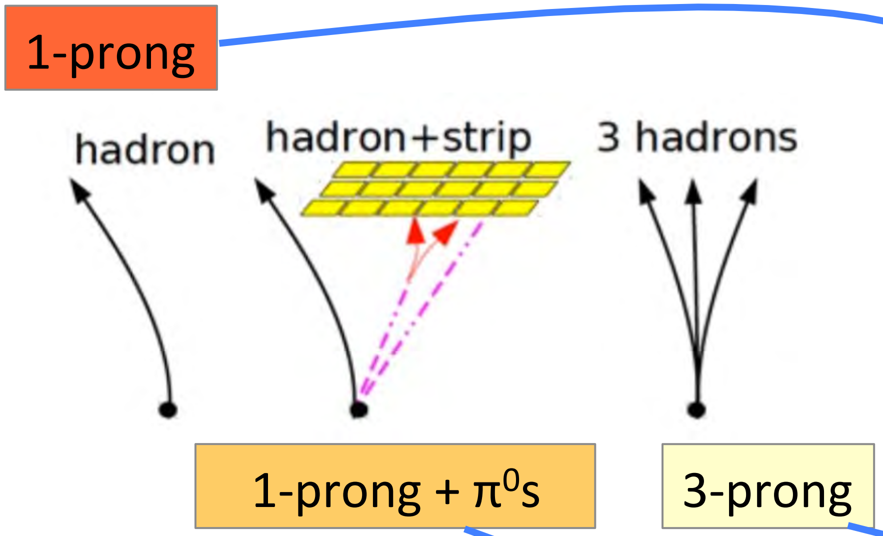
$$I_{rel} = \frac{\Sigma p_T(charged) + max[\Sigma E_T(neutral) + \Sigma E_T(photon) - \Delta\beta, 0]}{p_T(\mu\ or\ e)}$$

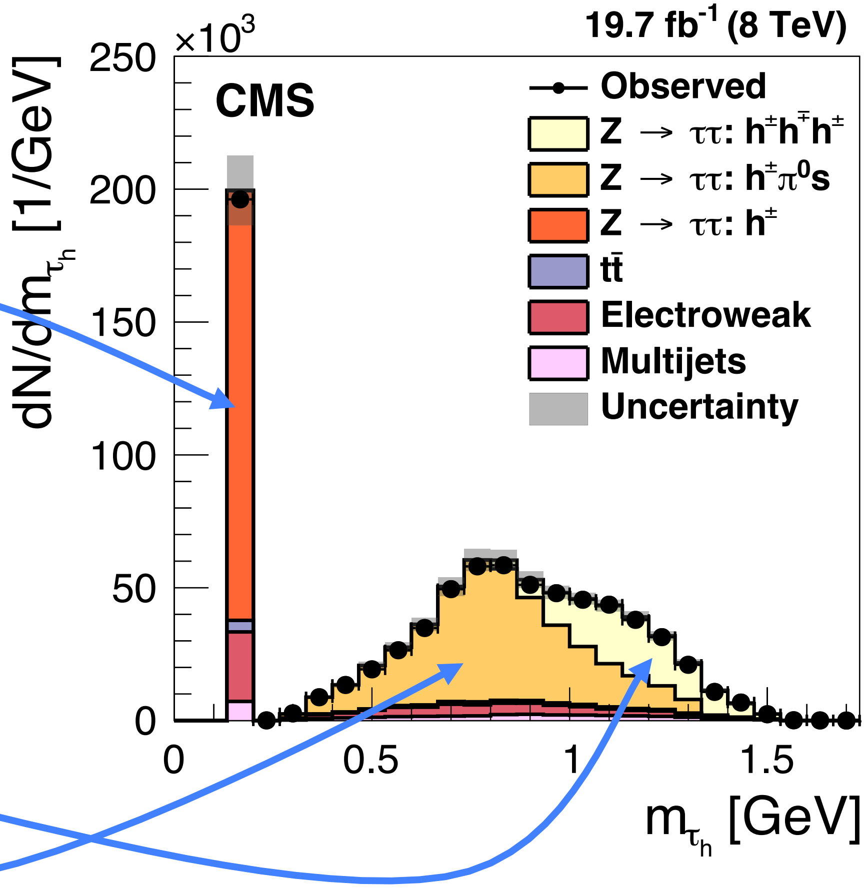
Tau Decays and Reconstruction

- Tau leptons decay either leptonically (35%) or hadronically (65%)
- Hadronic decays are reconstructed from charged and neutral jet constituents in CMS's Particle Flow algorithm
  - 1-prong, 1-prong + π^0, 3-prong
- Hadrons Plus Strips algorithm:
  - Can ID each τ_h decay mode
  - Exploits intermediate resonances ρ(770) and a_1(1260)



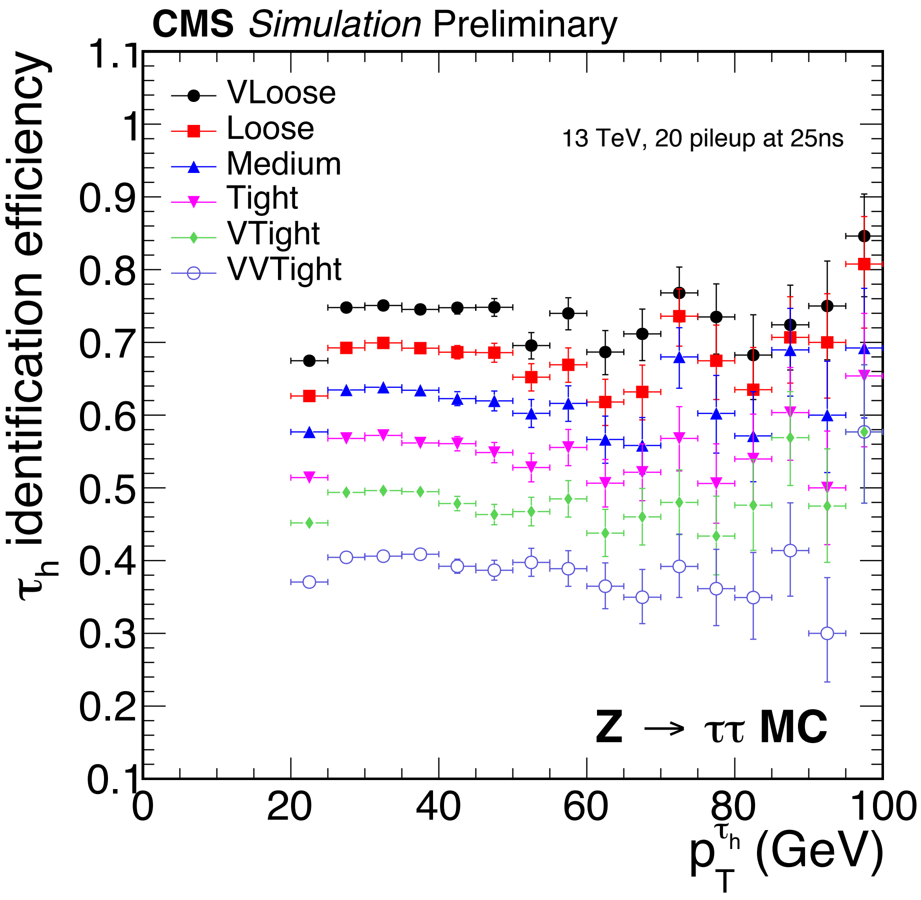
Decay mode	Meson resonance	$\mathcal{B}$ [%]
$\tau^- \to e^- \overline{\nu}_e \nu_\tau$		17.8
$\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$		17.4
$\tau^- \to h^- \nu_\tau$		11.5
$\tau^- \to h^- \pi^0 \nu_\tau$	$\rho(770)$	26.0
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	9.5
$\tau^- \to h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other modes with hadrons		3.2
All modes containing hadrons		64.8

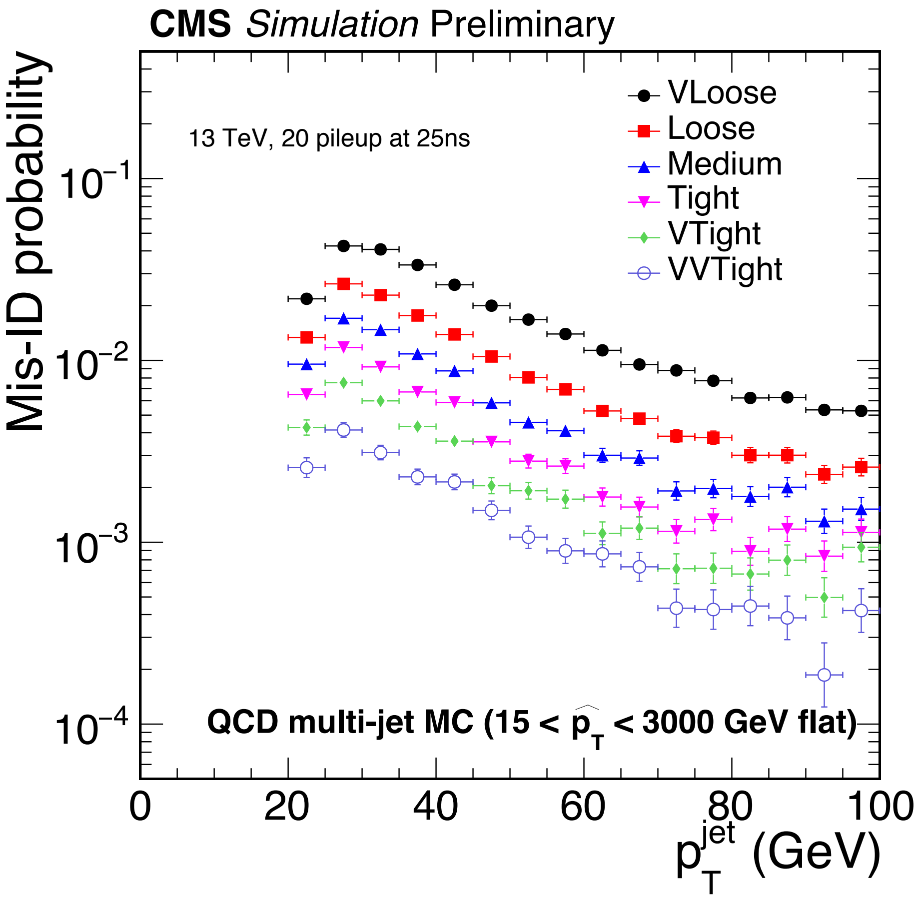




Tau Identification

- We need best tau identification with strong jet rejection
- Analyses use a multivariate-based ID that has large performance gains compared to Run 1 cut-based ID
  - Inputs: Isolation (pT sums), τ lifetime, distribution of signal and isolation candidates within the jet in ΔR, Δη, Δφ, and e/γ multiplicies

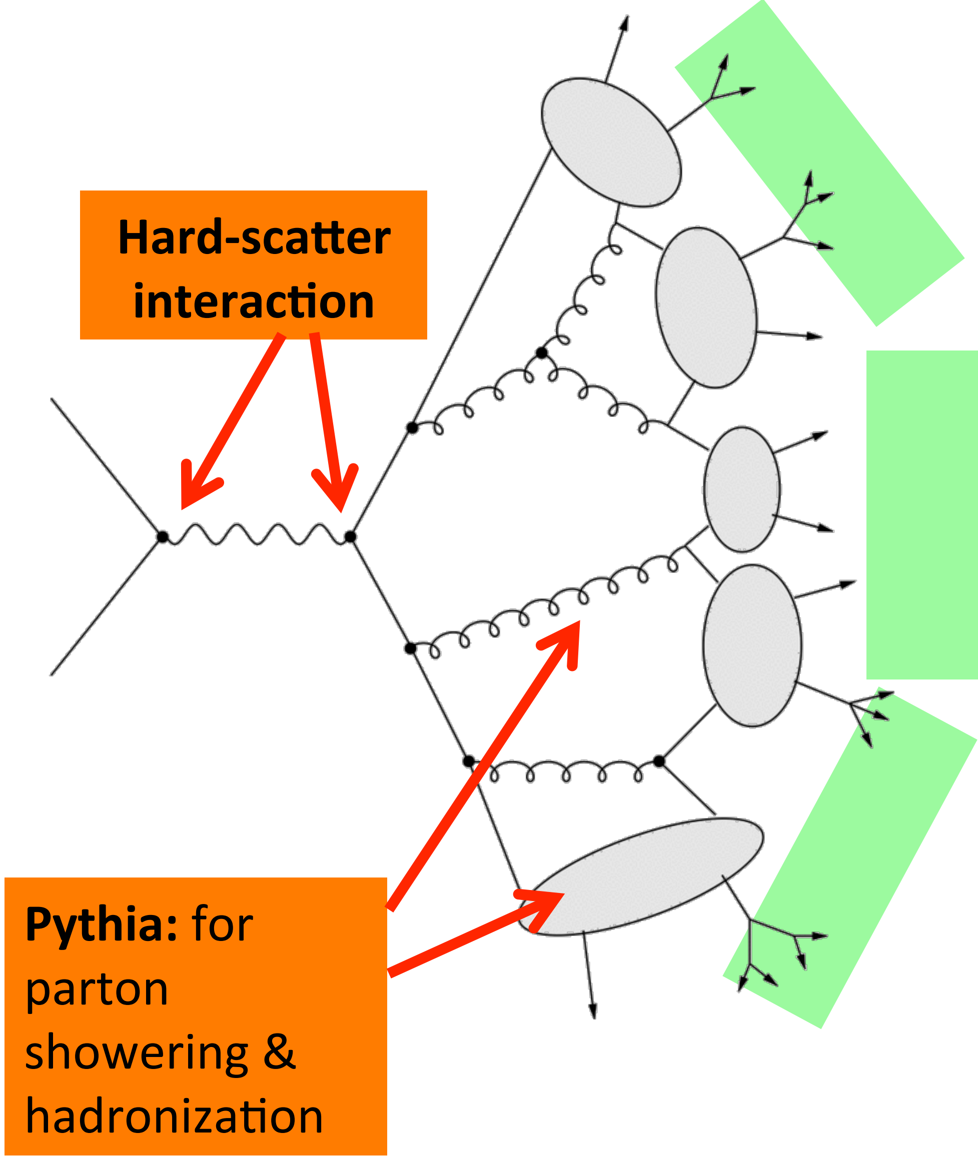




Monte Carlo

- **Purpose:** Create accurate simulated data to model expectations

Sample	Hard-scatter Interaction
All Higgs Samples	Powheg
ZZ, WZ, tt, single-top	Powheg
Z+jets, W+jets, ttV, VVV	Madgraph
ggZZ	MCFM
Min Bias	Pythia



Hard-scatter interaction

GEANT: particle interactions with detector

Pythia: for parton showering & hadronization

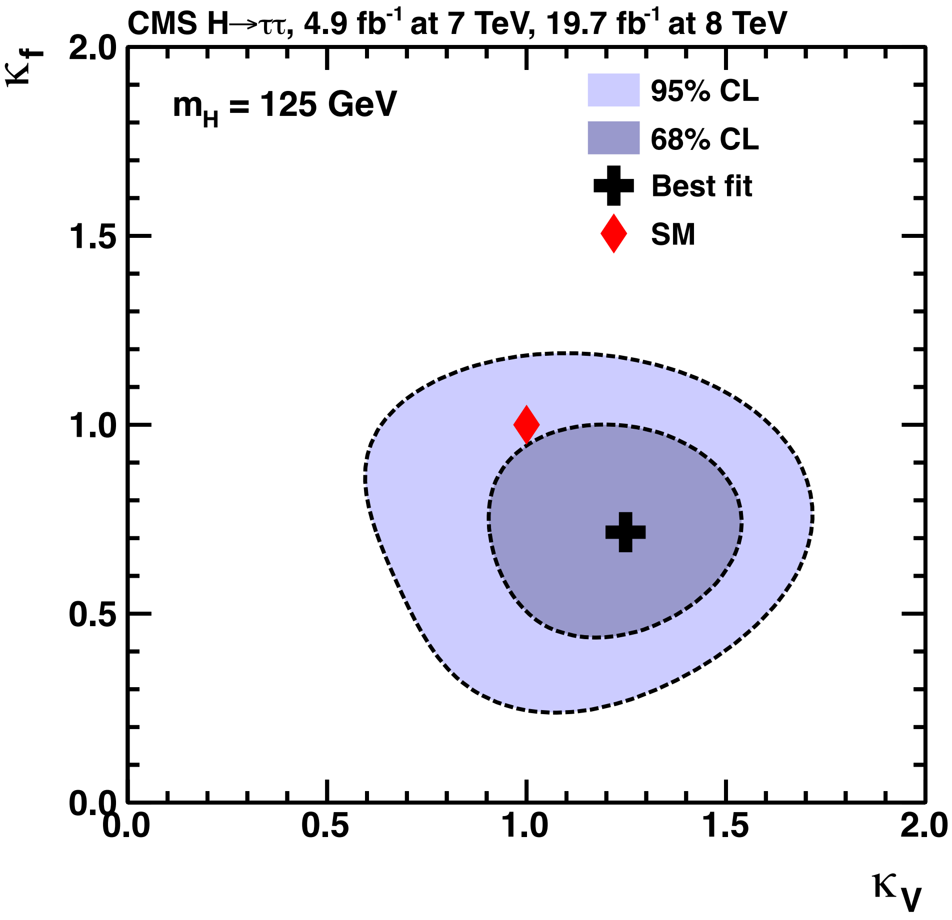
minimum bias data → simulate pile-up

Event reconstruction software: identical to actual CMS measured data

Previous H → ττ Results

Prior to Run-II, the most sensitive results came from the CMS and ATLAS collaborations using 7 & 8 TeV center-of-mass Run-I data

Experiment	Observed	Expected
CMS	3.2	3.7
ATLAS	4.5	3.4
CMS + ATLAS	5.5	5.0



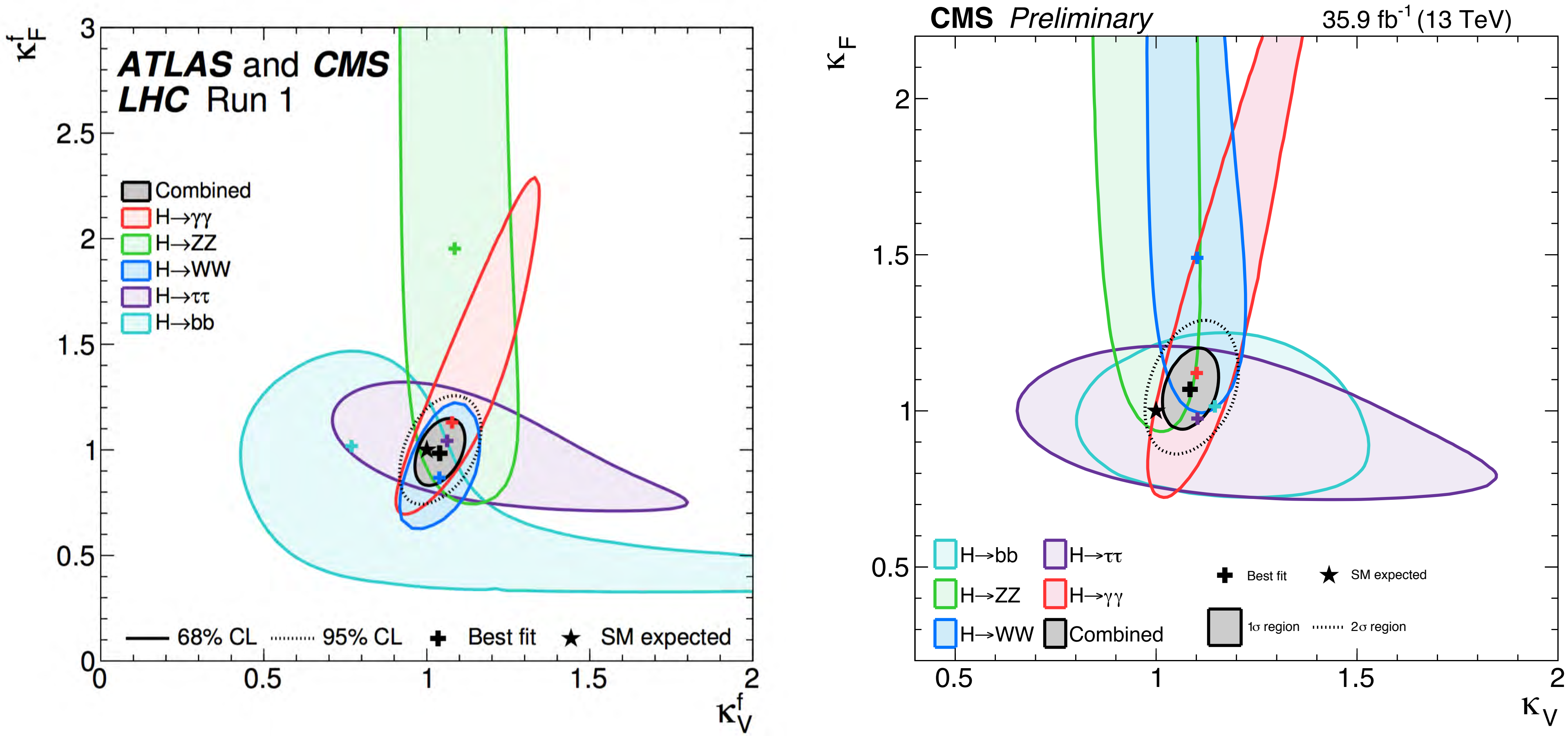
Left: Higgs boson couplings measurement from the CMS Run-I analysis. κ are ratios with respect to SM couplings, κ = 1 = SM expectation

- κ_F = coupling to fermions
- κ_V = coupling to vector bosons

- CMS – doi:10.1007/JHEP05(2014)104
- ATLAS – doi:10.1007/JHEP04(2015)117
- CMS + ATLAS – doi:10.1007/JHEP08(2016)045

Higgs Boson Couplings

- The same type of Higgs boson couplings distribution was produced for the Run-I combination (left). The κ_F and κ_V are decay process specific, (κ^f_F , κ^f_V)
- Run-II CMS results (right) using a common κ_F and κ_V scaling



Left: CMS + ATLAS – doi:10.1007/JHEP08(2016)045
Right: CMS Run-II - CMS-PAS-HIG-17-031

Gluon fusion and VBF analysis

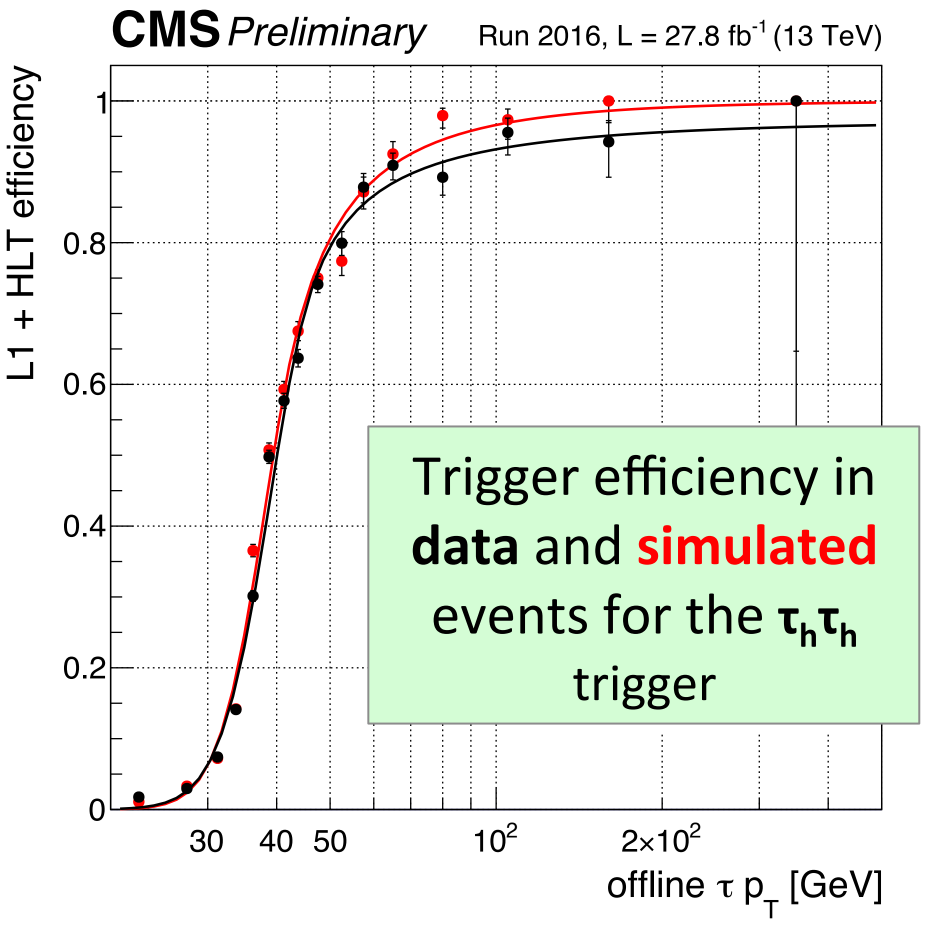
Analysis Overview:

- Targeting the two leading Higgs production mechanisms by studying di-tau final states
- Four different decay channels depending on tau decays:
  - **$e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$, $e\mu$**
- Three different signal region categories:
  - 0-Jet, VBF, Boosted

Analysis Triggers

- This analysis uses a variety of triggers targeting the different Higgs boson production mechanisms and Higgs decay final states
- All triggers **select events based on electrons, muons, or τ_h**

Channel	Trigger $(p_T/\|\eta\|)$ Req.
$\mu\tau_h$	$\mu(22/2.4)$
	$\mu(22/2.1)$
	$\mu(19/2.1)\,\&\,\tau_h(20/2.1)$
$e\tau_h$	$e(25/2.1)$
$\tau_h\tau_h$	$\tau_h(35/2.1)\,\&\,\tau_h(35/2.1)$
	$\tau_h(35/2.1)\,\&\,\tau_h(35/2.1)$
$e\mu$	$e(12/2.5)\,\&\,\mu(23/2.4)$
	$e(12/2.5)\,\&\,\mu(23/2.4)$
	$e(23/2.5)\,\&\,\mu(8/2.4)$
	$e(23/2.5)\,\&\,\mu(8/2.4)$



Trigger efficiency in **data** and **simulated** events for the **$\tau_h\tau_h$** trigger

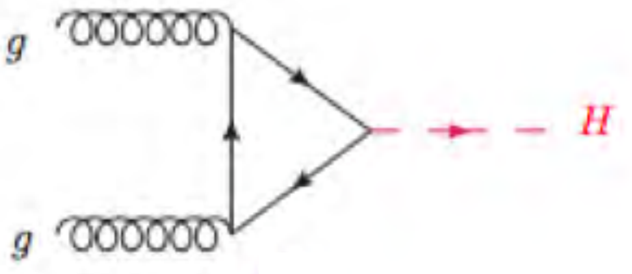
ZH **trigger selection requirements**

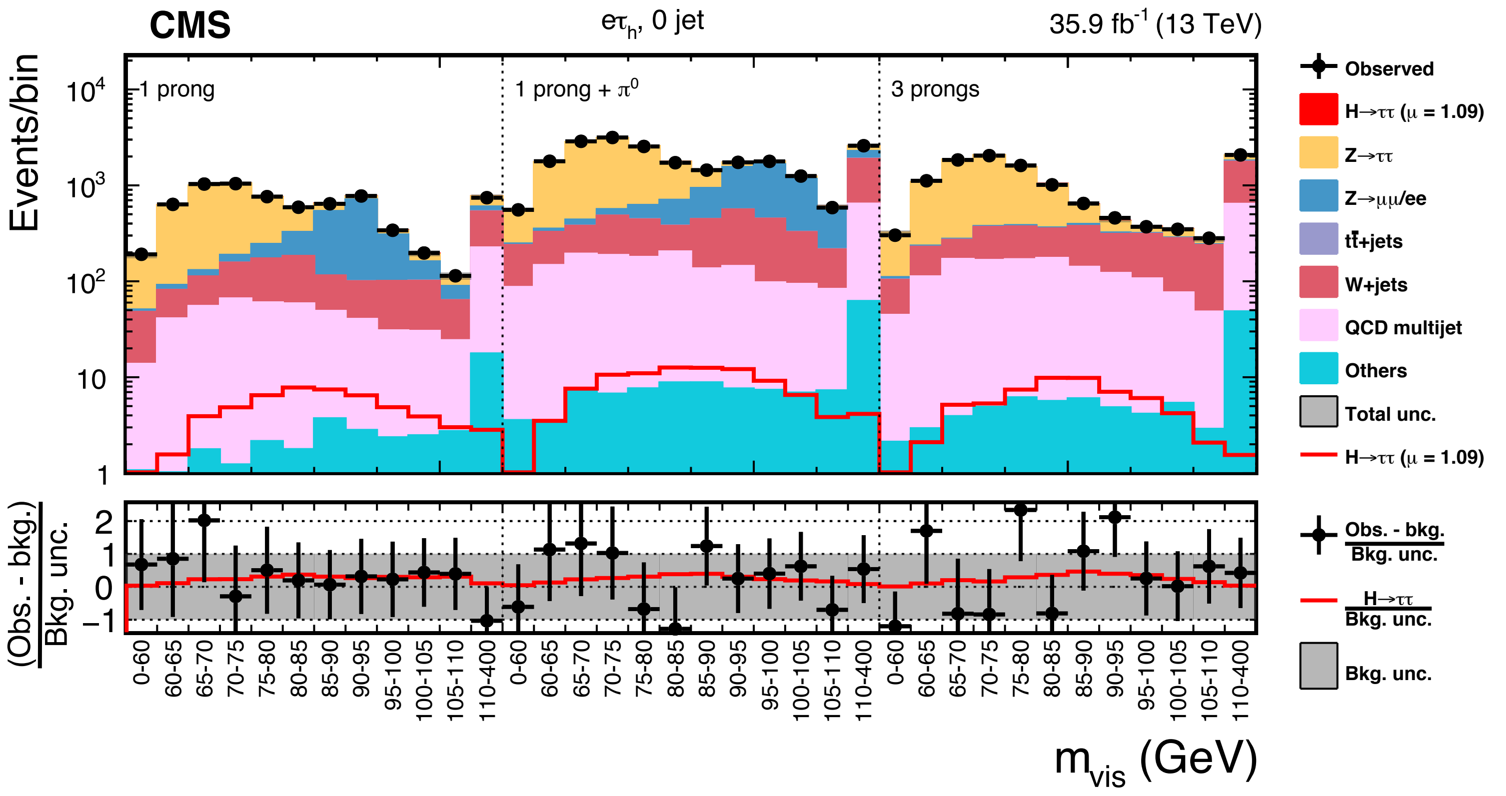
Final State	Trigger (p_T/η)	Lepton Selection: p_T
$ee\mu\tau_h$ $eee\tau_h$ $ee\tau_h\tau_h$ $eee\mu$	$e(23/2.5)\,\&\,e(12/2.5)$, or $e(27/2.5)$	$p_T^e > 24\,\&\,p_T^e > 13$, or $p_T^e > 28$
$\mu\mu\mu\tau_h$ $\mu\mu e\tau_h$ $\mu\mu\tau_h\tau_h$ $\mu\mu e\mu$	$\mu(17/2.4)\,\&\,\mu(8/2.4)$, or $\mu(24/2.4)$	$p_T^\mu > 18\,\&\,p_T^\mu > 10$, or $p_T^\mu > 25$

WH triggers are the single electron and single muon from above

Signal Extraction – 0-jet

	0-jet	VBF	Boosted
		Selection	
$\tau_h\tau_h$	No jet	$\geq$2 jets, $p_T^{\tau\tau} > 100$ GeV, $\Delta\eta_{jj} > 2.5$	Others
$\mu\tau_h$	No jet	$\geq$2 jets, $m_{jj} > 300$ GeV, $p_T^{\tau\tau} > 50$ GeV, $p_T^{\tau_h} > 40$ GeV	Others
$e\tau_h$	No jet	$\geq$2 jets, $m_{jj} > 300$ GeV, $p_T^{\tau\tau} > 50$ GeV	Others
$e\mu$	No jet	2 jets, $m_{jj} > 300$ GeV	Others
		Observables	
$\tau_h\tau_h$	$m_{\tau\tau}$	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$\mu\tau_h$	τ_h decay mode, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$e\tau_h$	τ_h decay mode, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$e\mu$	p_T^{μ}, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$





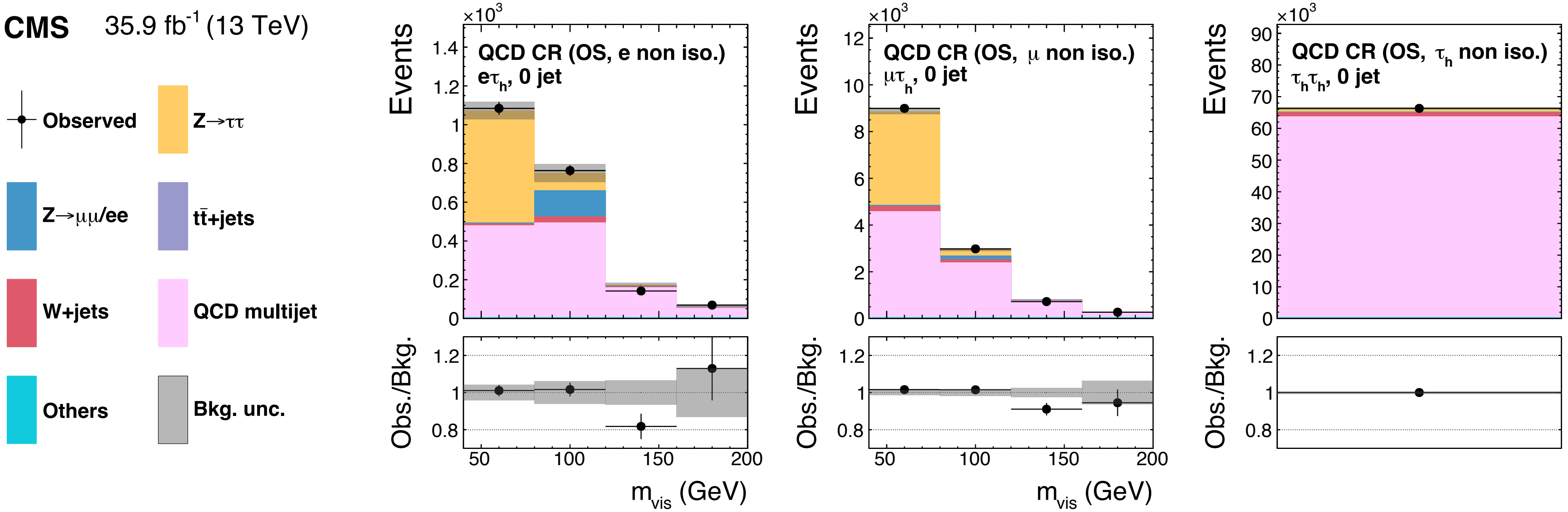
Main Backgrounds & Control Regions

Z→ee/μμ/ττ:

- LO MC sample
- Corrections p_T^Z, $mass_Z$, & m_{jj} from Z→μμ CR

QCD multijet:

- Fully data driven
- $e\tau_h/\mu\tau_h$ – taken from region with same-charge taus
- $\tau_h\tau_h$ – taken from region with opposite charge, anti-isolated taus
- CR included in simultaneous fit



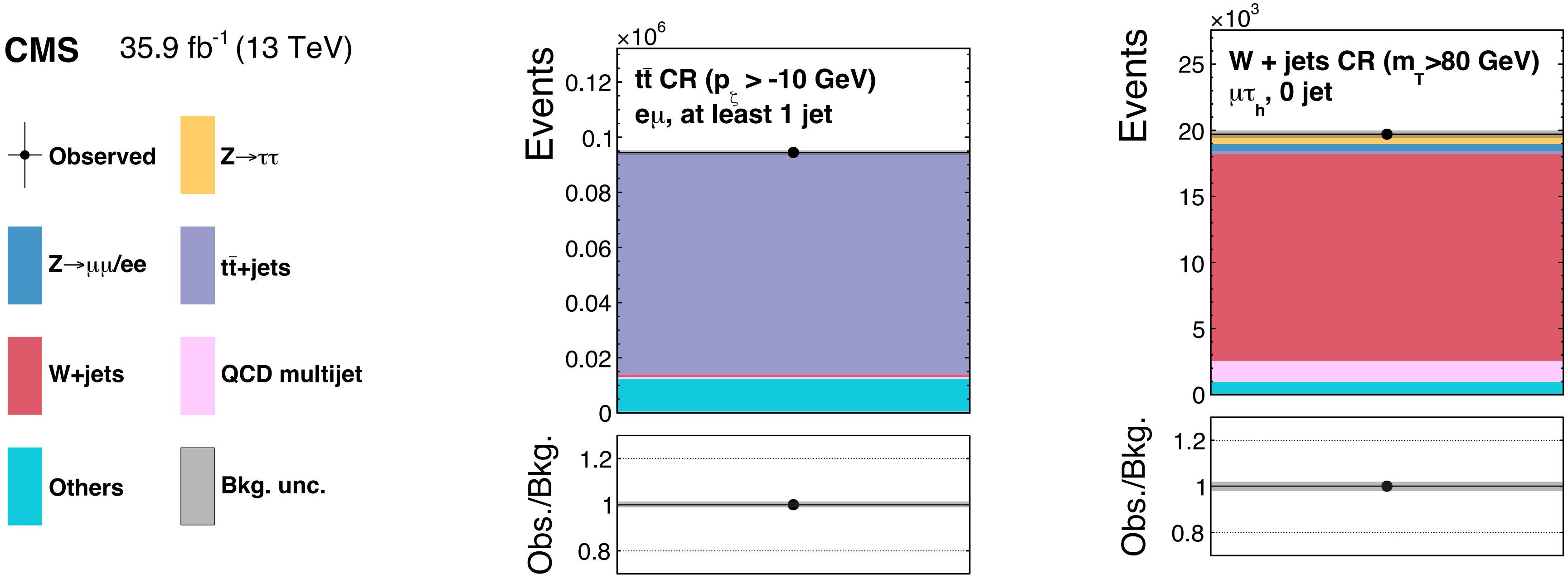
Main Backgrounds

ttbar:

- MC sample
- Single CR in eμ ttbar enriched CR

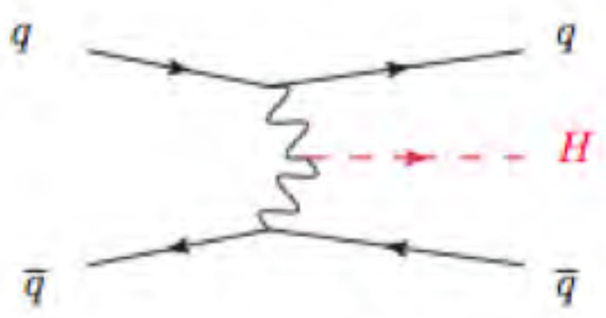
W+Jets:

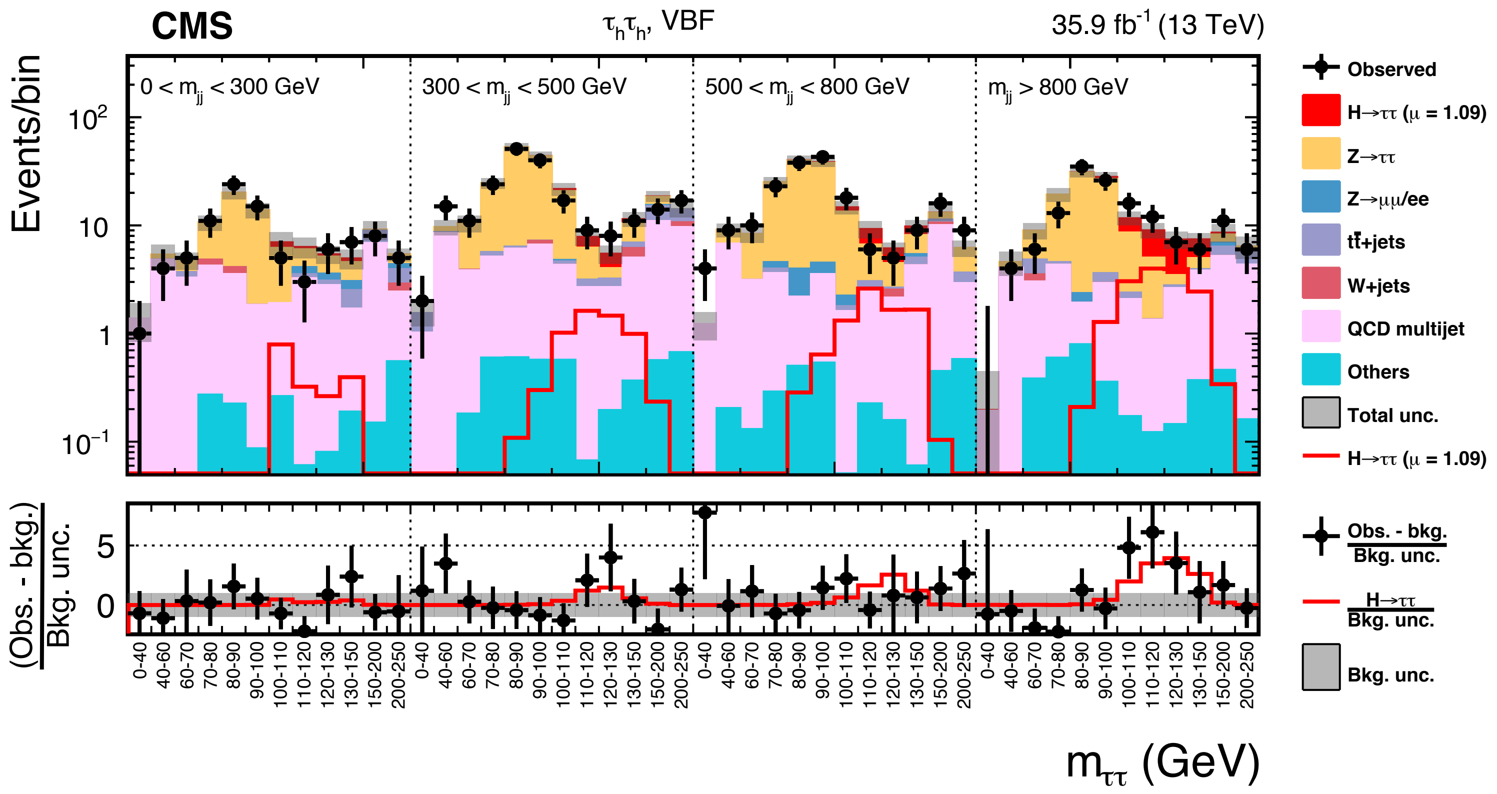
- Shape from LO MC
- $e\tau_h/\mu\tau_h$ – Normalized with data in high m_T CR [mT(lep,MET)>80 GeV]
- $\tau_h\tau_h/e\mu$ – from MC with corrections



Signal Extraction – VBF

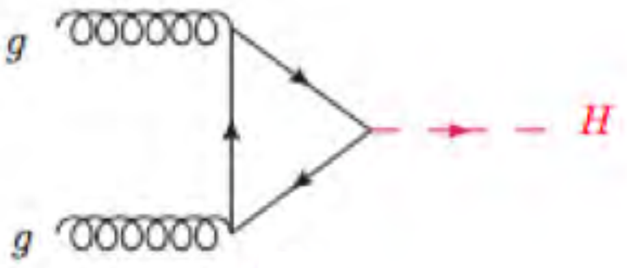
	0-jet	VBF	Boosted
		Selection	
$\tau_h\tau_h$	No jet	$\geq$2 jets, $p_T^{\tau\tau} > 100\,\text{GeV}$, $\Delta\eta_{jj} > 2.5$	Others
$\mu\tau_h$	No jet	$\geq$2 jets, $m_{jj} > 300\,\text{GeV}$, $p_T^{\tau\tau} > 50\,\text{GeV}$, $p_T^{\tau_h} > 40\,\text{GeV}$	Others
$e\tau_h$	No jet	$\geq$2 jets, $m_{jj} > 300\,\text{GeV}$, $p_T^{\tau\tau} > 50\,\text{GeV}$	Others
$e\mu$	No jet	2 jets, $m_{jj} > 300\,\text{GeV}$	Others
		Observables	
$\tau_h\tau_h$	$m_{\tau\tau}$	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$\mu\tau_h$	τ_h decay mode, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$e\tau_h$	τ_h decay mode, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$e\mu$	p_T^{μ}, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$

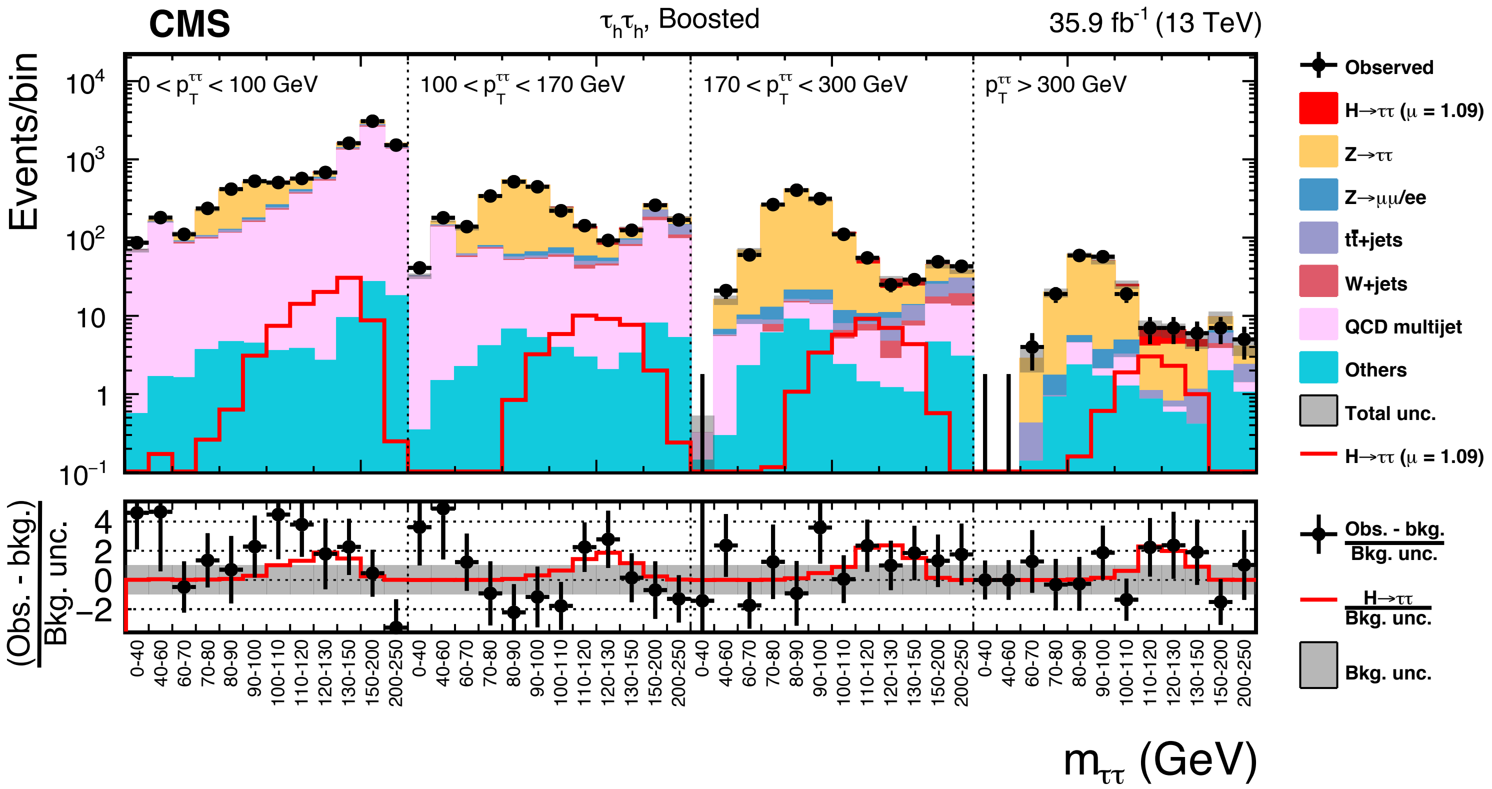




Signal Extraction – Boosted

	0-jet	VBF	Boosted
		Selection	
$\tau_h\tau_h$	No jet	$\geq$2 jets, $p_T^{\tau\tau} > 100\,\text{GeV}$, $\Delta\eta_{jj} > 2.5$	Others
$\mu\tau_h$	No jet	$\geq$2 jets, $m_{jj} > 300\,\text{GeV}$, $p_T^{\tau\tau} > 50\,\text{GeV}$, $p_T^{\tau_h} > 40\,\text{GeV}$	Others
$e\tau_h$	No jet	$\geq$2 jets, $m_{jj} > 300\,\text{GeV}$, $p_T^{\tau\tau} > 50\,\text{GeV}$	Others
$e\mu$	No jet	2 jets, $m_{jj} > 300\,\text{GeV}$	Others
		Observables	
$\tau_h\tau_h$	$m_{\tau\tau}$	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$\mu\tau_h$	τ_h decay mode, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$e\tau_h$	τ_h decay mode, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$
$e\mu$	p_T^{μ}, m_{vis}	m_{jj}, $m_{\tau\tau}$	$p_T^{\tau\tau}$, $m_{\tau\tau}$





Associated production analysis

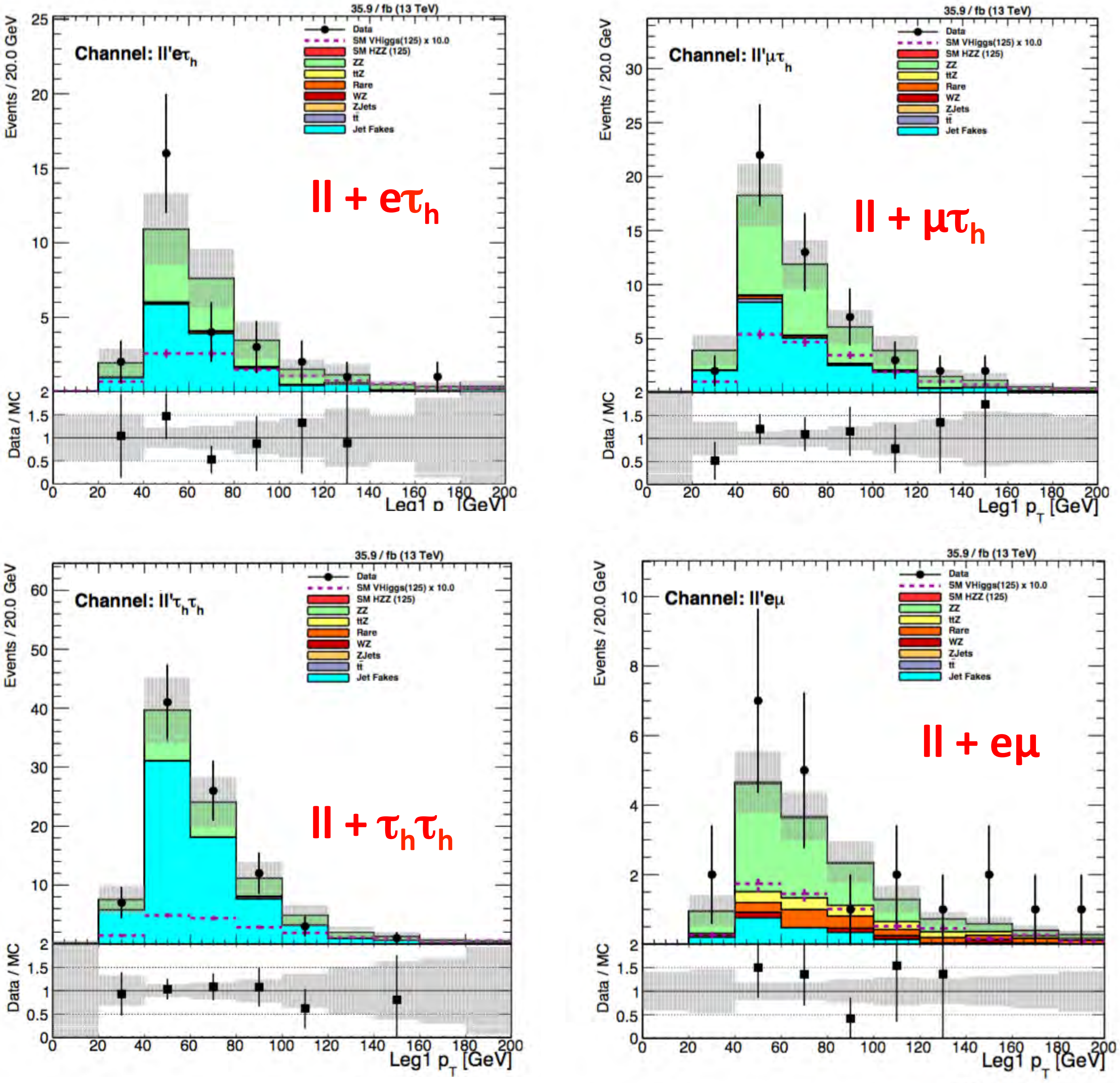
Analysis Overview:

- Targeting the third and fourth leading Higgs production processes by studying three and four lepton final states
- WH – three lepton final states
  - W decaying leptonically: **e || μ**
    - Hadronic: **$\mu\tau_h\tau_h$ & $e\tau_h\tau_h$**
    - Semileptonic: **$e\mu\tau_h$ & $e\mu\tau_h$**
- ZH – four lepton final states
  - Two different Z decay channels: **ee || μμ**
  - Four different Higgs decay channels depending on tau decays:
    - **$e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$, $e\mu$**

ZH – Example Distributions

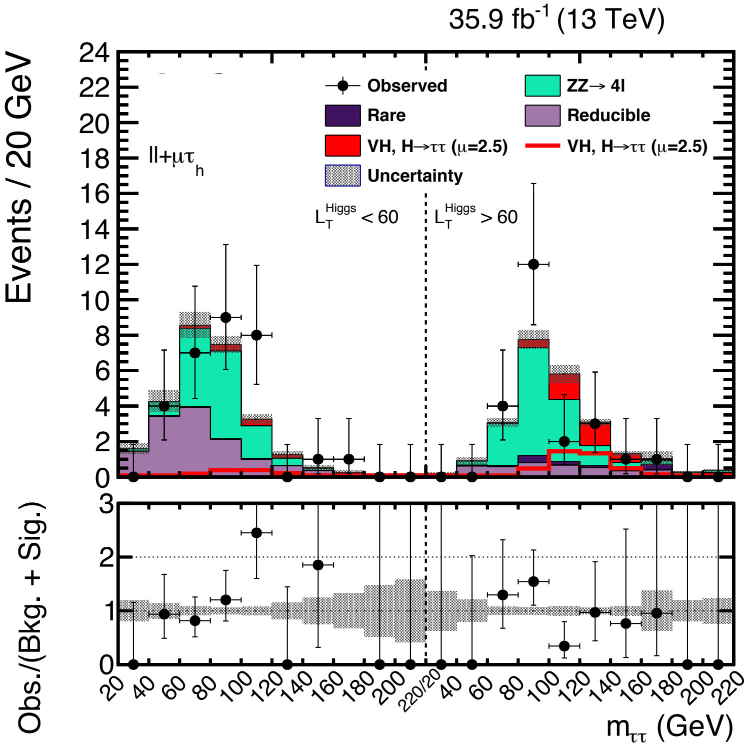
Pre-fit, statistical uncertainty only, showing good yield and shape agreement for **leading Z lepton pT**. Blinded based on sensitivity per bin.

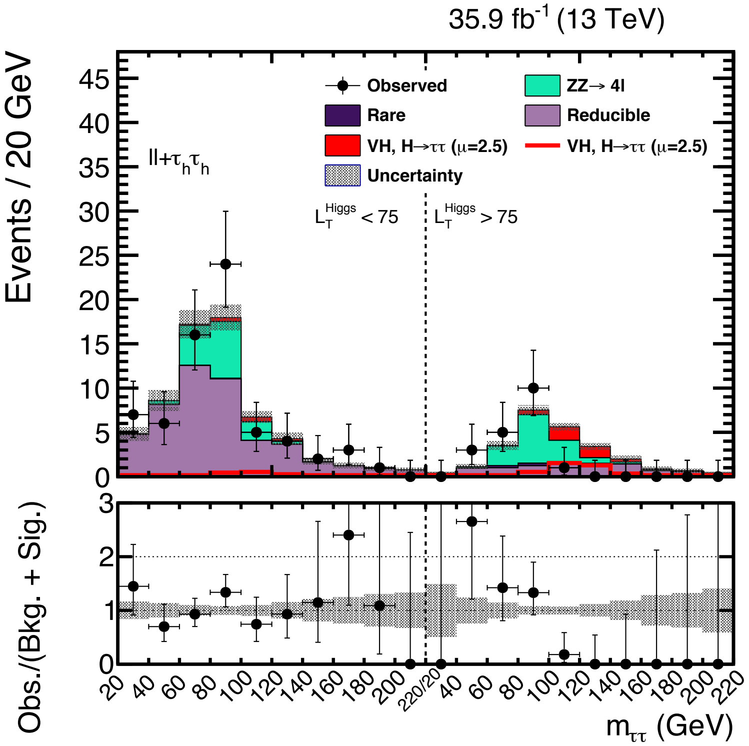
- Good yield and shape agreement for the Jet Fake backgrounds in cyan which are measured from the Higgs same-sign region
- All other backgrounds taken from simulation

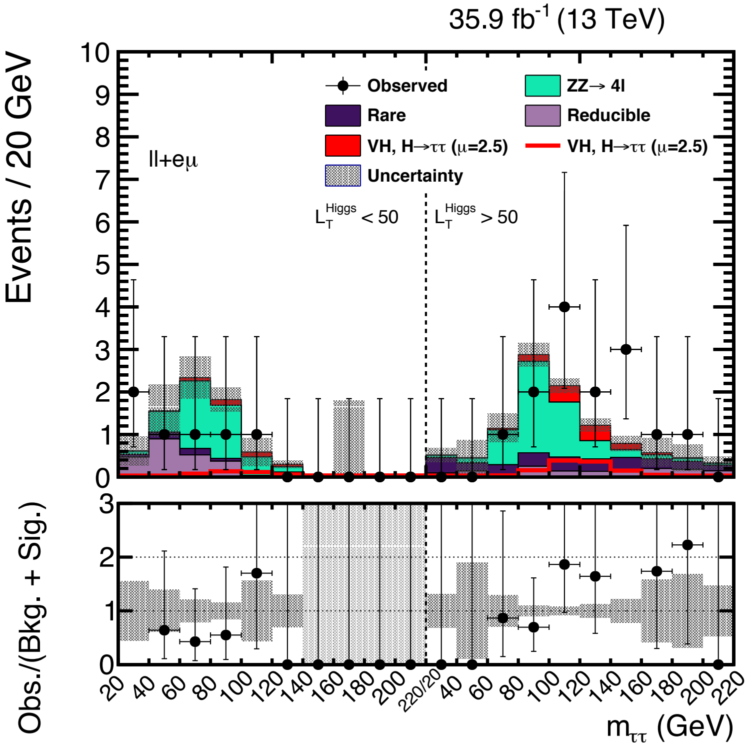


Signal Region Splitting

- In the ZH final states, splitting the signal region into two categories based on the scalar sum of the Higgs boson leptons (LT) improves the analysis sensitivity
- The left half of each distribution is the “Low-LT” region, right is the “High-LT” region

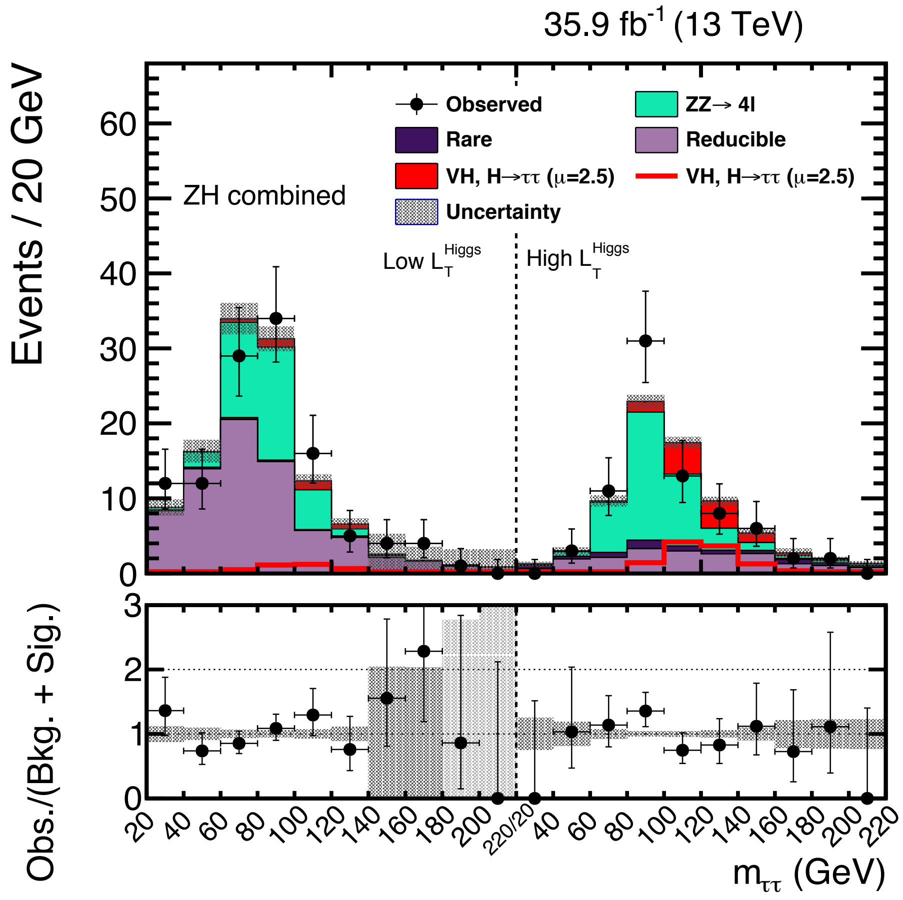


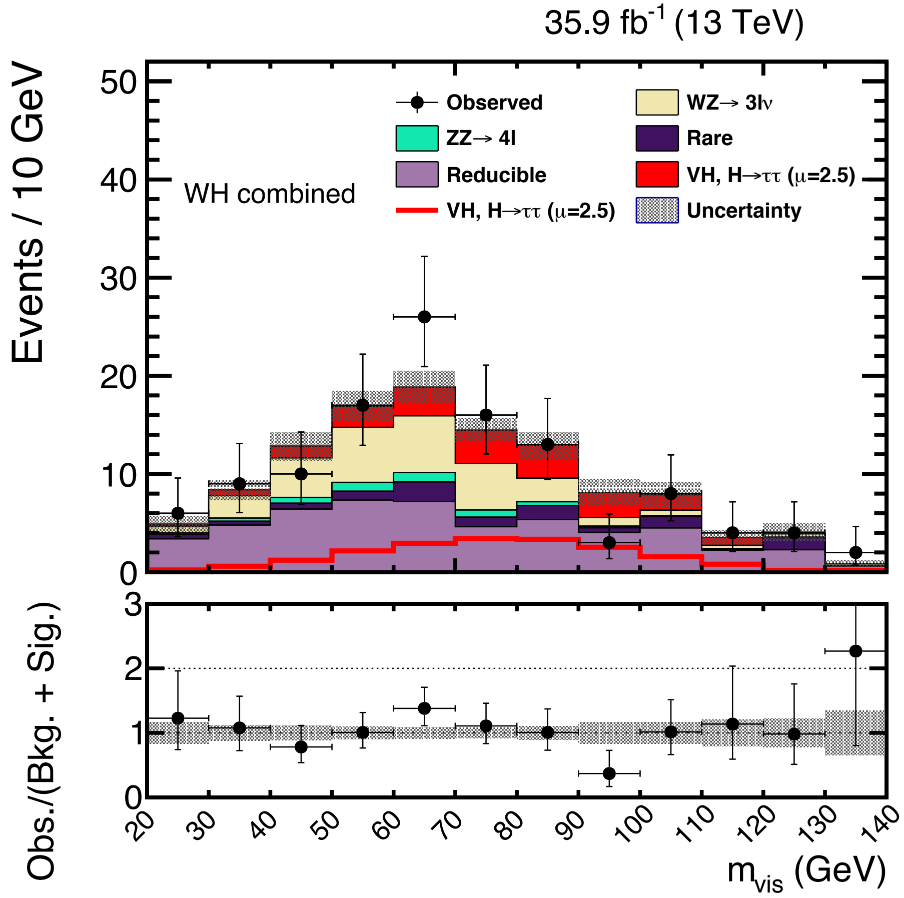




VH Combined Mass Distributions

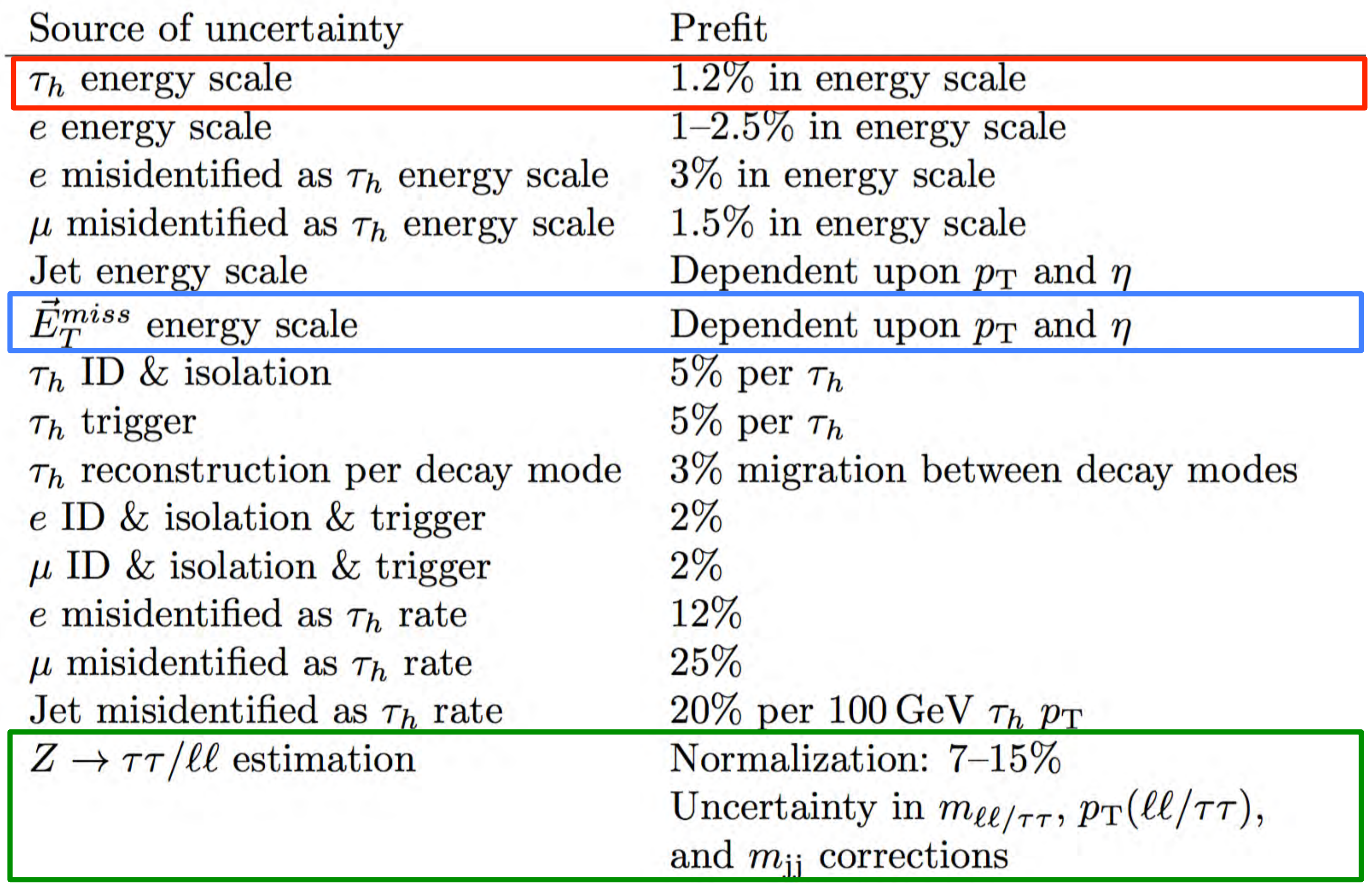
- All eight ZH final states can be combined into a single ZH mass distribution for visualization
- Below, the four WH final states are also be combined for visualization
- The distributions show the best-fit signal strength μ = 2.5 for the VH signals





Systematics

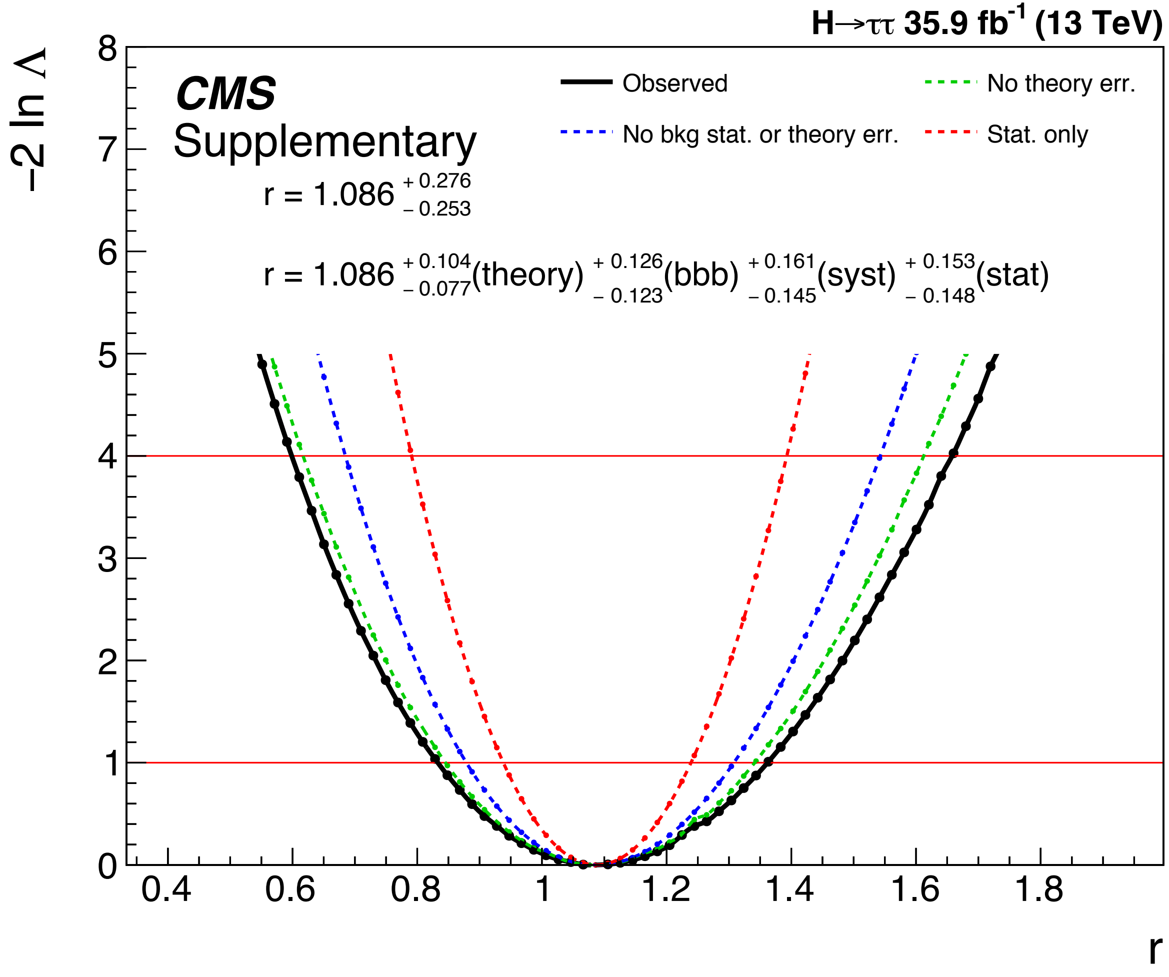
- Many of the uncertainties which have the largest impact on the final results are listed below
- Two of the leading sources of uncertainty are the tau energy scale and MET energy scale, uncertainties which directly affect the shape of the signal extraction variable
- Uncertainties on the modeling of the ZTT background and signal theory uncertainties also have a large impact



Source of uncertainty	Prefit
τ_h energy scale	1.2% in energy scale
e energy scale	1–2.5% in energy scale
e misidentified as τ_h energy scale	3% in energy scale
μ misidentified as τ_h energy scale	1.5% in energy scale
Jet energy scale	Dependent upon p_T and η
$\vec{E}_T^{miss}$ energy scale	Dependent upon p_T and η
τ_h ID & isolation	5% per τ_h
τ_h trigger	5% per τ_h
τ_h reconstruction per decay mode	3% migration between decay modes
e ID & isolation & trigger	2%
μ ID & isolation & trigger	2%
e misidentified as τ_h rate	12%
μ misidentified as τ_h rate	25%
Jet misidentified as τ_h rate	20% per 100 GeV τ_h p_T
$Z \to \tau\tau/\ell\ell$ estimation	Normalization: 7–15% Uncertainty in $m_{\ell\ell/\tau\tau}$, $p_T(\ell\ell/\tau\tau)$, and m_{jj} corrections

Impact of Uncertainties

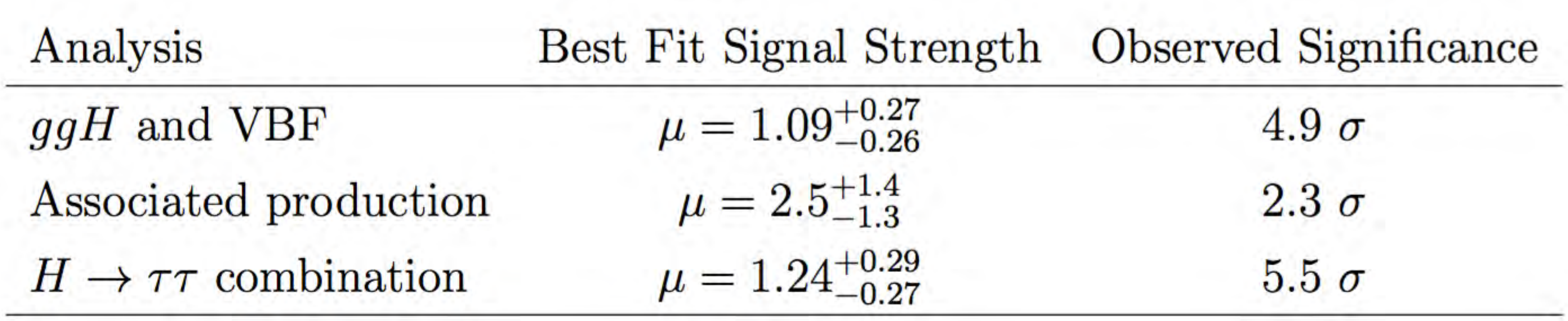
- H → ττ analyses is increasingly systematically limited
- Likelihood scan of the signal strength measurement with the uncertainty split into four groups:
  1. theoretical uncertainties
  2. bin-by-bin statistical uncertainties on the backgrounds
  3. other systematic uncertainties, and
  4. the statistical uncertainty of the data gathered.



Analysis Comparison

- The results of gluon fusion + VBF analysis are combined with the associated production analysis
- The combination, covering the four leading Higgs boson production processes, results in an observed 5.5 standard deviation significance of H → ττ process
- The best-fit signal strength μ = 1.24 is compatible with the standard model expectation within one standard deviation

Signal strength and significance by analysis and combination

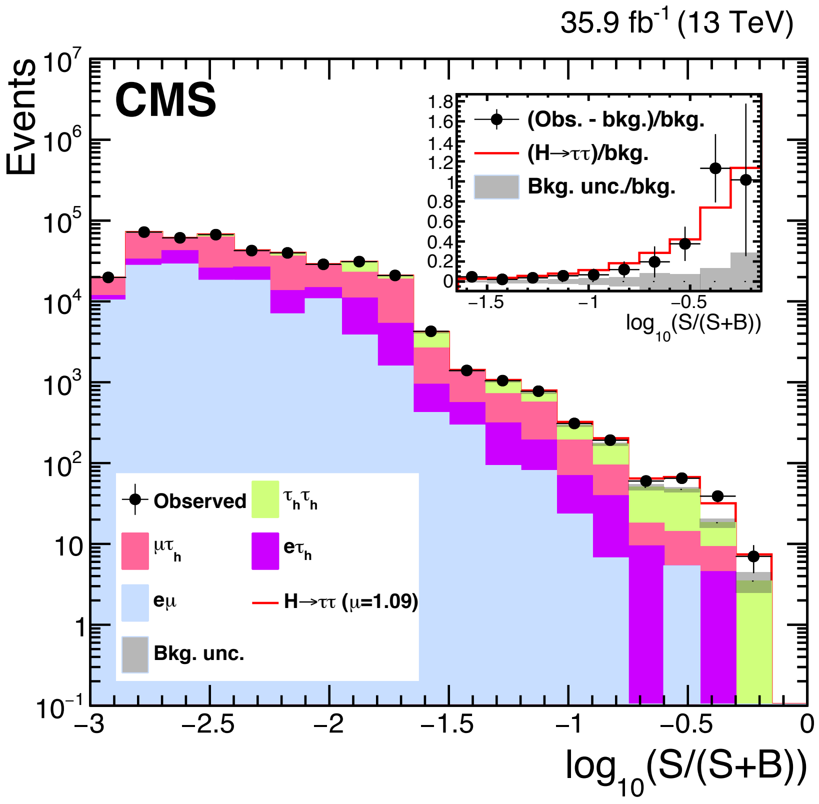


Analysis	Best Fit Signal Strength	Observed Significance
ggH and VBF	$\mu = 1.09^{+0.27}_{-0.26}$	$4.9\ \sigma$
Associated production	$\mu = 2.5^{+1.4}_{-1.3}$	$2.3\ \sigma$
$H \to \tau\tau$ combination	$\mu = 1.24^{+0.29}_{-0.27}$	$5.5\ \sigma$

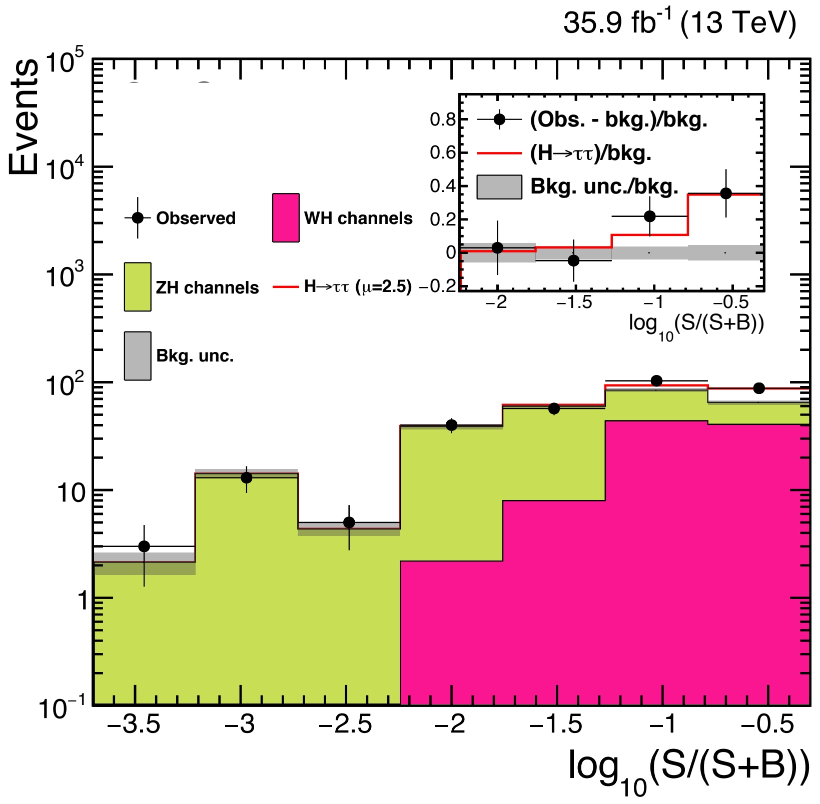
Significance per Bin

- An excess of observed events with respect to the SM background expectation is visible in the most sensitive bins of the analysis
- Below: distributions of the decimal logarithm of the ratio between the best fit signal and the sum of the best fit signal and best fit background expectations in each bin of the mass distributions used to extract the results

ggH & VBF Channels

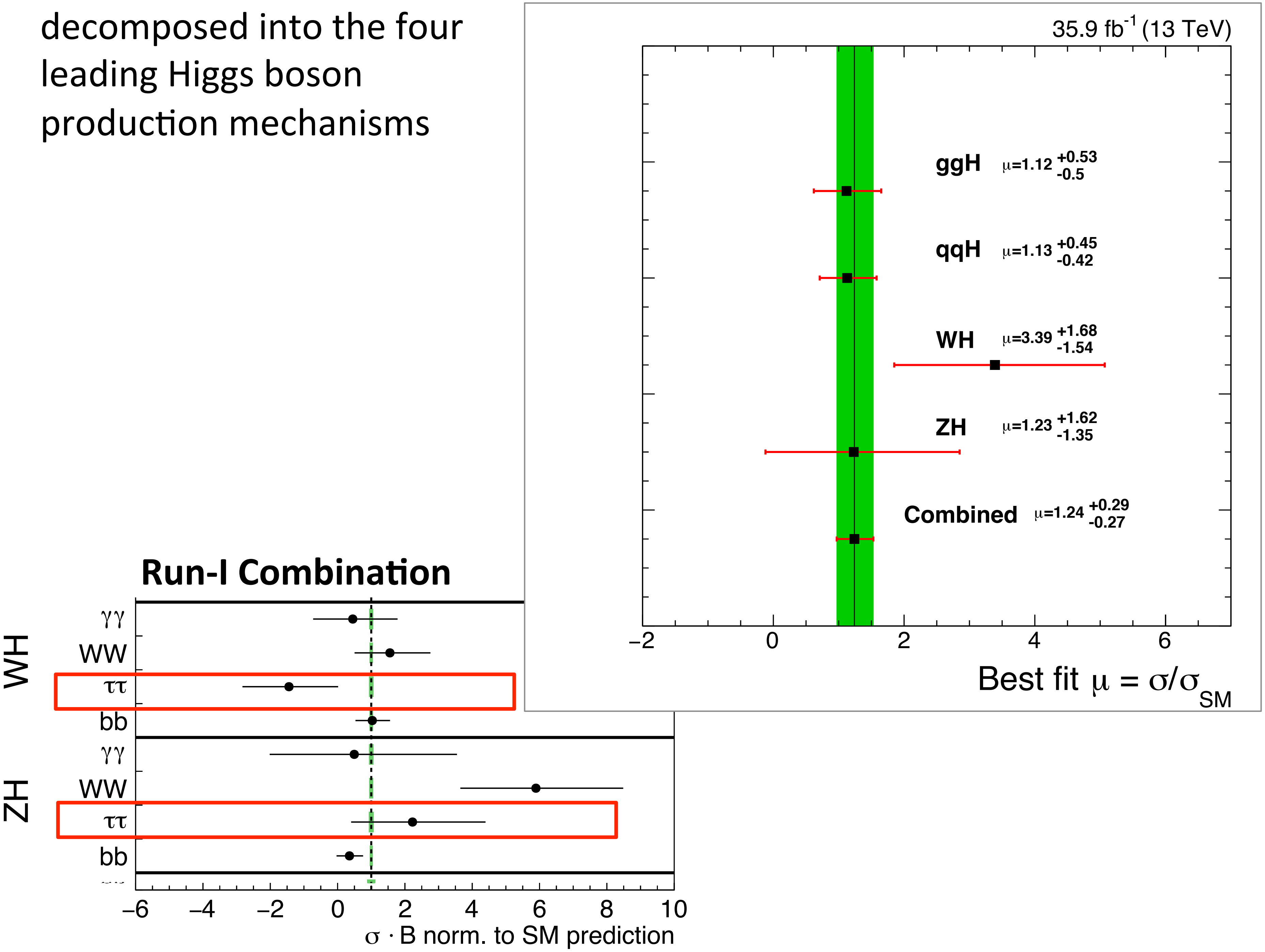


WH & ZH Channels



Signal Strength by Production

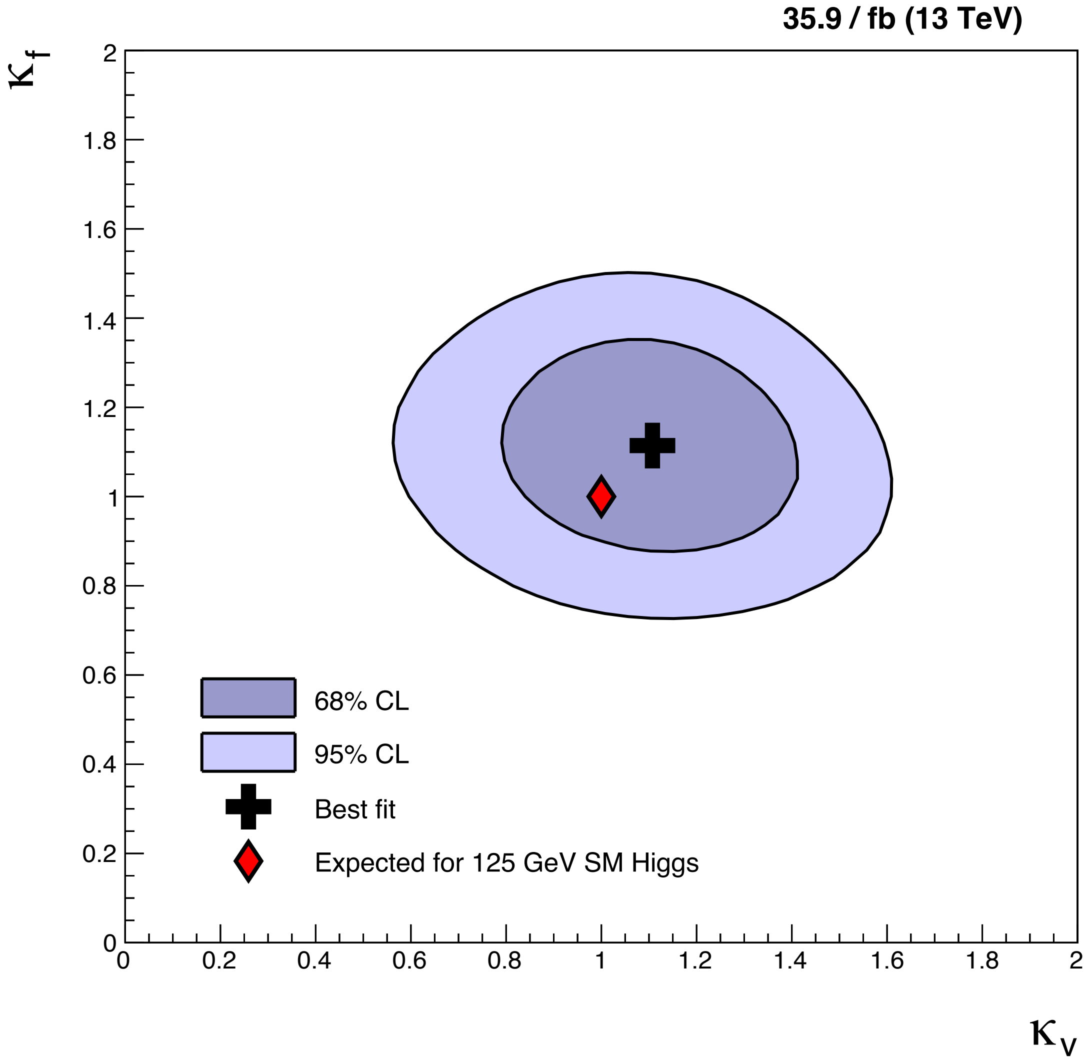
- The signal strength can be decomposed into the four leading Higgs boson production mechanisms



Higgs Boson Couplings

- Couplings of the Higgs boson to fermions and vector bosons can be studied based on Higgs production and decay processes
  - Couplings to vector bosons denoted: κ_V
  - Couplings to fermions denoted: κ_f
- The 2D scan of κ_V and κ_f shows measured Higgs couplings consistent with the standard model

- The measurement agrees with the standard model predictions within one standard deviation



Scan of the negative log-likelihood difference as a function of κ_V and κ_f, for m_H = 125.09 GeV

Conclusion

- This 5.5 standard deviation observation of the H → ττ process and the consistency of the Higgs boson couplings with the standard model provide confirmation of the Higgs boson Yukawa couplings to fermions
- This is evidence that the Higgs field provides mass for the τ lepton in addition to the vector bosons
- The signal strengths for the four targeted Higgs production processes are measured and show consistency with standard model predictions

Future Outlook

- We plan to incorporate new techniques into the analysis once we have the full Run-II dataset, hopefully ~150 fb^{-1}
  - Probe details such as the Higgs boson p_T spectrum with new signal model techniques
- Study potential anomalous couplings of the Higgs boson

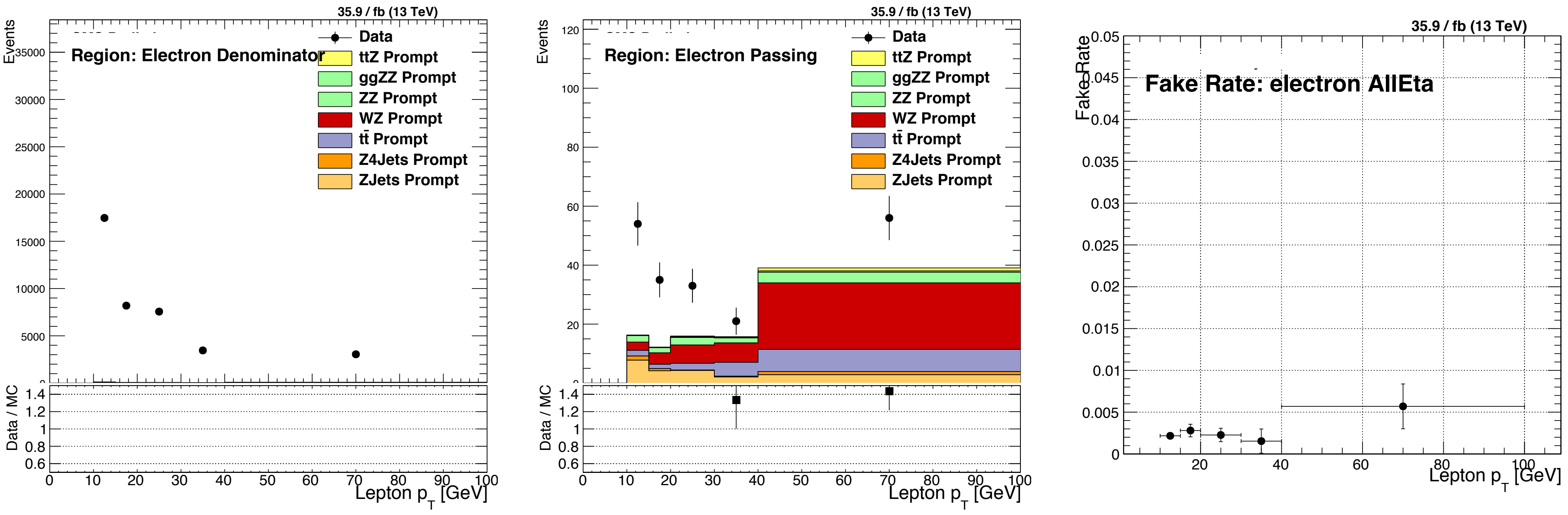
Bonus Material

ZH Fake Rate

- In **Run-I** we used a **100% data-driven Fake Rate** approach to model reducible backgrounds, mainly DY+Jets, WZ+Jets, and ttbar
- In **Run-2** the optimized electron and muon WPs chosen for the Higgs candidate legs in LLET, LLMT, and LLEM final states result in **a large prompt lepton contribution in measurement region** passing signal cuts
- Because we want the fake rate method to estimate the backgrounds from Jet → e/mu/tau, we much take care to treat the prompt background appropriately
- To achieve this, use data-driven method aided by prompt MC subtraction
  - In the denominator and passing regions used to construct the fake rates estimate **fake contribution = data – prompt matched MC**
- Yield is estimated from the measured fake rates by applying a weight to data events which fail the signal region lepton ID and/or isolation requirements
- The shape of the jet fake background is taken from the Higgs same-sign region and uses relaxed isolation and ID criteria

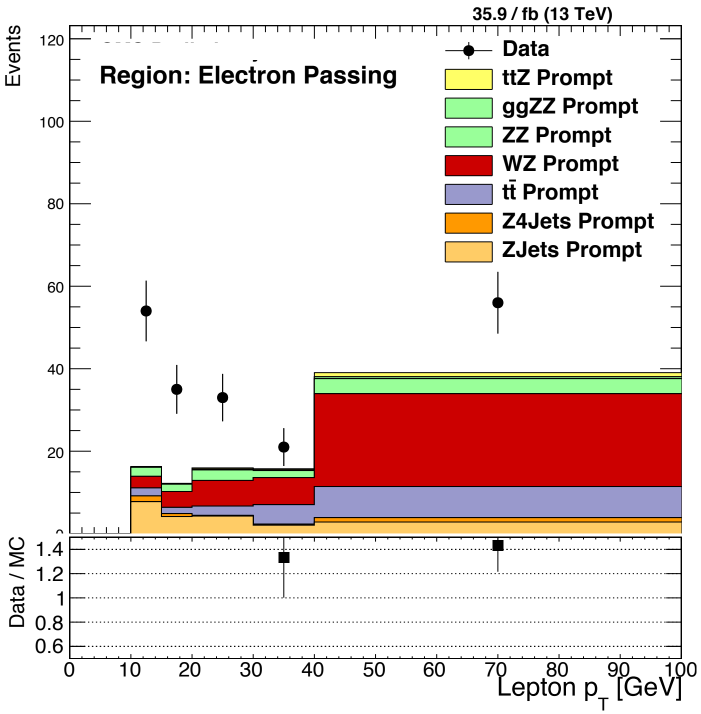
Measuring Fake Rates

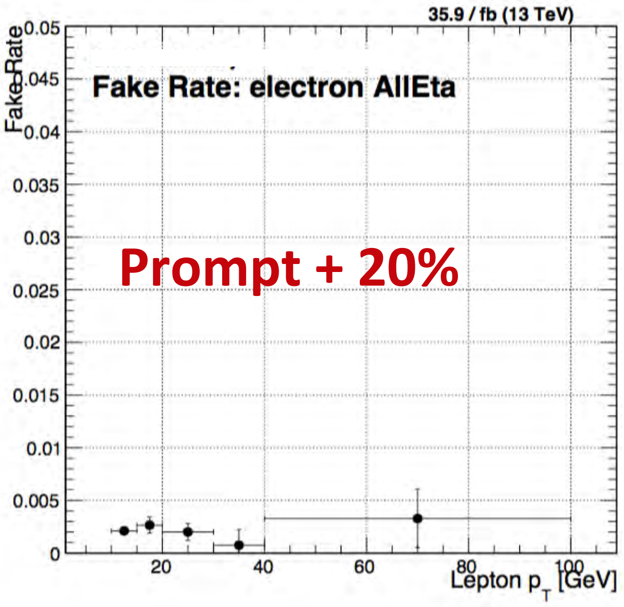
- The fake rates are measured in the Higgs same-sign region to avoid prompt lepton contributions as much as possible
- Fake rates are parameterized by lepton pT
- In data the prompt contribution is estimated from MC and subtracted
- Uncertainty in fake rate related to prompt background subtraction
- **Below:** Showing electron fake rates estimation, (left) denominator selection, (middle) passing region, (right) resulting fake rates



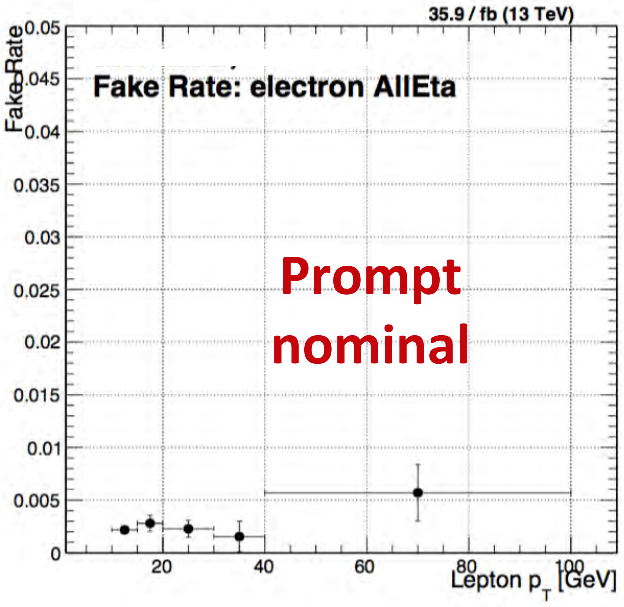
Fake Rate Uncertainties

- The prompt contributions in jet→e/μ are considerable in the passing region at higher lepton pT
- We add a shape based uncertainty which fluctuates the prompt contributions +/- 20% (10%) in ZH (WH) FRs
- **Below:** Showing electron fake rates with different scaling for prompt backgrounds

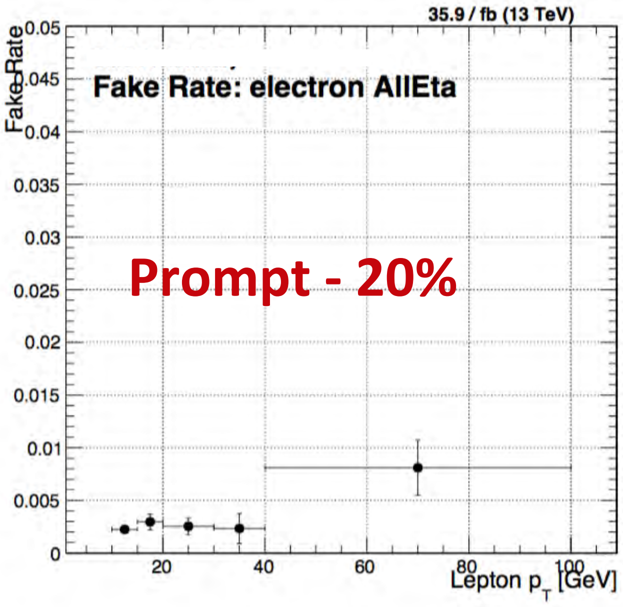




Prompt + 20%



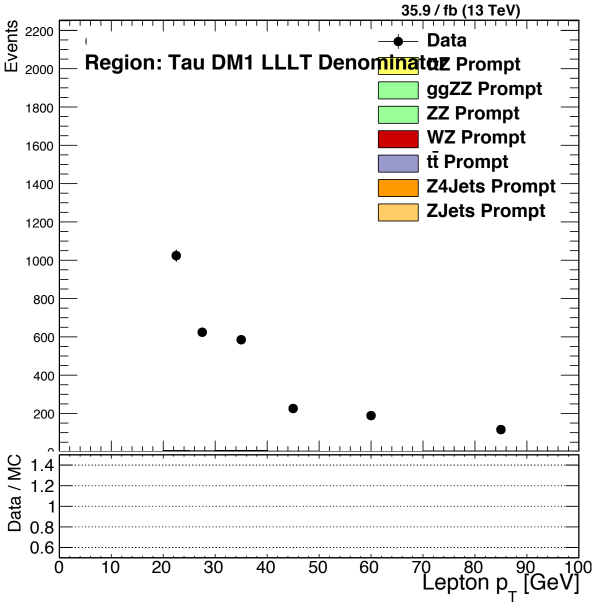
Prompt nominal

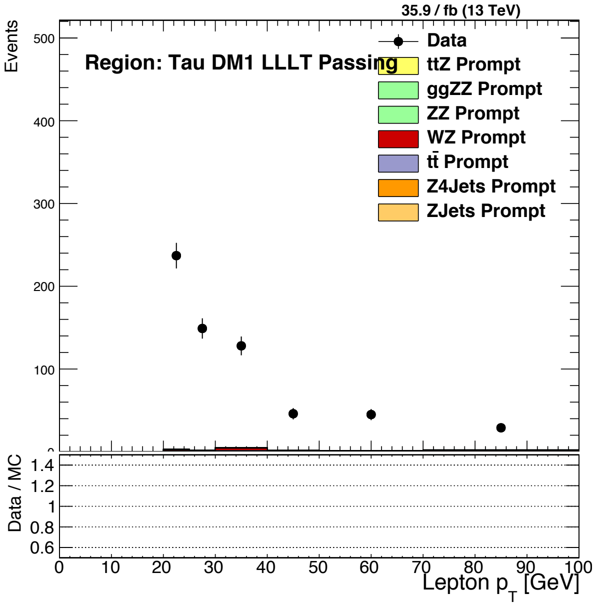


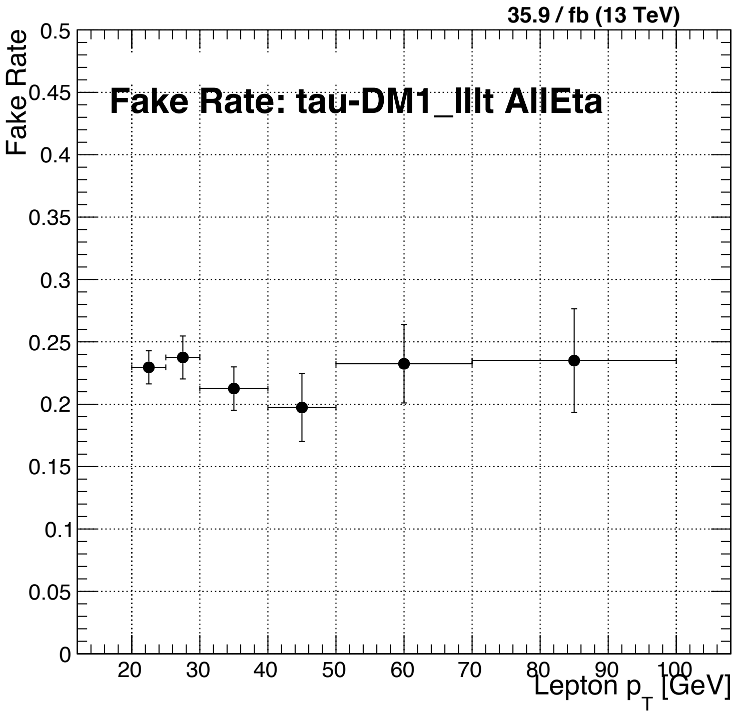
Prompt - 20%

Tau Fake Rates

- Tau fake rates are split by tau decay mode in denominator selection
- A very very loose tau ID is applied in denominator selection as well
- **Below:** 1prong+1piZero fake rate estimation in LLLT events, (left) denominator selection, (middle) passing region, (right) resulting fake rates
- The contribution from prompt leptons is negligible in the tau fake rate measurements







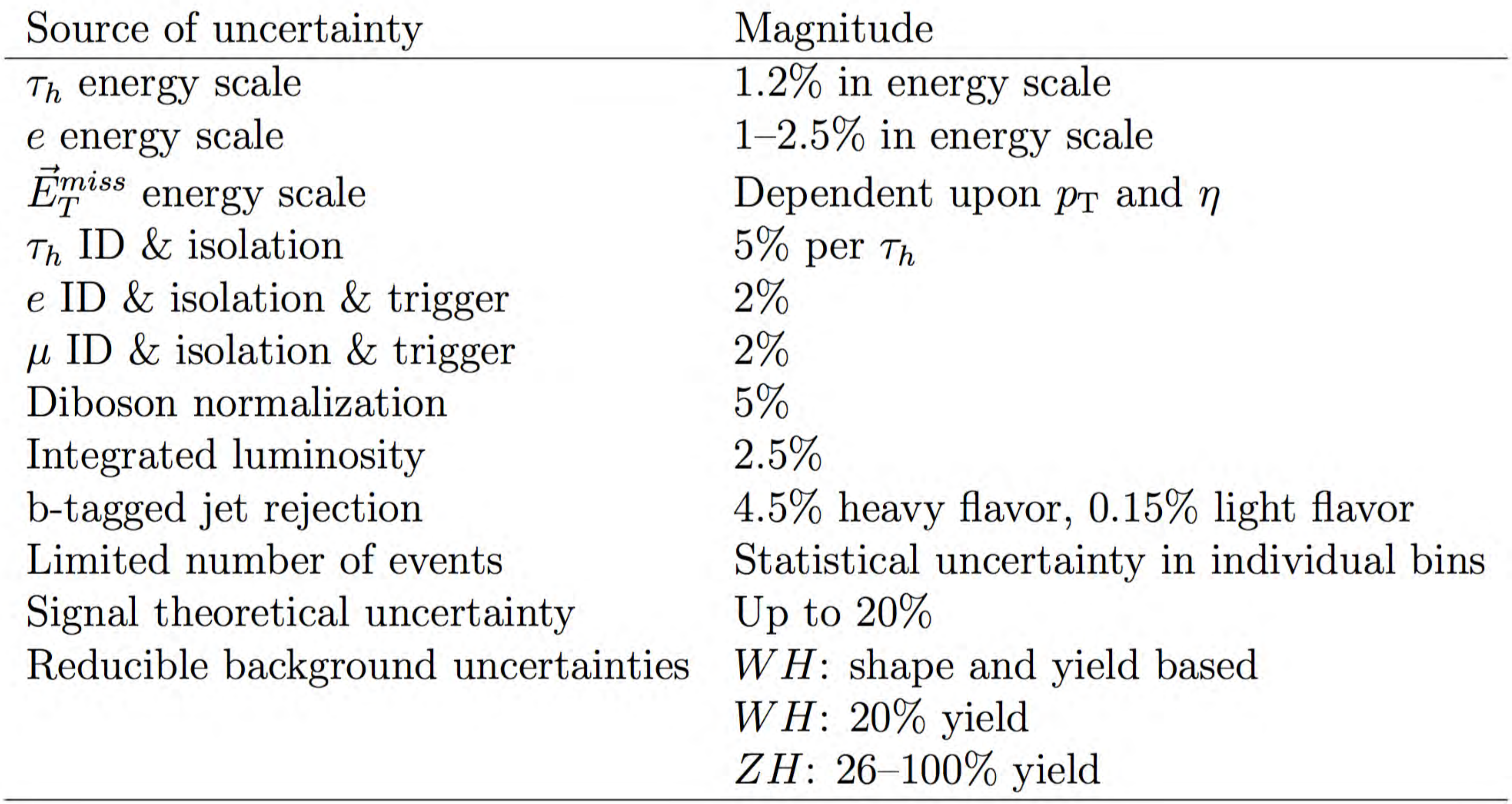
Yield & Shape of Jet Fake Backgrounds

- The yield of the Jet Fake background in the signal region is estimated from the measured fake rates by applying a weight to data events which fail the signal region lepton ID and/or isolation requirements

$$\text{Fake Rate Weight 1} = \frac{F\left(\text{lepton } p_{\mathrm{T}}\ lep_1\right)}{\left(1 - F\left(\text{lepton } p_{\mathrm{T}}\ lep_1\right)\right)}$$

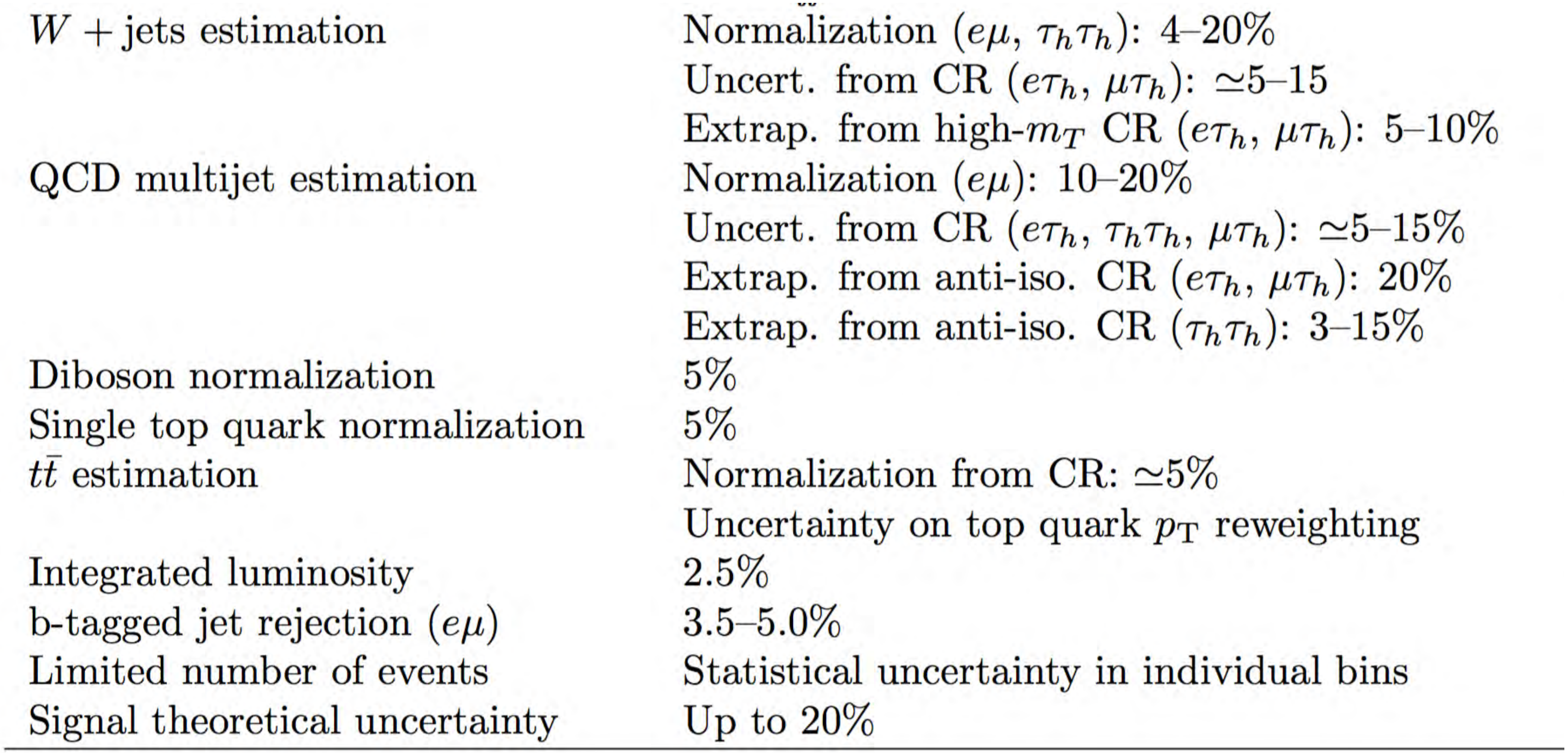
- In the ZH final states, the shape of the jet fake background is taken from the Higgs same-sign region and uses relaxed isolation and ID to achieve a high statistics shape which is compatible with jet fakes from MC

VH Systematics



Source of uncertainty	Magnitude
τ_h energy scale	1.2% in energy scale
e energy scale	1–2.5% in energy scale
$\vec{E}_T^{miss}$ energy scale	Dependent upon p_{T} and η
τ_h ID & isolation	5% per τ_h
e ID & isolation & trigger	2%
μ ID & isolation & trigger	2%
Diboson normalization	5%
Integrated luminosity	2.5%
b-tagged jet rejection	4.5% heavy flavor, 0.15% light flavor
Limited number of events	Statistical uncertainty in individual bins
Signal theoretical uncertainty	Up to 20%
Reducible background uncertainties	WH: shape and yield based
	WH: 20% yield
	ZH: 26–100% yield

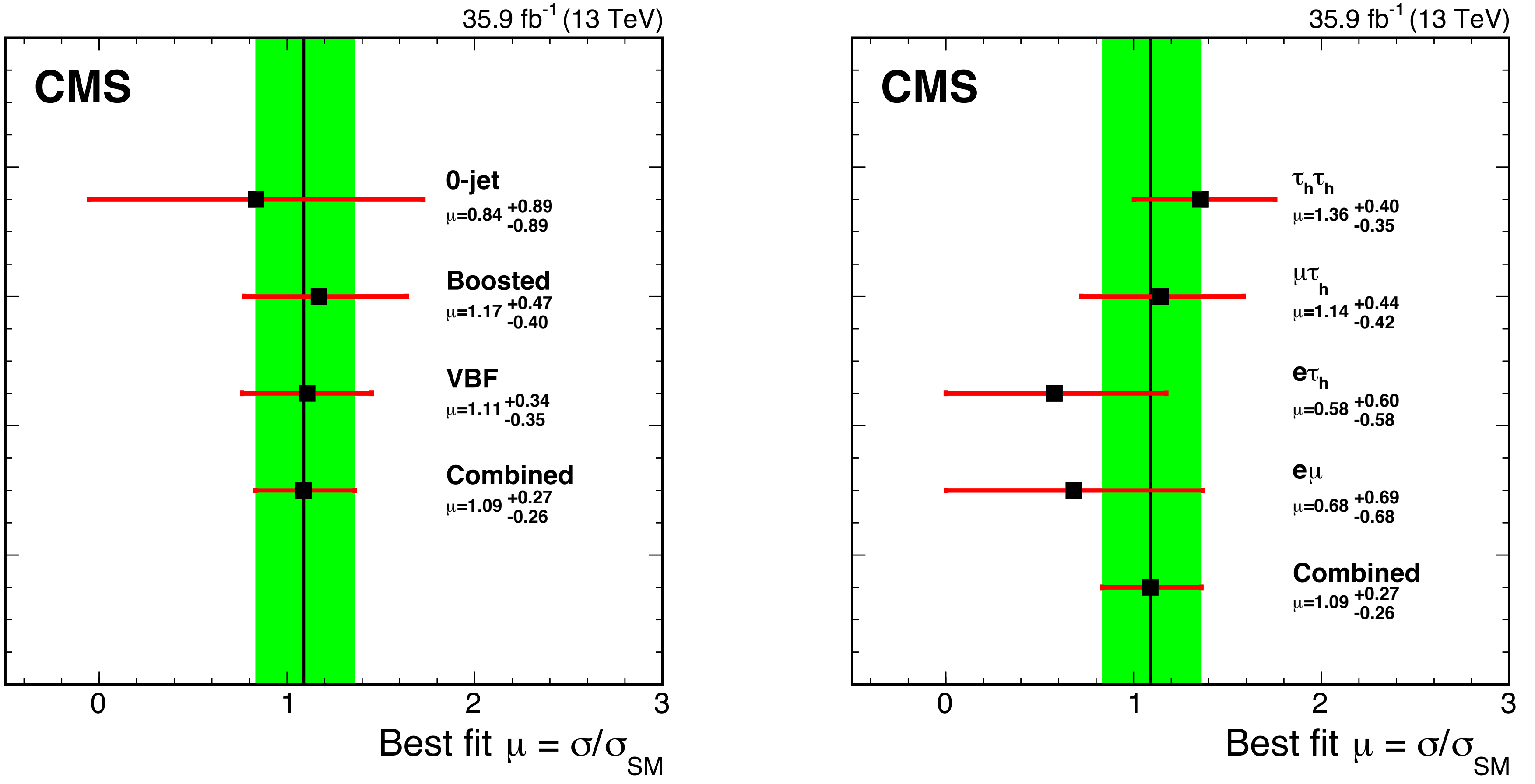
HTT Systematics - 2



$W+$jets estimation	Normalization ($e\mu$, $\tau_h\tau_h$): 4–20%
	Uncert. from CR ($e\tau_h$, $\mu\tau_h$): $\simeq$5–15
	Extrap. from high-m_T CR ($e\tau_h$, $\mu\tau_h$): 5–10%
QCD multijet estimation	Normalization ($e\mu$): 10–20%
	Uncert. from CR ($e\tau_h$, $\tau_h\tau_h$, $\mu\tau_h$): $\simeq$5–15%
	Extrap. from anti-iso. CR ($e\tau_h$, $\mu\tau_h$): 20%
	Extrap. from anti-iso. CR ($\tau_h\tau_h$): 3–15%
Diboson normalization	5%
Single top quark normalization	5%
$t\bar{t}$ estimation	Normalization from CR: $\simeq$5%
	Uncertainty on top quark p_{T} reweighting
Integrated luminosity	2.5%
b-tagged jet rejection ($e\mu$)	3.5–5.0%
Limited number of events	Statistical uncertainty in individual bins
Signal theoretical uncertainty	Up to 20%

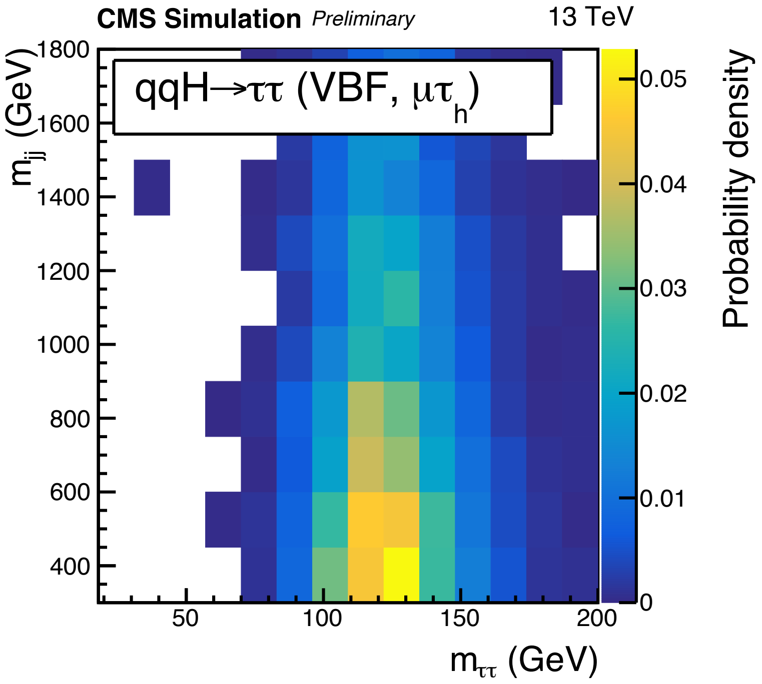
Di-Tau Signal Strengths

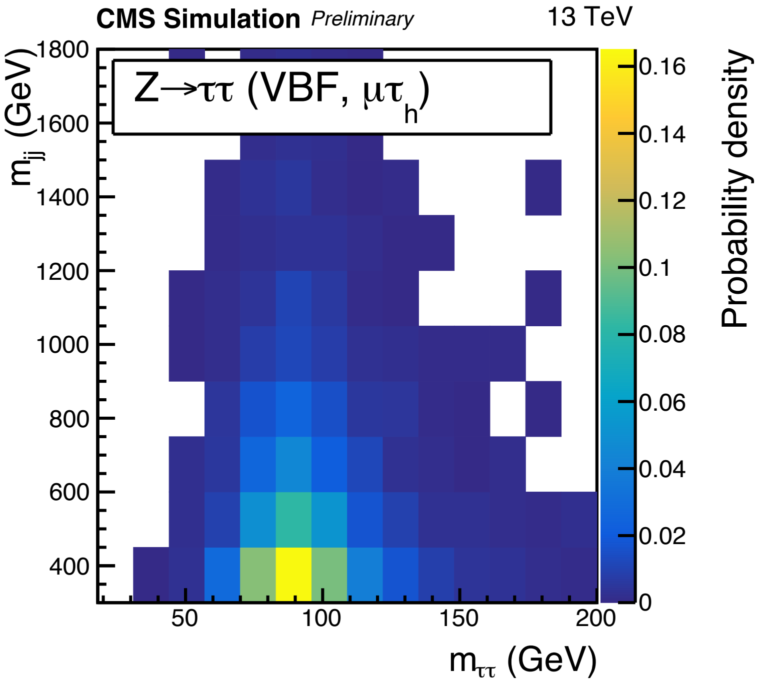
- Best fit signal strength μ per category (left) and channel (right), for m_H = 125.09 GeV



Separation of Signal and Background

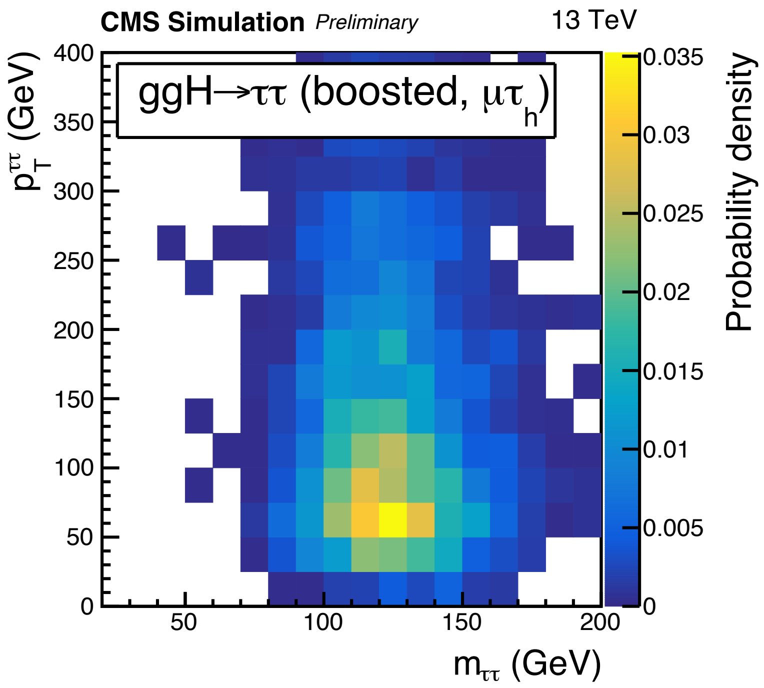
- Additional separation of signal and background can be gained by specifically targeting different higgs production kinematics for VBF & gluon fusion
- In the VBF targeted category, we consider 2D distributions of di-tau mass ($m_{\tau\tau}$) and m_{jj}

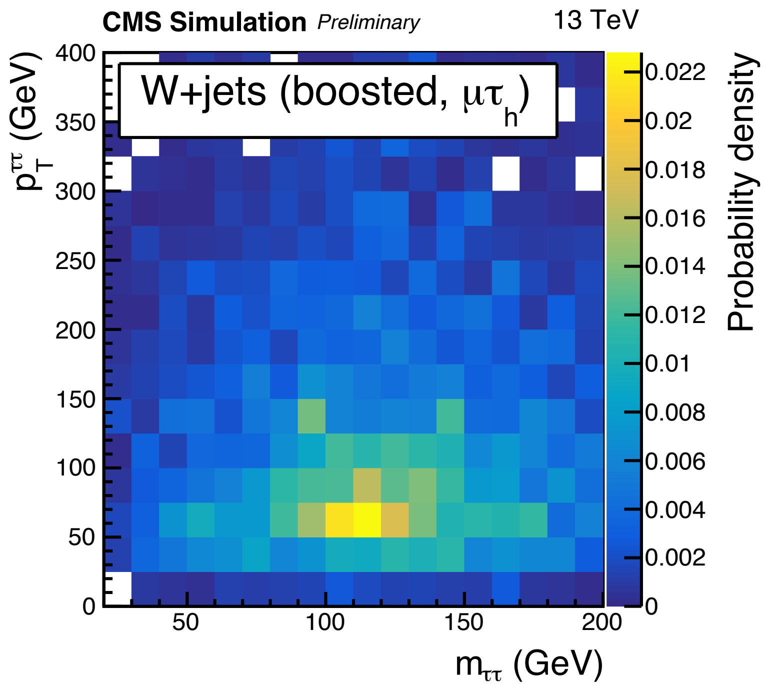




Separation of Signal and Background

- Additional separation of signal and background can be gained by specifically targeting different higgs production kinematics for VBF & gluon fusion
- In the Boosted category, we consider 2D distributions of di-tau mass ($m_{\tau\tau}$) and Higgs pT ($p_T^{\tau\tau}$)





Yukawa Interactions

- A Yukawa coupling is an interaction between a scalar field and a Dirac field similar to

$$V \approx h_f \bar{\psi}_f \phi \psi_f$$

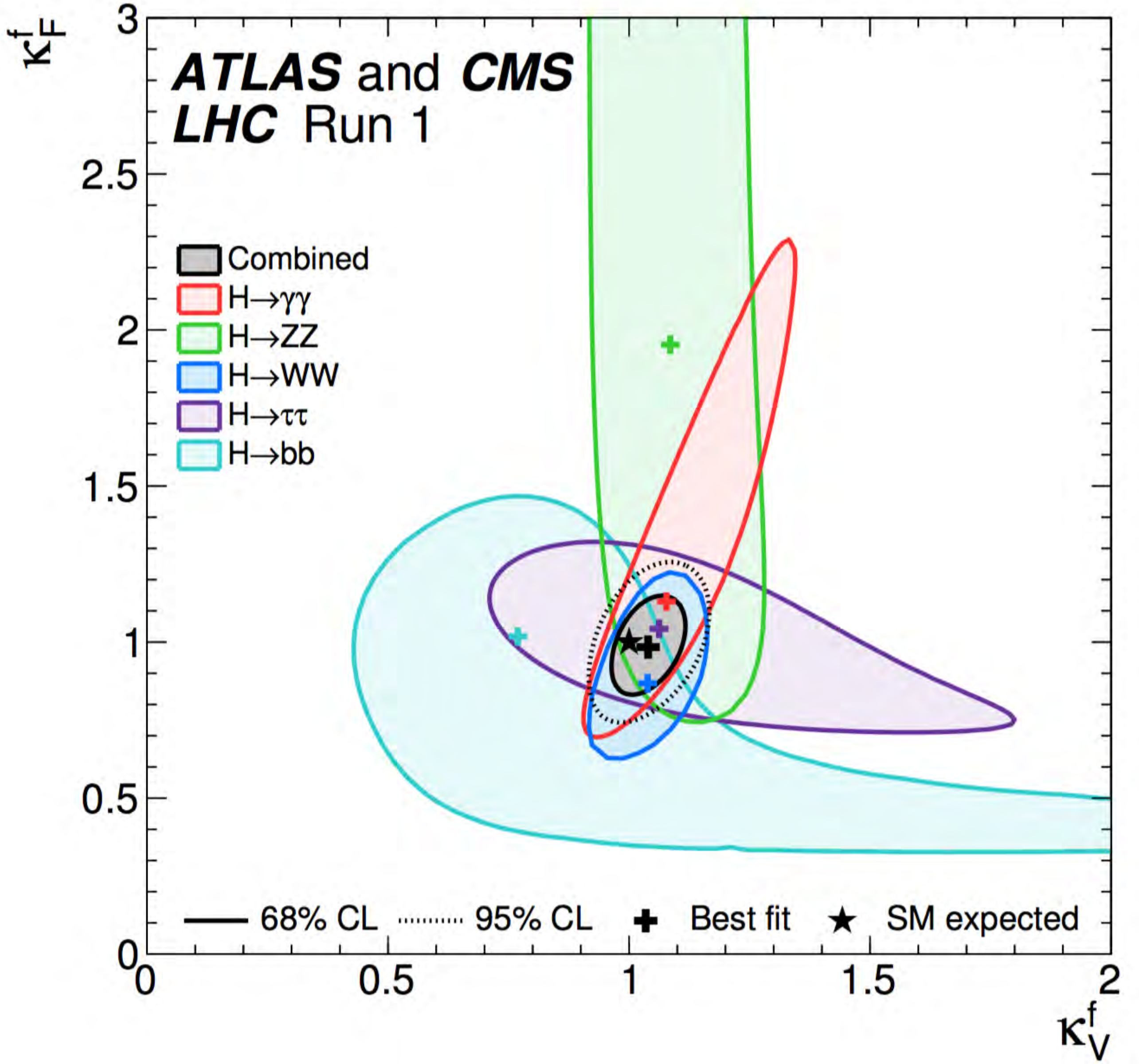
  - Dirac fields, ψ, describe fermions
  - The scalar field, ϕ, is taken to be that of the Higgs boson
- Yukawa interaction linking together the fermions and the Higgs boson, results in massive fermions where their mass can be written as,

$$m_f = \frac{h_f v}{\sqrt{2}}$$

- m_f covers the masses for the nine charged fermions
- h_f is the Yukawa coupling of that fermion to the Higgs boson
- v is the vacuum expectation value from the previous slide, v = 246 GeV

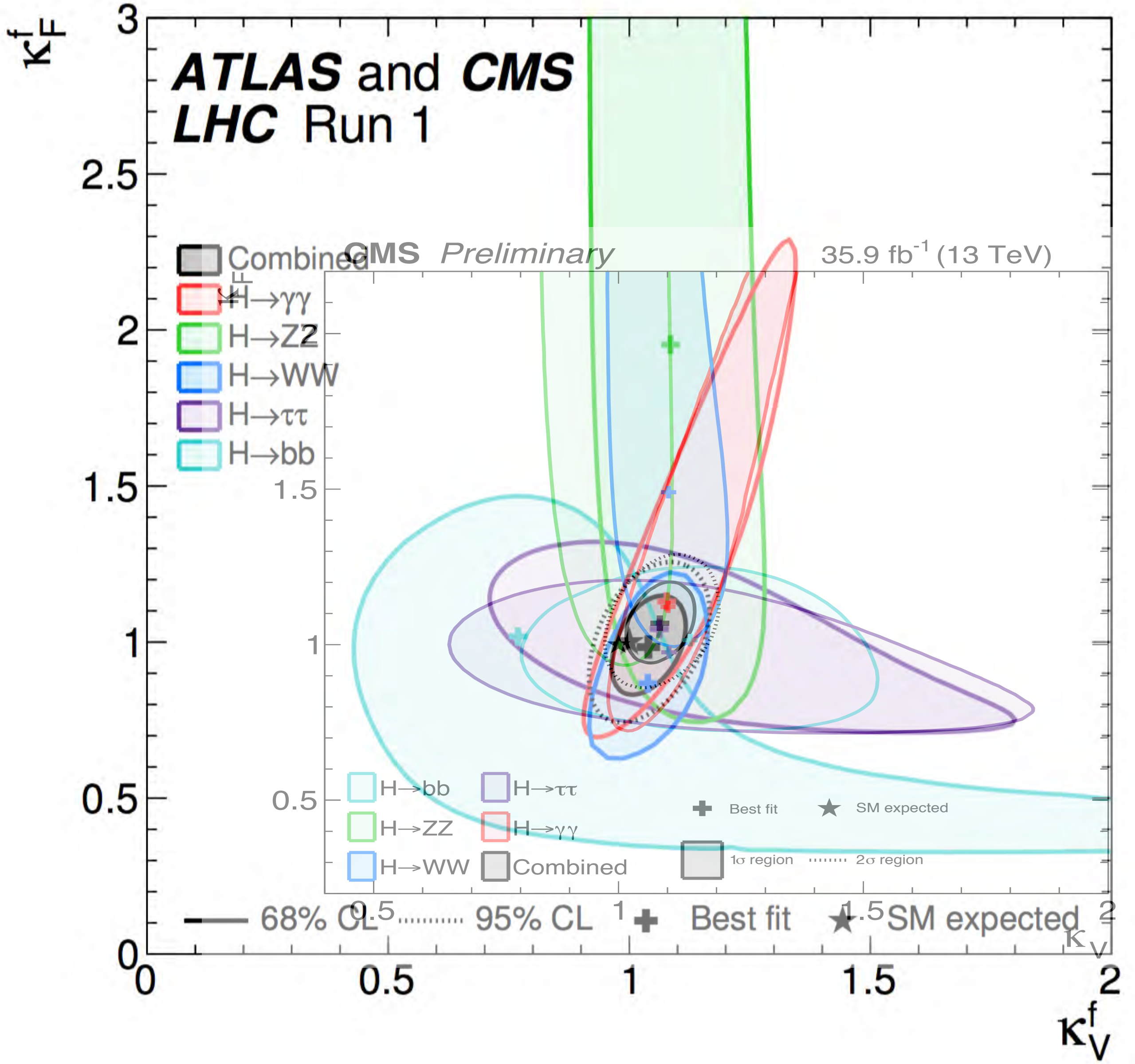
H → ττ Couplings Progression - I

- Run-I CMS + ATLAS Combination



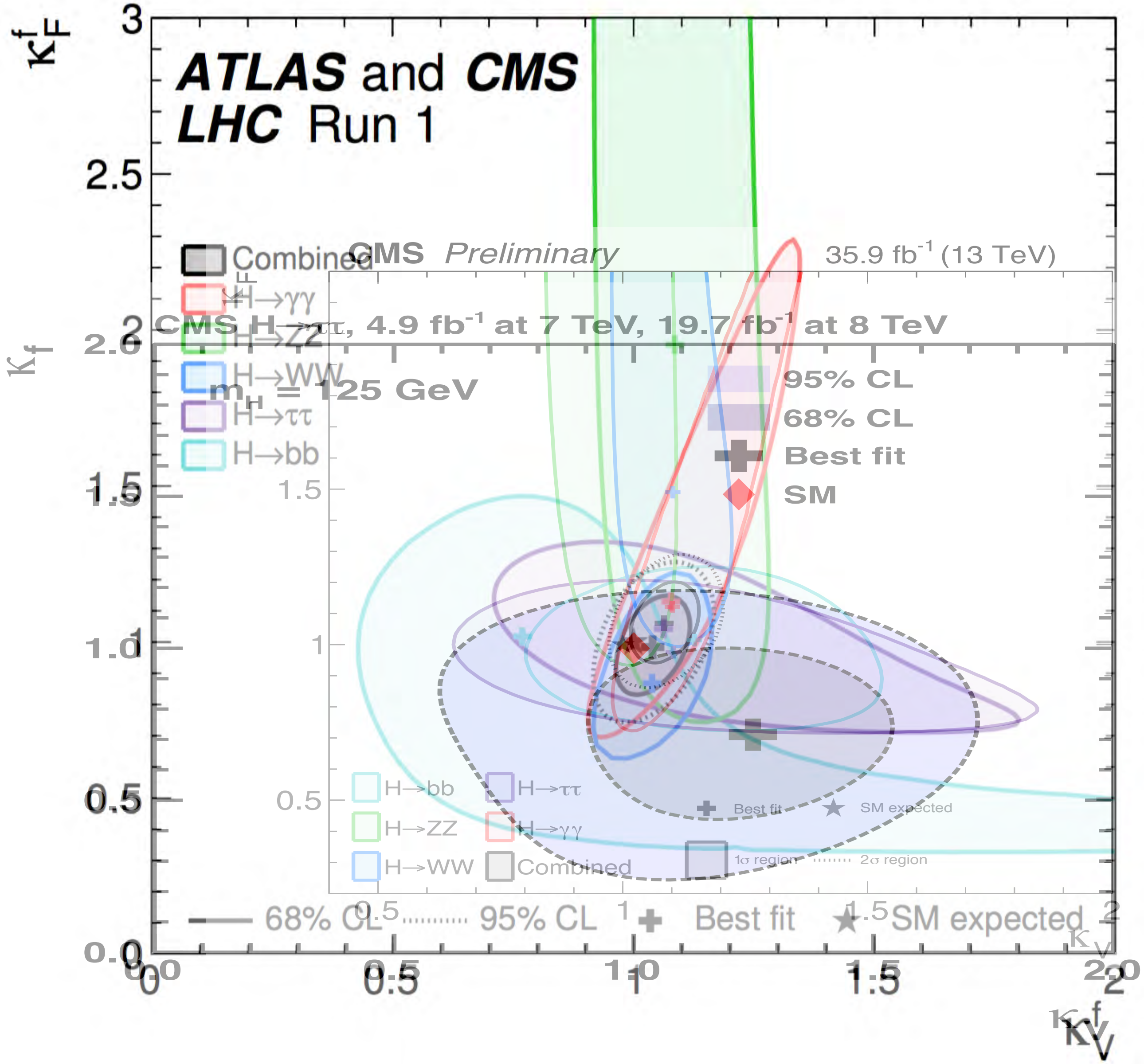
H → ττ Couplings Progression - II

- + CMS Run-II 2016 combination without VH



H → ττ Couplings Progression - III

- + CMS Run-I distribution



H → ττ Couplings Progression - IV

- + CMS 2016 Run-II distribution