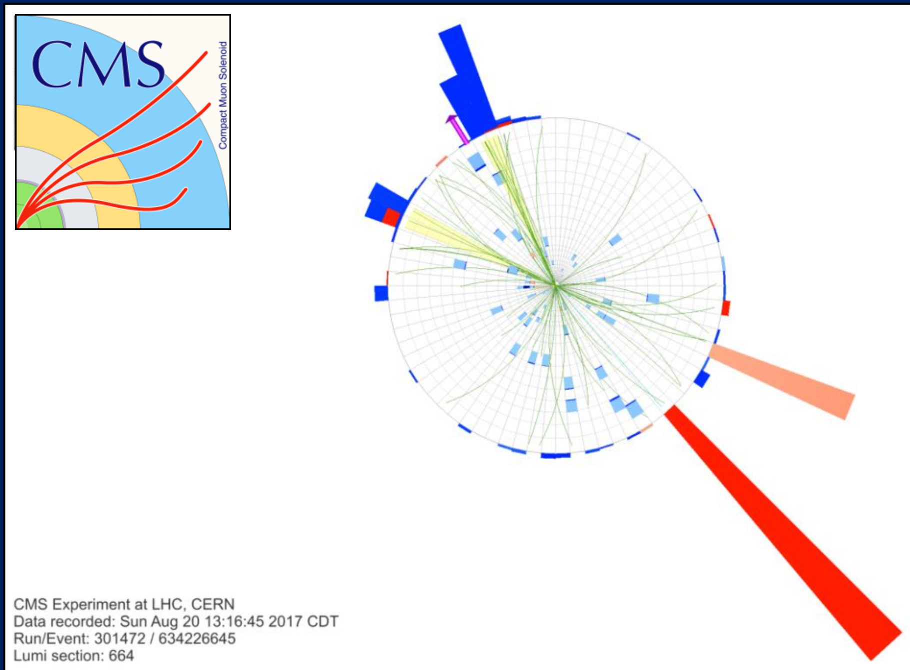


Observation of Higgs boson decay to bottom quarks with CMS data

HIG-18-016 - arXiv:1808.08242,
provisionally accepted for publication in PRL

Luca Perrozzi (ETH Zurich)
on behalf of the CMS Collaboration

LPCC seminar - CERN
Aug. 28st 2018

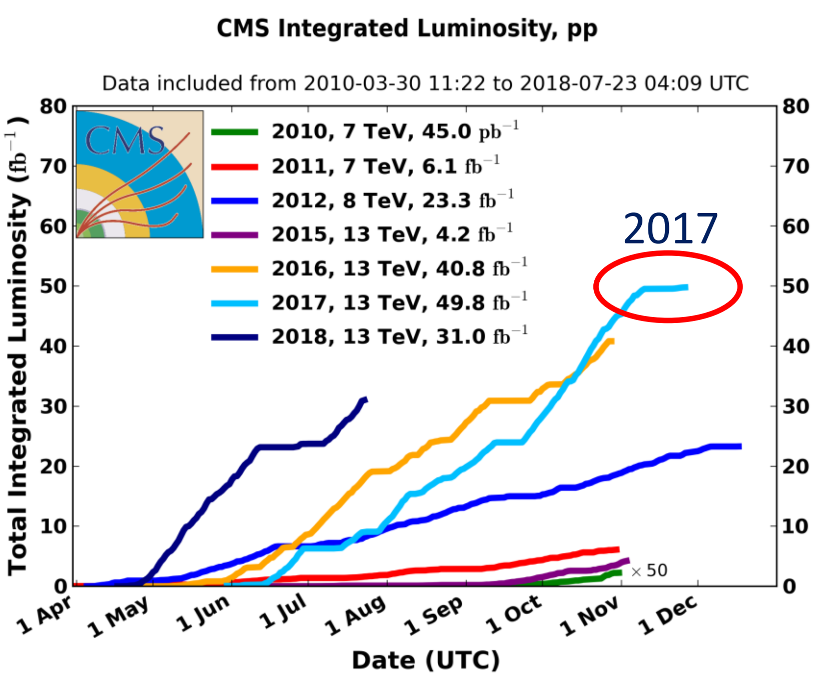
ETH zürich

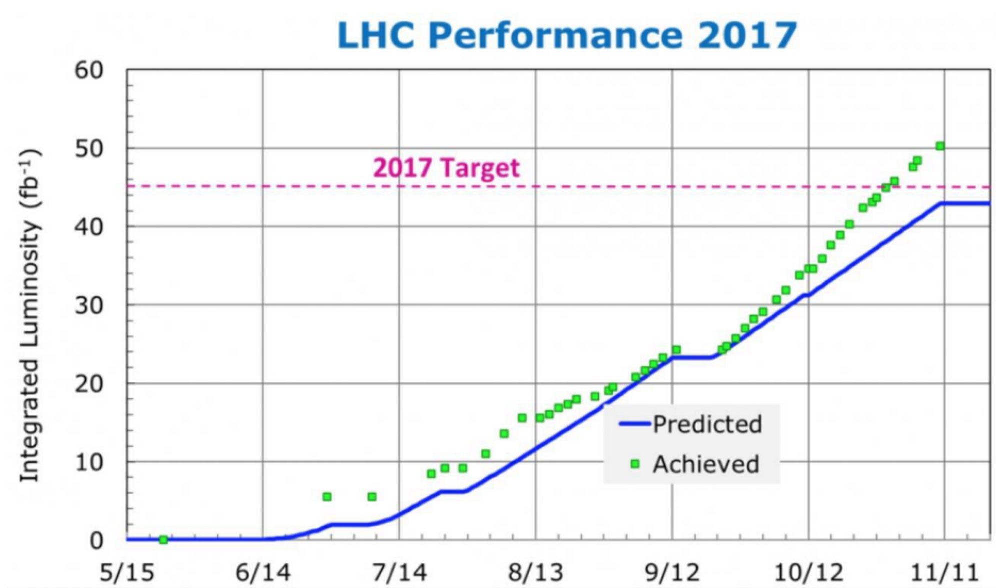
Outline

- **LHC and CMS**
- **H→bb: motivation and challenges**
- **VH(bb) analysis with 2017 data**
- **Combination with previous results**
- **Prospects**

The Large Hadron Collider

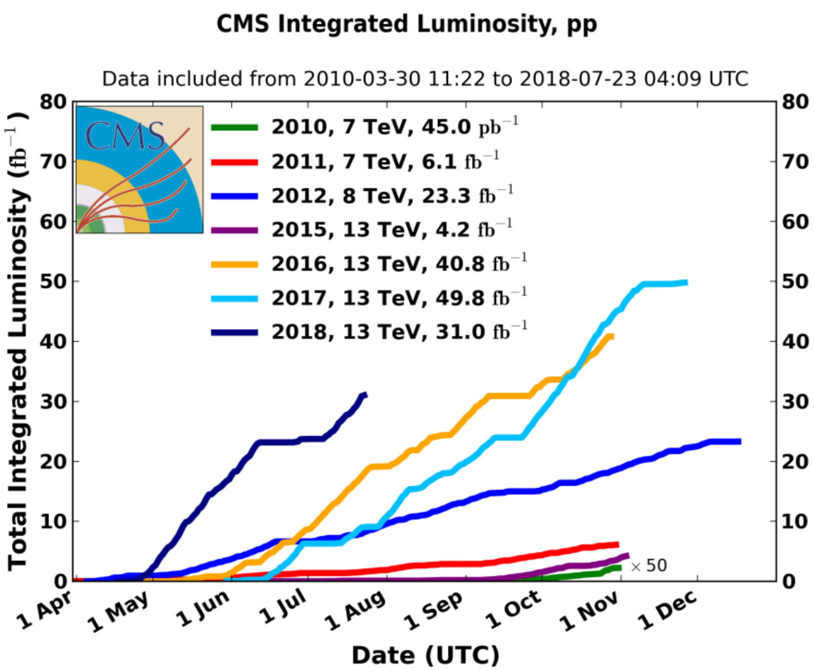
- After successful Run 1, **LHC has produced >3 years of 13 TeV data with stunning performance**
- Expected integrated luminosity **>150 fb^{-1} by the end of 2018**
- **DESIGN peak luminosity exceeded by a factor of 2**
  - Average pileup ~38 in 2017 and 2018
- Incredible machine availability, **>50% of time in stable operation**





The Large Hadron Collider

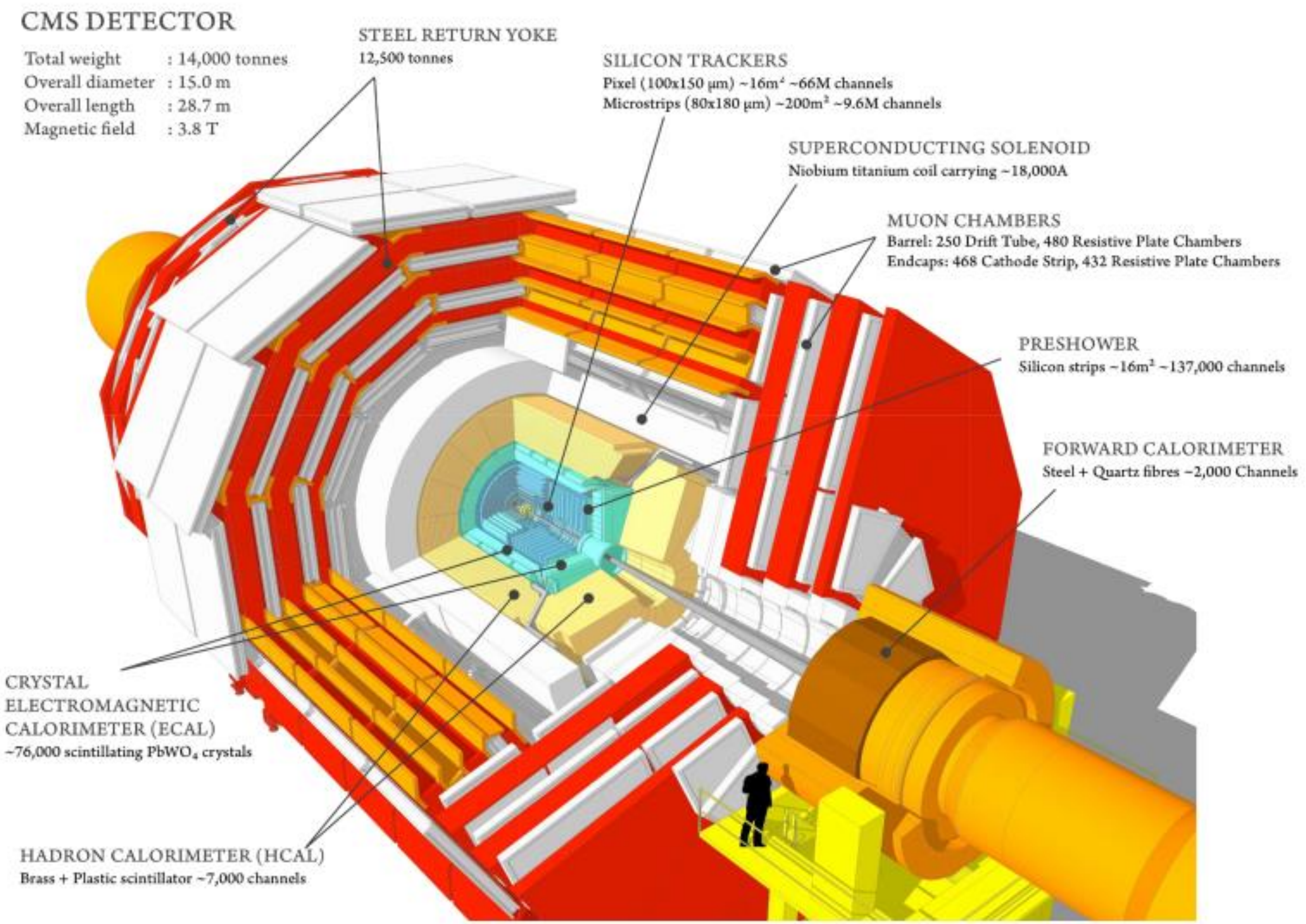
- After successful Run 1, **LHC has produced >3 years of 13 TeV data with stunning performance**
- Expected integrated luminosity **>150 fb^{-1} by the end of 2018**
- **DESIGN peak luminosity exceeded by a factor of 2**
  - Average pileup ~38 in 2017 and 2018
- Incredible machine availability, **>50% of time in stable operation**



A big
THANK YOU!
to CERN and the LHC team

Compact Muon Solenoid

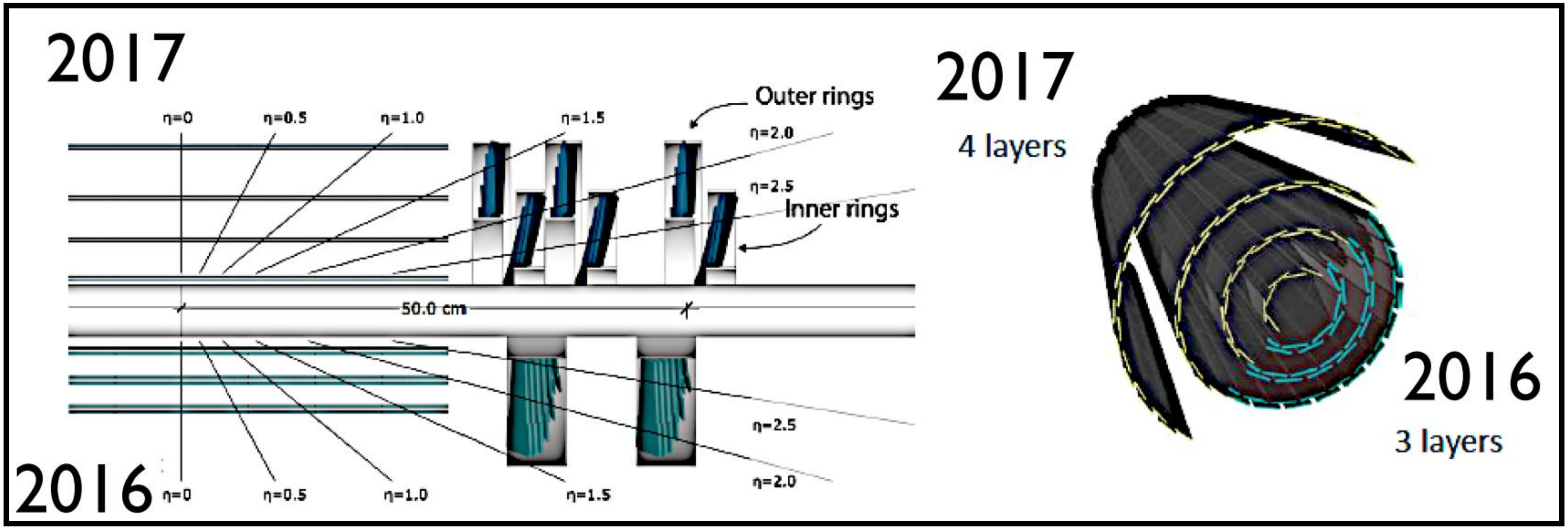
- Multipurpose detector at the LHC: silicon tracking, electromagnetic & hadronic calorimeters, a 3.8 T superconducting solenoid, & muon tracking chambers



Compact Muon Solenoid

- Multipurpose detector at the LHC: silicon tracking, electromagnetic & hadronic calorimeters, a 3.8 T superconducting solenoid, & muon tracking chambers

- **Pixel detector upgraded for 2017 data taking**
  - Large impact on b-tagging performance (as discussed later)

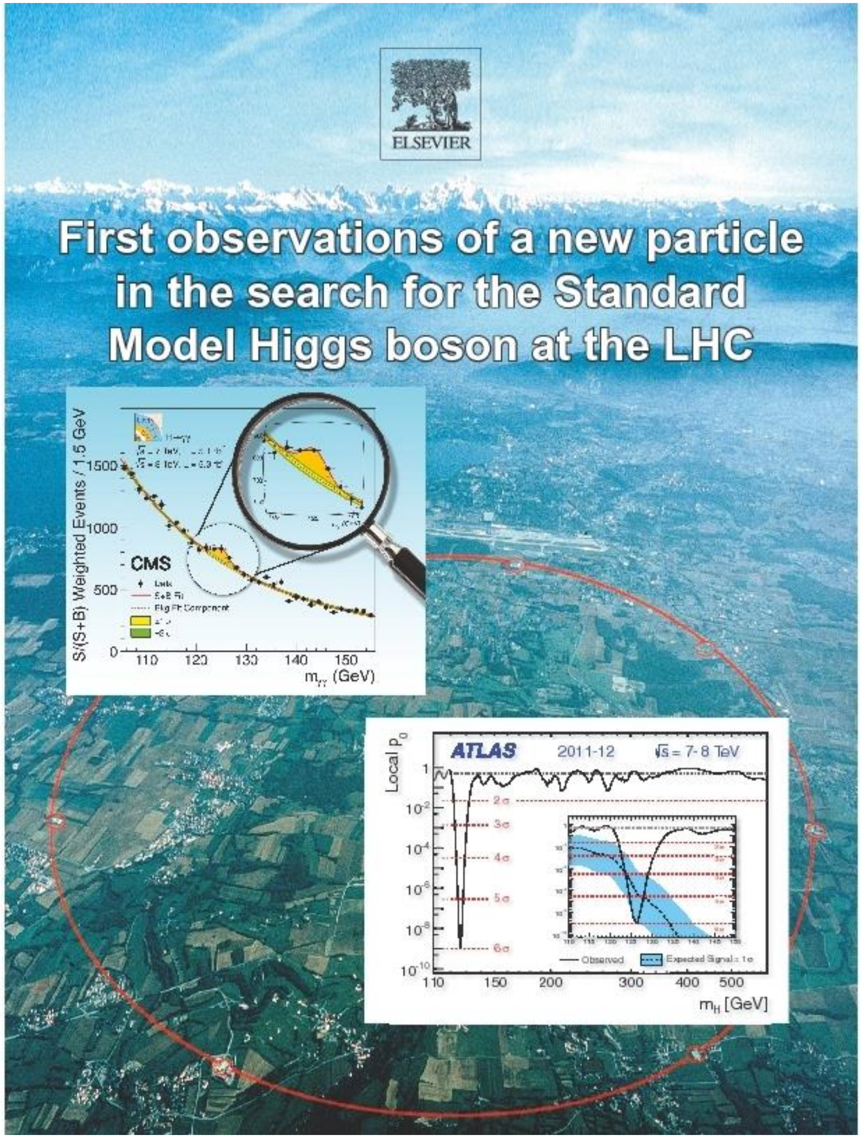


FROM THE HIGGS DISCOVERY TO H→bb

The Higgs boson discovery

July 4th 2012

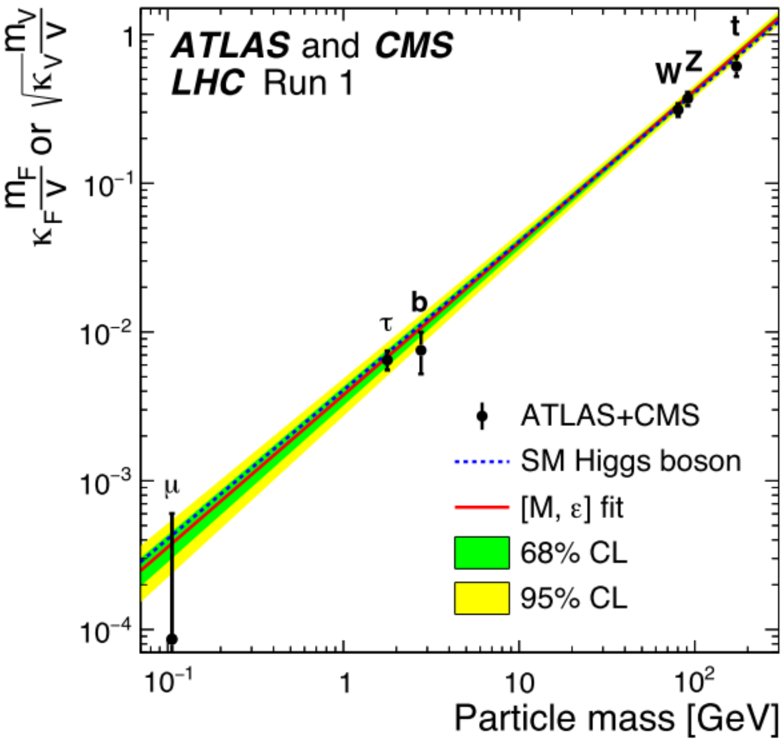
- **CMS and ATLAS reported independently the first observation of the Higgs boson**
  - Standard Model picture complete
- **Result driven by γγ and ZZ decay modes**
  - High mass resolution
  - High signal to background ratio
  - High signal trigger efficiency



ATLAS: PLB 716 (2012) 1-29
CMS: PLB 716 (2012) 30

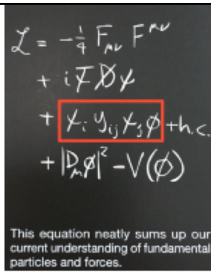
Established Higgs properties

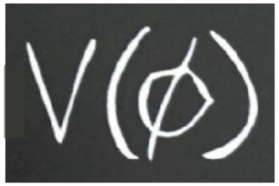
- **Mass:** $125.09 \pm 0.21\,(\mathrm{stat.}) \pm 0.11\,(\mathrm{syst.})\ \mathrm{GeV}$ ATLAS+CMS: PRL 114 (2015) 191803
- **Spin/Parity:** 0^+ ATLAS: EPJC 75 (2015) 476 CMS: PRD 92 (2015) 012004
- **Width:** < 1 GeV (direct)
  CMS: JHEP 11 (2017) 047
  < 0.015 GeV (indirect)
  ATLAS: arXiv:1808.01191 submitted to PLB
- **Observed direct coupling** to:
  - **Vector bosons** ATLAS: PLB 716 (2012) 1-29 CMS: PLB 716 (2012) 30
  - **τ leptons** ATLAS: ATLAS-CONF-2018-021 CMS: PLB 779 (2018) 283
  - **top quarks** ATLAS: PLB 784 (2018) 173 CMS: PRL 120 (2018) 231801



All measurements compatible with SM predictions

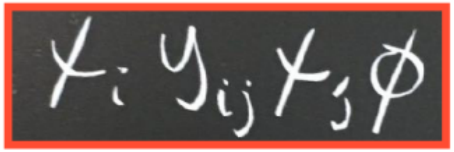
Why do Yukawa couplings matter?
(1) A part of the Higgs sector that's unlike any other experimentally-probed interaction





(-$\mu^2\varphi^2$ + $\lambda\varphi^4$, HHH) the keystone of the Higgs mechanism and Standard Model, familiar as QFT toy model, never probed in nature

(HWW, HZZ): A gauge interaction, with scalars rather than fermions; much like what we've seen before

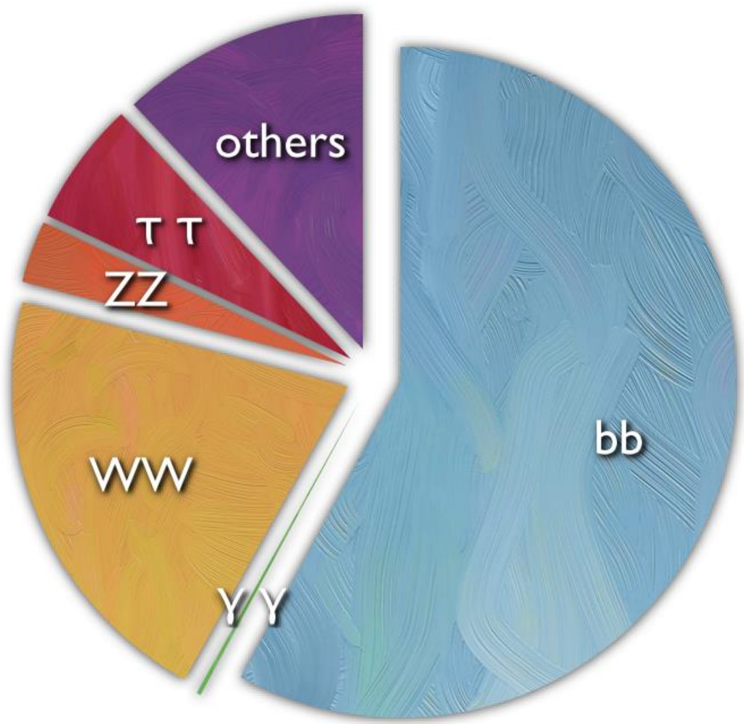


(Hbb, Htt, etc.): not a gauge interaction, and unlike anything we've probed before

G. Salam, LHCP '18

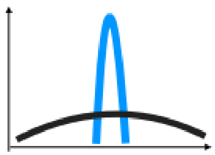
H→b$\bar{b}$: motivations

- **H→b$\bar{b}$ has the largest branching fraction (58%) for m_H=125 GeV**
- Unique final state to measure **coupling with down-type quarks**
- **Drives the uncertainty of the total Higgs boson width**
  - Limits the sensitivity to BSM contributions
- **Not yet observed**

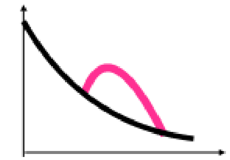


Challenges of H→b$\bar{b}$

- **H→b$\bar{b}$ compared with one of the discovery channels**



	$H \to 4\ell$	$H \to b\bar{b}$
Branching Ratio	0.03%	58%
mass resolution	1%	10%
S/B	2	0.05



- **H→b$\bar{b}$ search needs:**
  - Highly efficient **b-jets identification**
  - Excellent **resolution on m(bb)**
  - Use of full event information to **increase S/B**

- **Primary decay mode for searches at LEP and Tevatron**
  - More difficult at LHC due to increased background (S/B ~2.5x worse than Tevatron)

First H→bb̄ searches started at LEP…

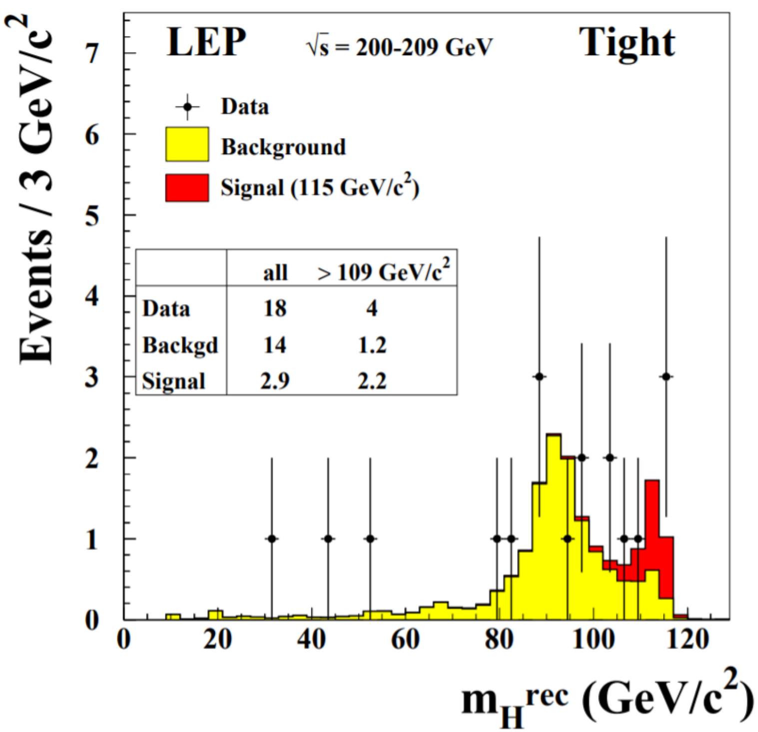
Physics Letters B 565 (2003) 61–75

Search for the Standard Model Higgs boson at LEP

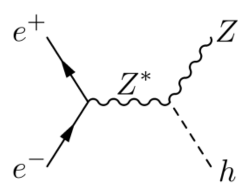
ALEPH Collaboration [1] DELPHI Collaboration [2] L3 Collaboration [3] OPAL Collaboration [4]
The LEP Working Group for Higgs Boson Searches [5]

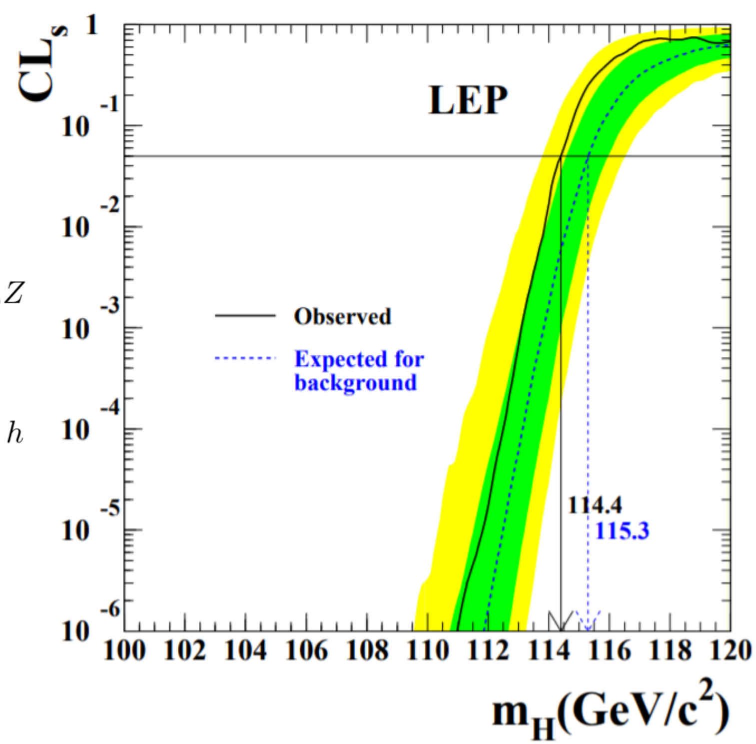
ELSEVIER
PHYSICS LETTERS B

m_H > 114.4 GeV @ 95%CL



	all	> 109 GeV/c²
Data	18	4
Backgd	14	1.2
Signal	2.9	2.2



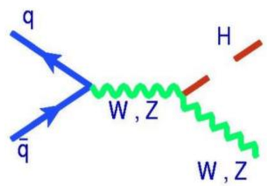


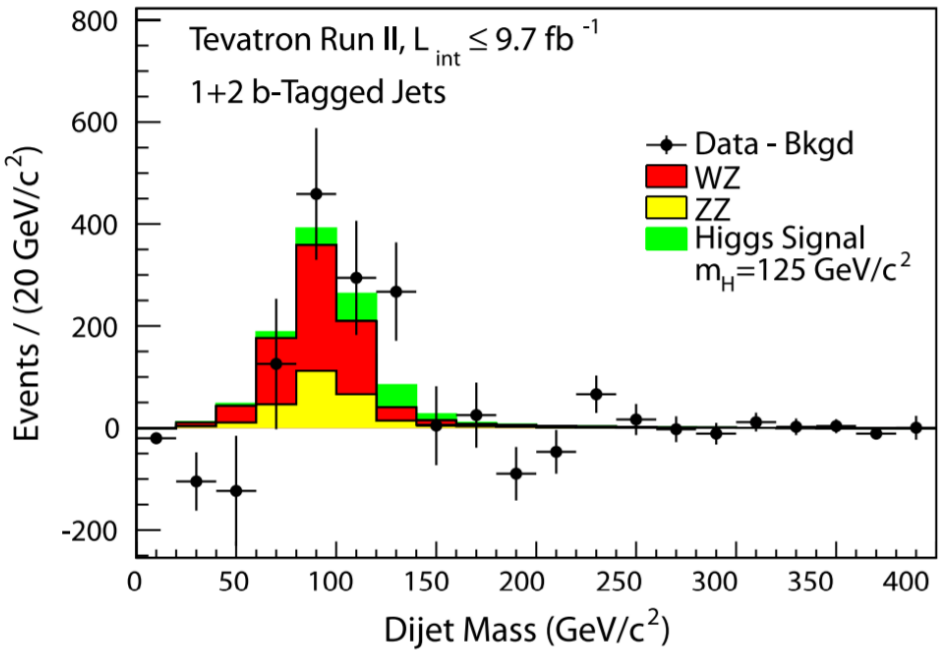
…and continued at Tevatron…

PRL 109, 071804 (2012) PHYSICAL REVIEW LETTERS week ending 17 AUGUST 2012

Evidence for a Particle Produced in Association with Weak Bosons and Decaying to a Bottom-Antibottom Quark Pair in Higgs Boson Searches at the Tevatron

(*CDF Collaboration)
(†D0 Collaboration)



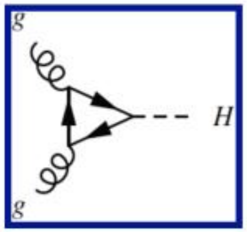


<u>Significance</u>

2.8σ observed @ 125 GeV

H→b$\bar{b}$ searches continue at the LHC

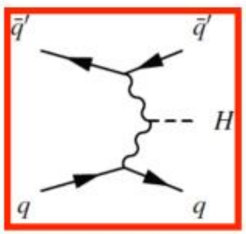
- Very large datasets at LHC give access to **several production modes to search for H→b$\bar{b}$**



Gluon Fusion (87%)

Overwhelming (**10^7 larger**) background of b-quark production due to strong interactions

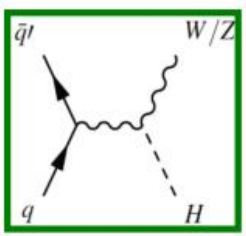
CMS: PRL 120 (2018) 071802



Vector-Boson Fusion (7%)

Very large background but a very distinctive topology
ISR photon to enhance S/B

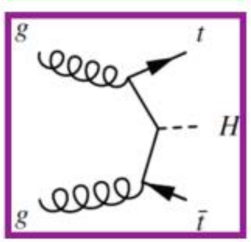
ATLAS: arXiv:1807.08639 submitted to PRD
ATLAS: JHEP 11 (2016) 112
CMS: HIG-16-003
CMS: PRD 92 (2015) 032008



Higgs-strahlung (4%)

leptons, E_T^{mis} to trigger and high p_T V suppress backgrounds

Most sensitive



Top Fusion ttH (1%)

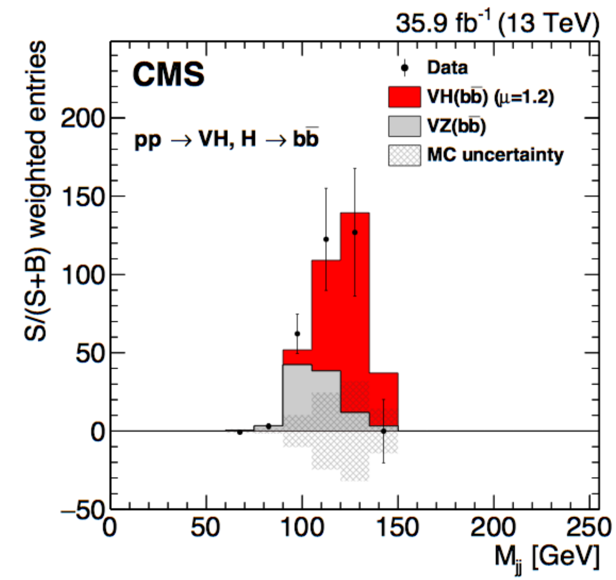
dominant background is t$\bar{t}$ + jets

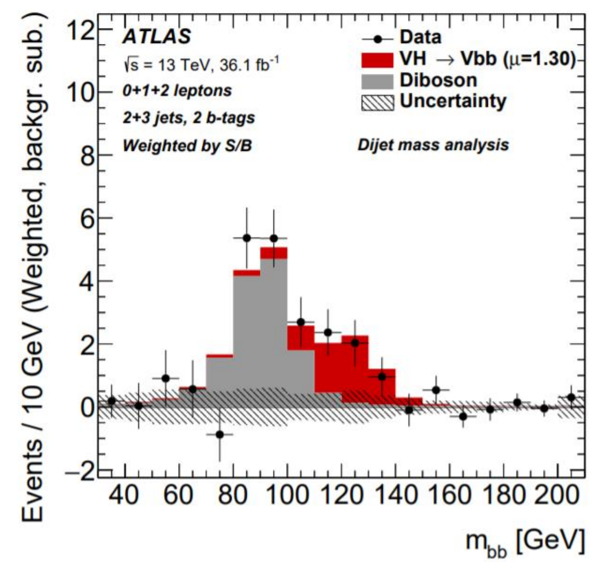
ATLAS: JHEP 05 (2016) 160
ATLAS: PRD 97, 072016 (2018)
CMS: JHEP 09 (2014) 087
CMS: arXiv:1804.03682 submitted to JHEP
CMS: JHEP 06 (2018) 101

VH($b\bar{b}$) results at LHC

- **VH(bb) evidence at LHC established with 2016 data by both ATLAS and CMS**
  - Detectors clearly demonstrated ability to deal with very high pile-up for such complex analysis
- **Signal strength uncertainty ~40%**

	signal strength	significance (exp)	significance (obs)
ATLAS Run 1 [1]	$0.52^{+0.40}_{-0.37}$	2.6σ	1.4σ
CMS Run 1 [2]	$0.89^{+0.47}_{-0.44}$	2.5σ	2.1σ
ATLAS+CMS Run 1 [3]	$0.79^{+0.29}_{-0.27}$	3.7σ	2.6σ
ATLAS 2015+2016 [4]	$1.20^{+0.42}_{-0.36}$	3.0σ	**3.5σ**
CMS 2016 [5]	$1.19^{+0.40}_{-0.38}$	2.8σ	**3.3σ**





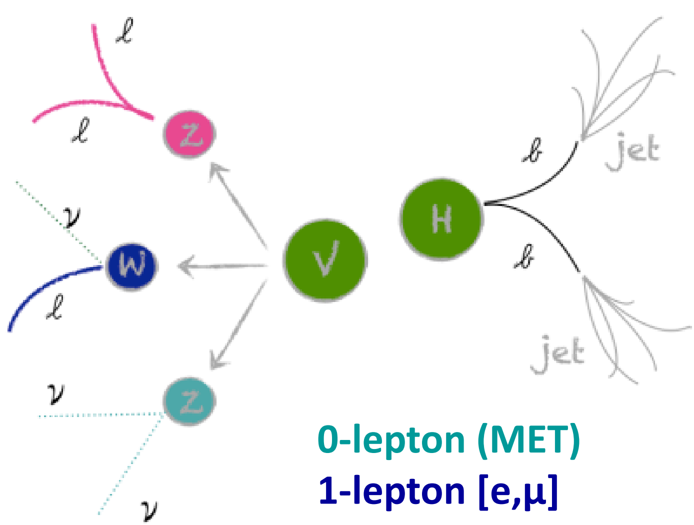
[1] JHEP 01 (2015) 069
[2] JHEP 08 (2016) 045
[3] JHEP 08 (2016) 045
[4] JHEP 12 (2017) 024
[5] PLB 780 (2018) 501

VH(bb) WITH 2017 DATA

VH(bb): Analysis strategy

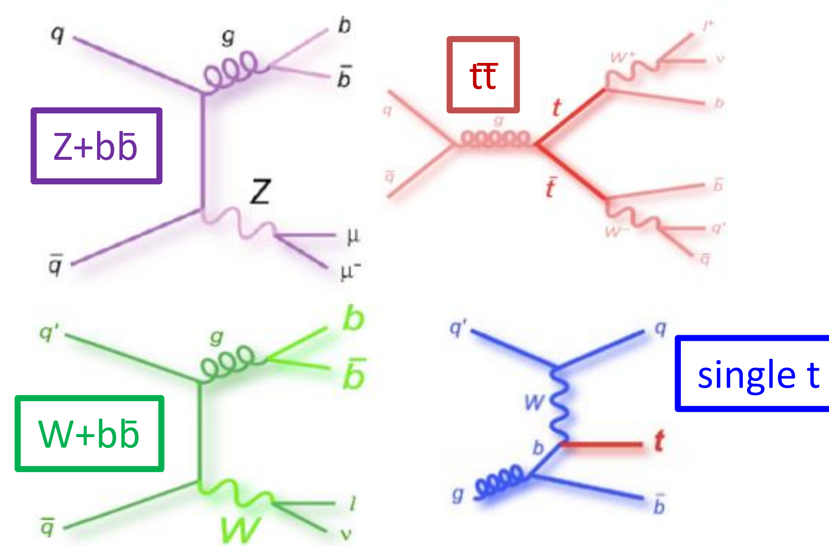
- **Analysis strategy:**
  - **3 channels** with 0, 1, and 2 leptons and 2 b-tagged jets
    - To target Z(νν)H(bb), W(lν)H(bb)and Z(ll)H(bb) processes
  - **Signal region designed to increase S/B**
    - **Large boost** for vector boson
    - **Multivariate analysis** exploiting the most discriminating variables ($m_{b\bar{b}}$, $\Delta R_{b\bar{b}}$, b-tag)
  - **Control regions** to validate backgrounds and control/constrain normalizations

signal



0-lepton (MET)
1-lepton [e,μ]
2-leptons [ee,μμ]

irreducible backgrounds

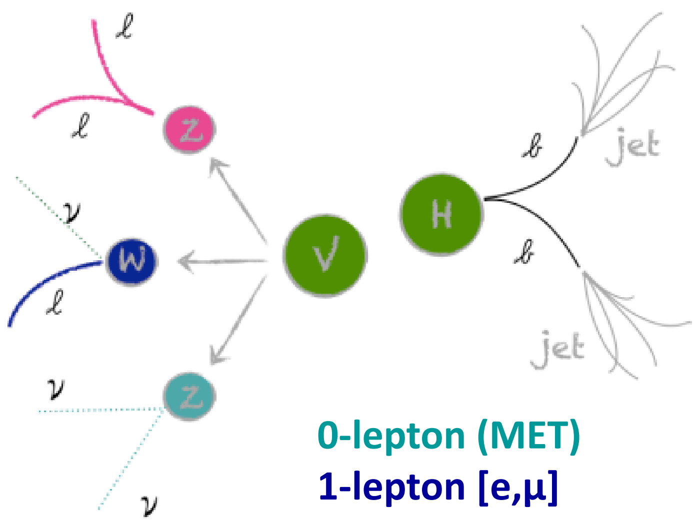


normalization from data, shapes from MC

VH(bb): Analysis strategy

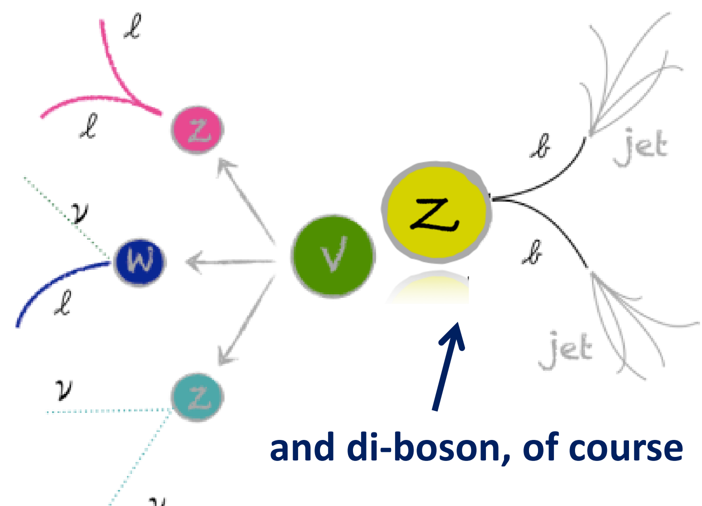
- **Analysis strategy:**
  - **3 channels** with 0, 1, and 2 leptons and 2 b-tagged jets
    - To target Z(νν)H(bb), W(lν)H(bb)and Z(ll)H(bb) processes
  - **Signal region designed to increase S/B**
    - **Large boost** for vector boson
    - **Multivariate analysis** exploiting the most discriminating variables ($m_{b\bar{b}}$, $\Delta R_{b\bar{b}}$, b-tag)
  - **Control regions** to validate backgrounds and control/constrain normalizations

signal



0-lepton (MET)
1-lepton [e,μ]
2-leptons [ee,μμ]

irreducible backgrounds



and di-boson, of course

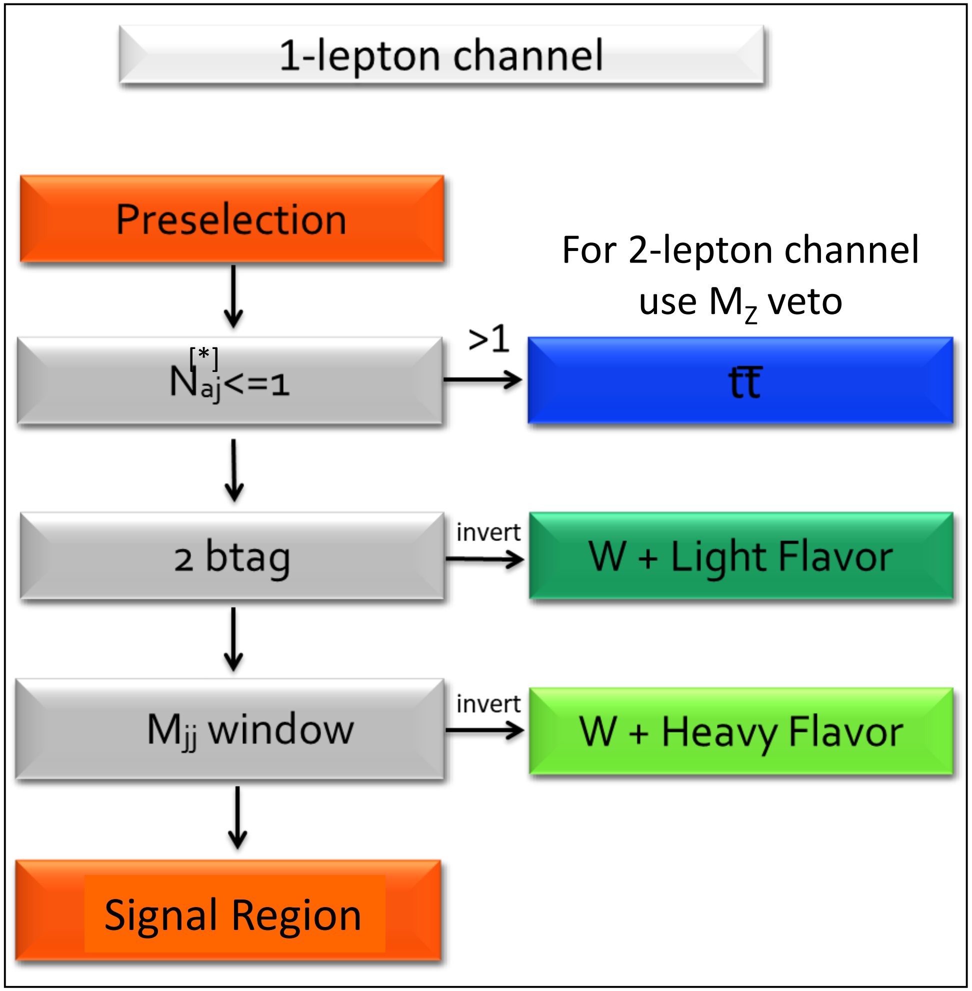
used to validate the analysis strategy

Event Selection/Categorization

- **Selections** (jets, leptons, b-tagging) **optimized separately by channel**
  - 4 analysis categories:
    - 0-lepton: **p_T(Z) > 170 GeV**
    - 1-lepton: **p_T(W) > 150 GeV**
    - 2-lepton: **p_T(Z) > 150 GeV**
    - 2-lepton: **50 < p_T(Z) < 150 GeV**

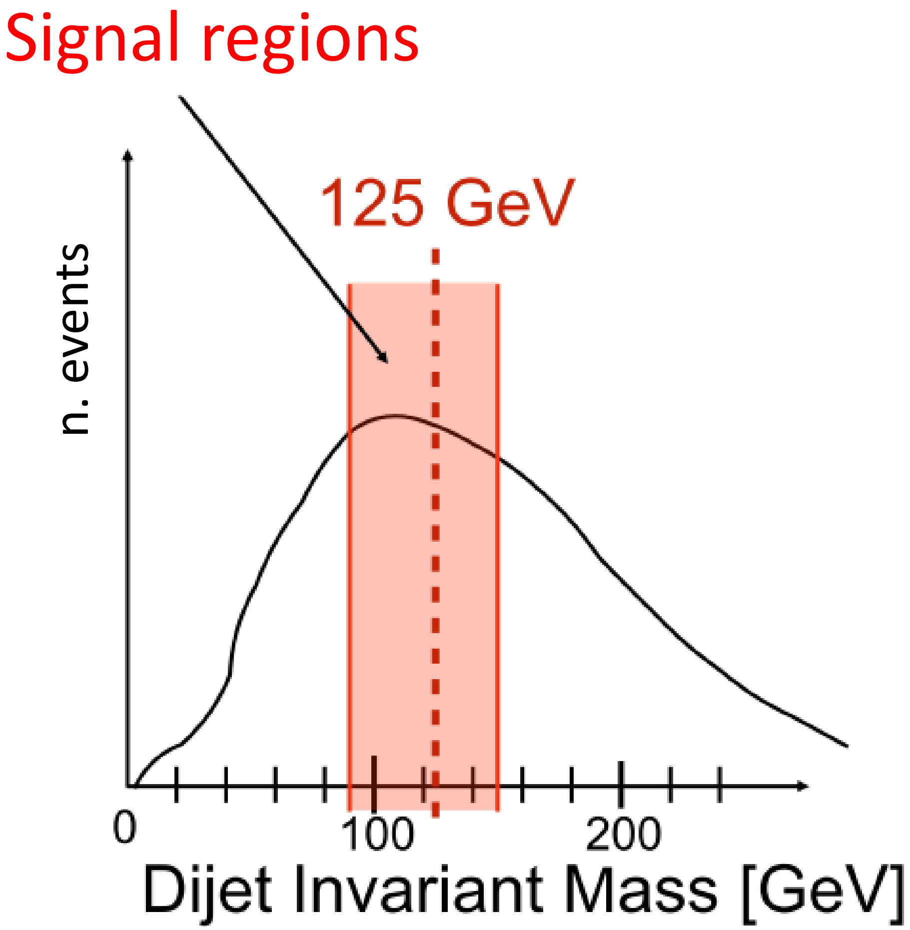
- **Control regions** designed to map closely each signal region
  - Inverted selections to **enhance purity in targeted backgrounds: tt̄, V+light flavor, and V+heavy flavor**

Control region example

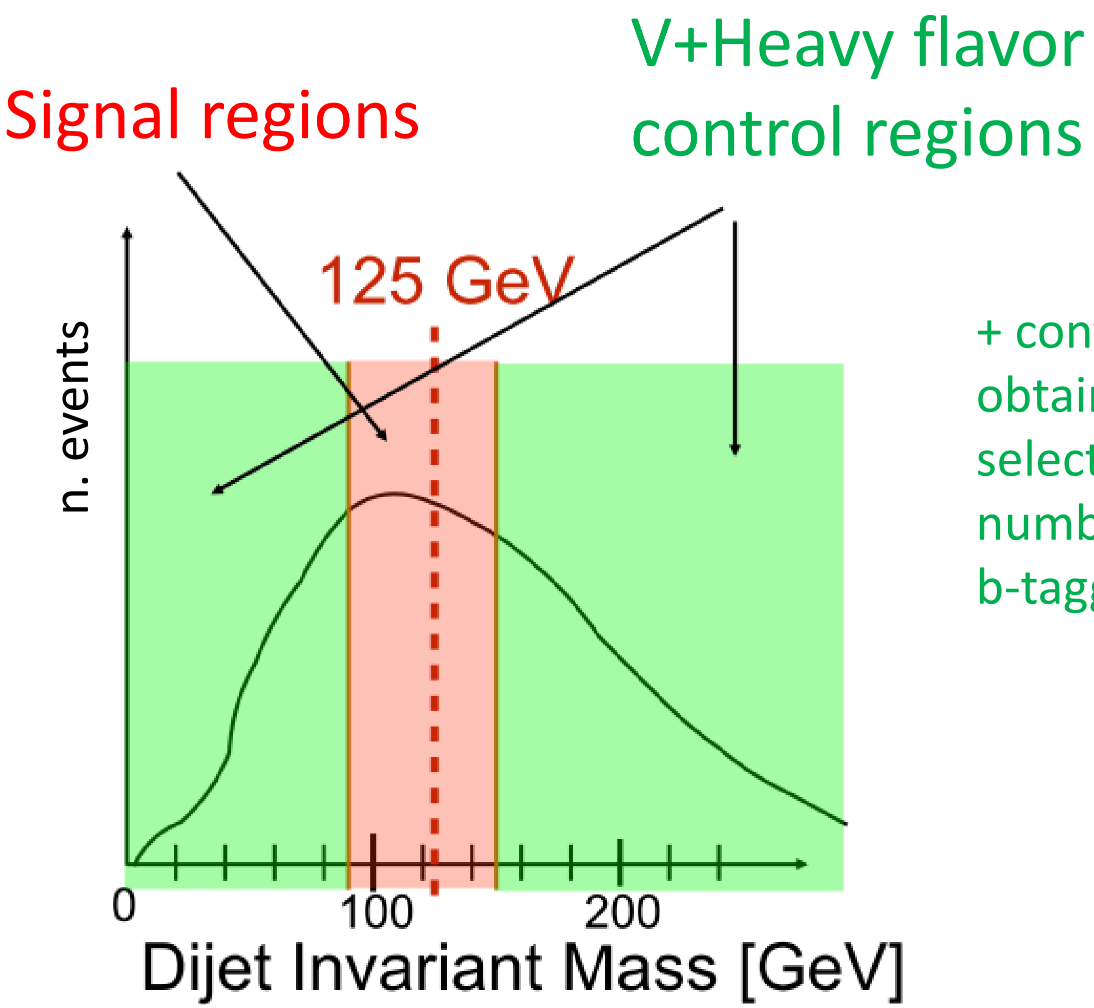


[*] Number of additional jets in the event

Fit strategy in a nutshell



Fit strategy in a nutshell

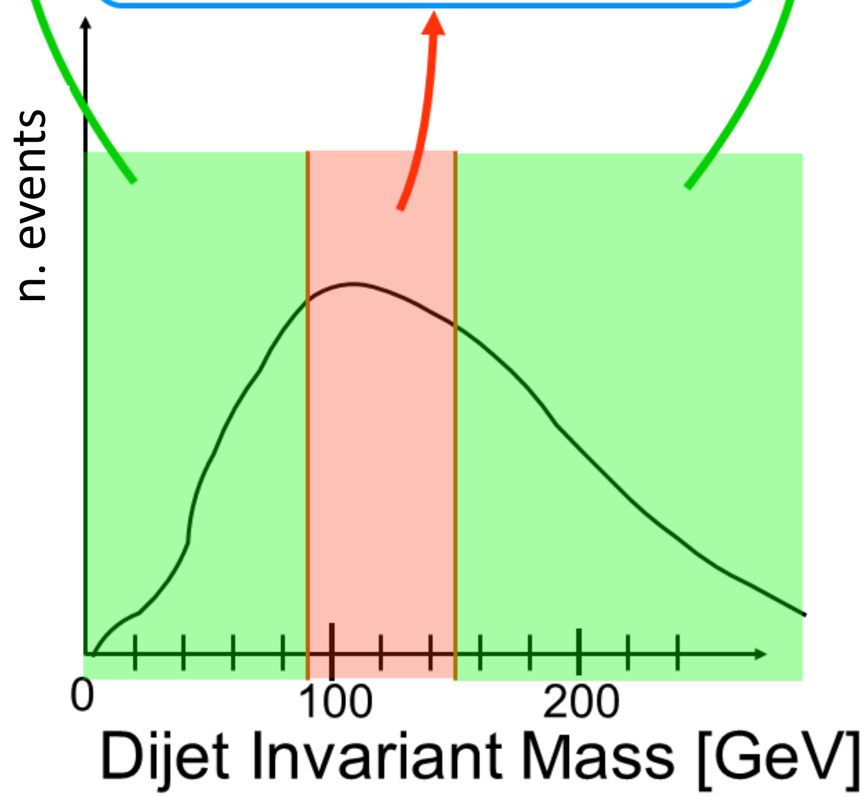


+ control regions obtained inverting selections on number of jets, b-tagging, m(ll)

Fit strategy in a nutshell

Simultaneous fit to extract:

- **Data/MC background normalization factors**
- **Signal strength, significance**



+ control regions obtained inverting selections on number of jets, b-tagging, m(ll)

VH(bb) in 2017: main features

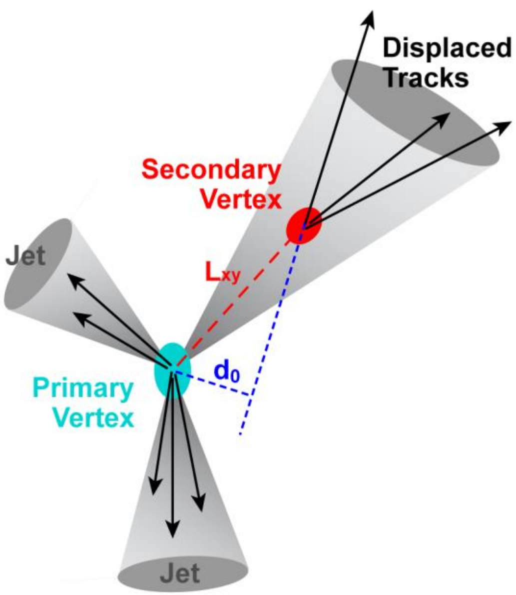
- **Improved mass resolution from:**
  - Better **b-jet identification**
  - New **b-jet energy regression**
  - **Kinematic fit in 2-lepton channel**
  - FSR jet recovery

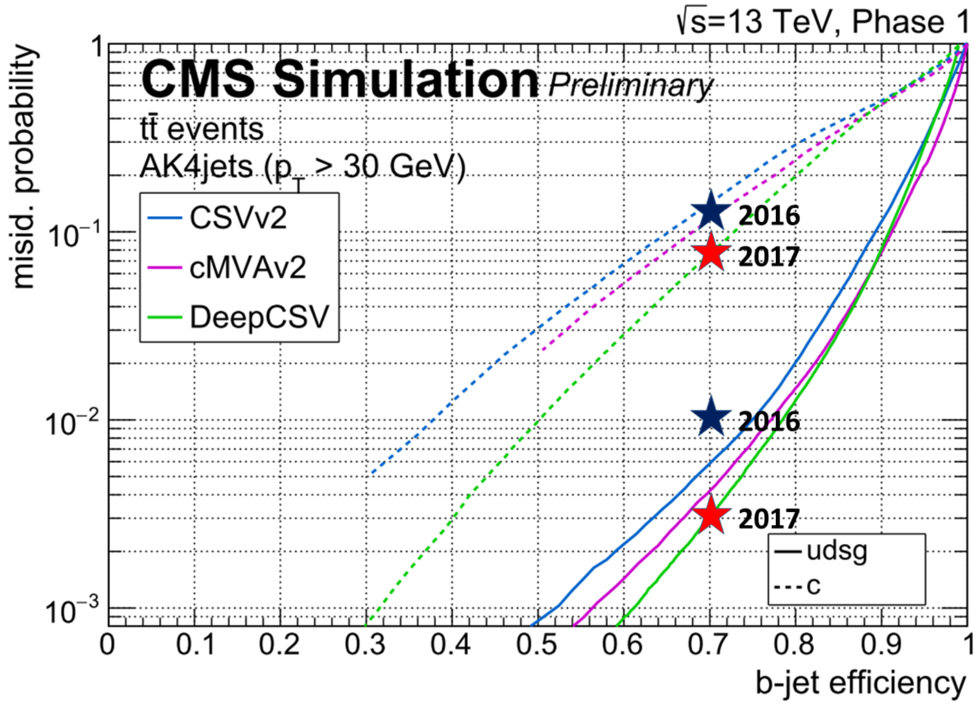
- Use of **deep neural network** (DNN) to discriminate:
  - Signal from background, in Signal Regions
  - Background components among each other, in Control Regions

- Combined effect: O(5-10%) increase of the analysis sensitivity wrt 2016, depending on channel

b-jet identification

- **Continuous effort to improve b-tagging at CMS**
  - **New pixel detector** (4 layers)
  - **DNN algorithm (DeepCSV) with additional per-track information**
  - **Contamination from q/g < 1% for efficiency ~70%**
- MC corrections derived on data with $t\bar{t}$ events
- Good agreement between data and MC verified in all analysis regions

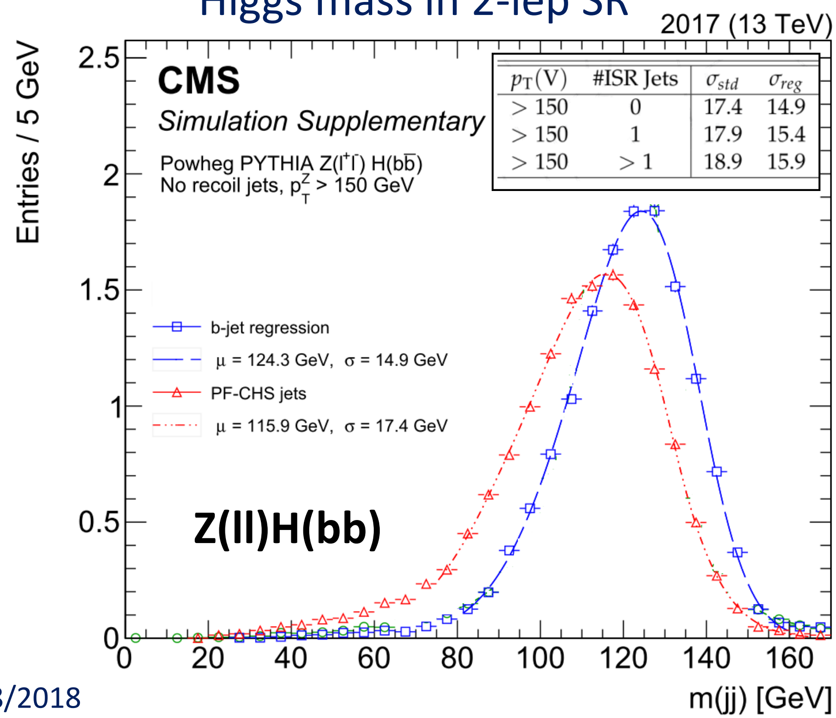




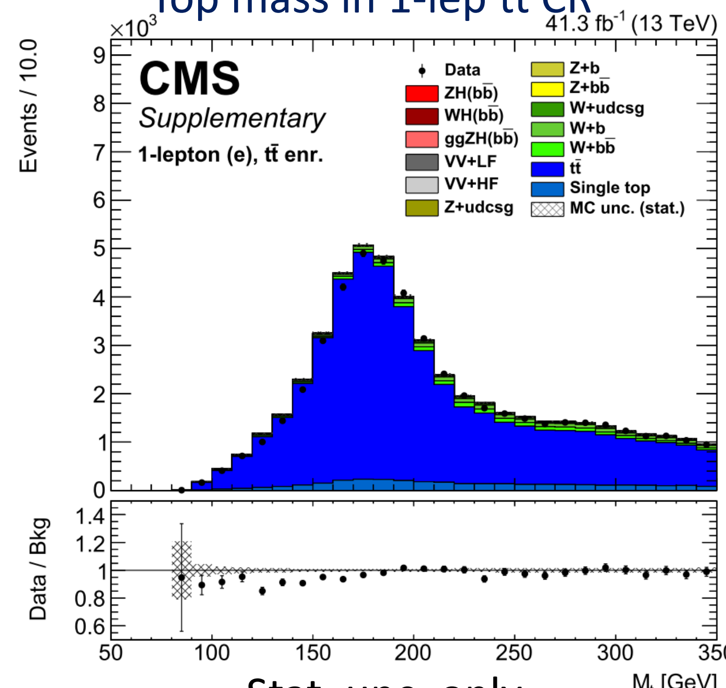
b-jet energy regression

- **Regression mainly recovers missing energy in the jet due to neutrino**
  - Switch from Boosted Decision Trees to DNN algorithm
- **Extended set of input variables** now including lepton flavor (μ/e), jet mass, fragmentation-like variable, energy fractions in ΔR rings
- **Significant m(bb) resolution improvement without sculpting of the background**
  - σ/peak down to 11.9% in 2017 wrt 13.2% in 2016

Higgs mass in 2-lep SR



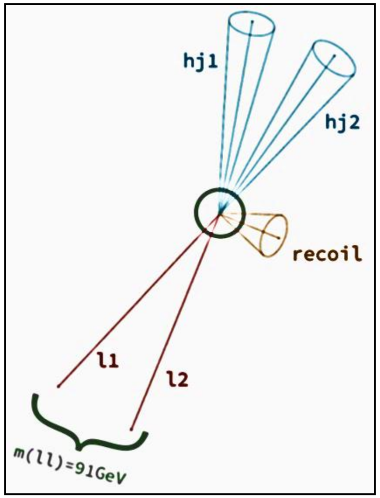
Top mass in 1-lep tt CR

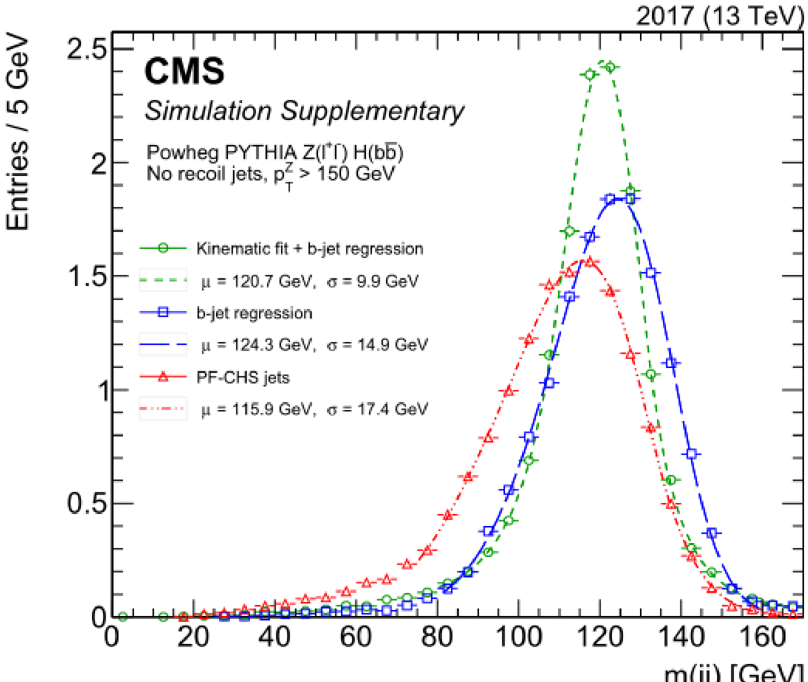


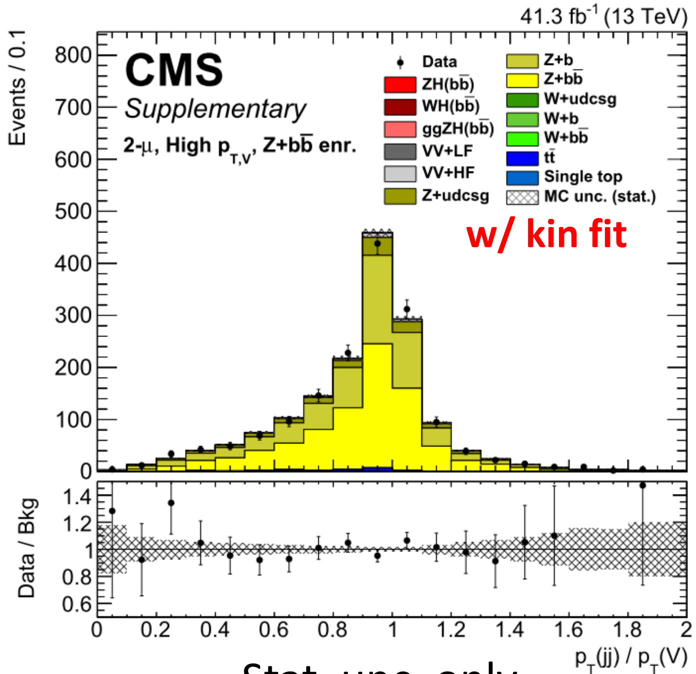
Stat. unc. only

Kinematic fit in 2-lepton channel

- **No intrinsic missing energy in the Z(ll)H($b\bar{b}$) process**
- Improve jet p_T measurement through kinematic fit procedure
  - Constrain dilepton system to Z mass
  - **Balance the ll+bb+j system in the (p_x,p_y) plane**
- Improvement up to 36% on m($b\bar{b}$) resolution





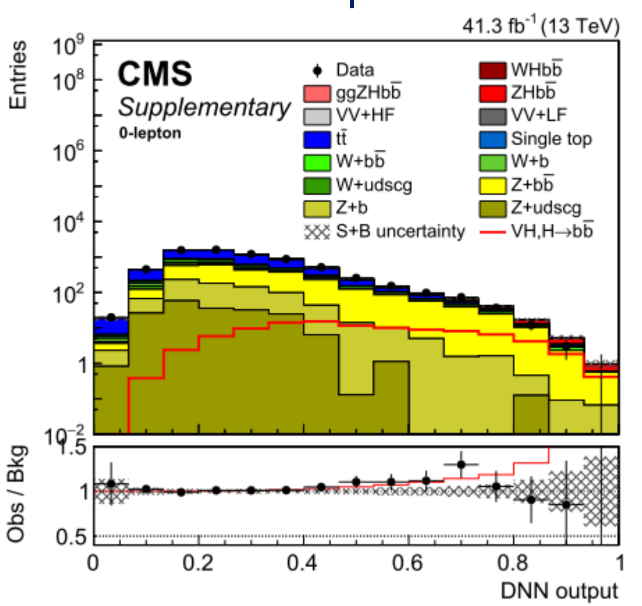


Stat. unc. only

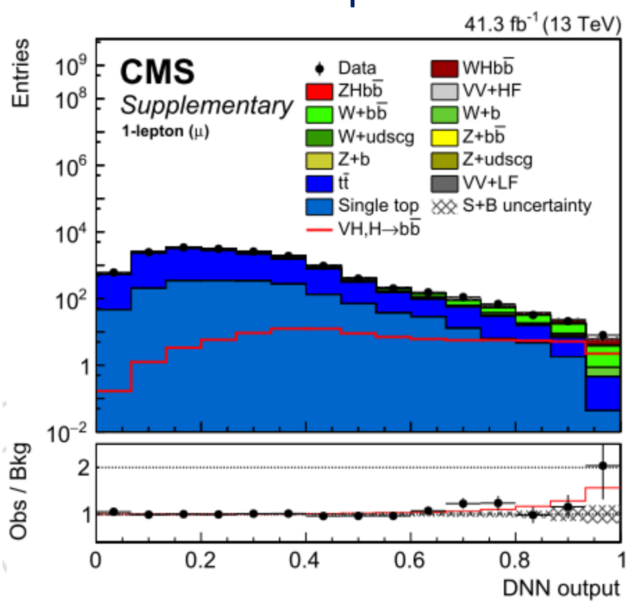
Signal vs Background discriminator

- **DNN discriminator used to extract signal**
  - **Input variables: b-jet properties, di-jet kinematics, event topology,** carefully validated through data/MC comparison
  - Trained separately in each channel
  - Performance optimization with blind analysis

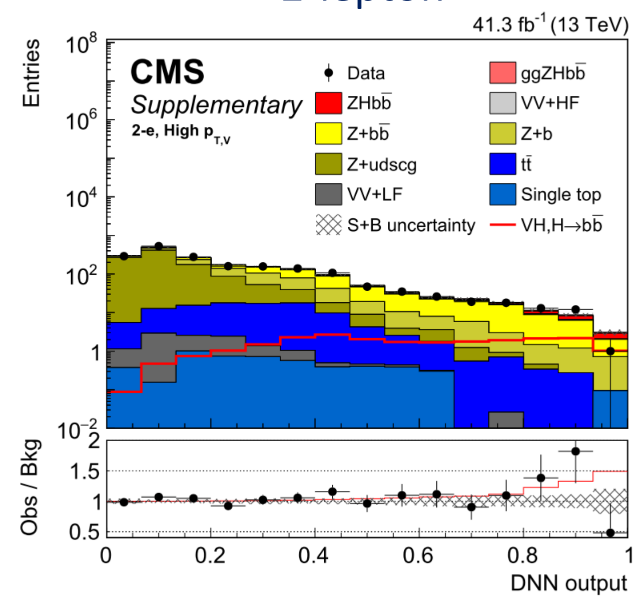
0-lepton



1-lepton



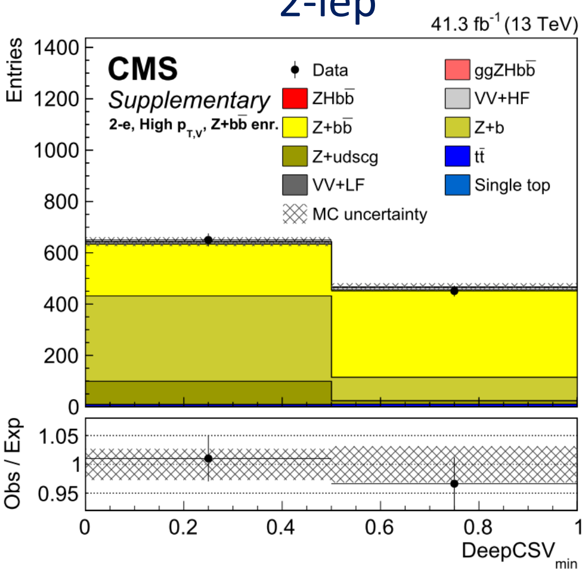
2-lepton



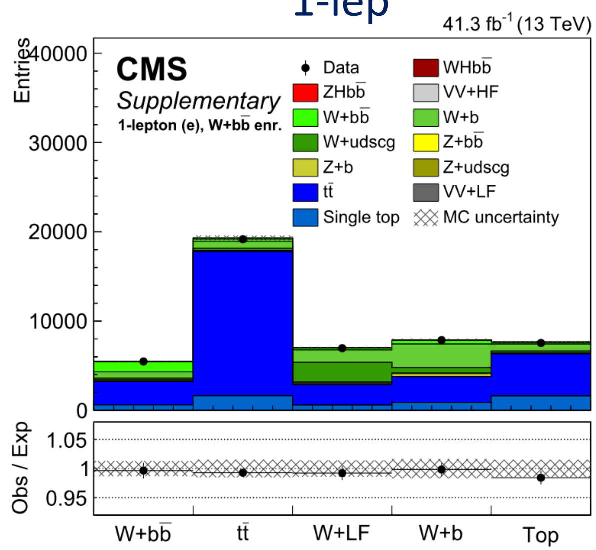
Heavy Flavor control region discriminators

- Reminder: leading systematic uncertainty from normalization of V+(b)b
- **2-lepton channel control region very pure**
  - Fit **b-tag** shape (DeepCSV) to discriminate processes
- **0- and 1-lepton channel control regions less pure**
  - Fit **DNN multi-categorizer** to distinguish among background components
    - Use same input variables as Signal vs Background discriminator

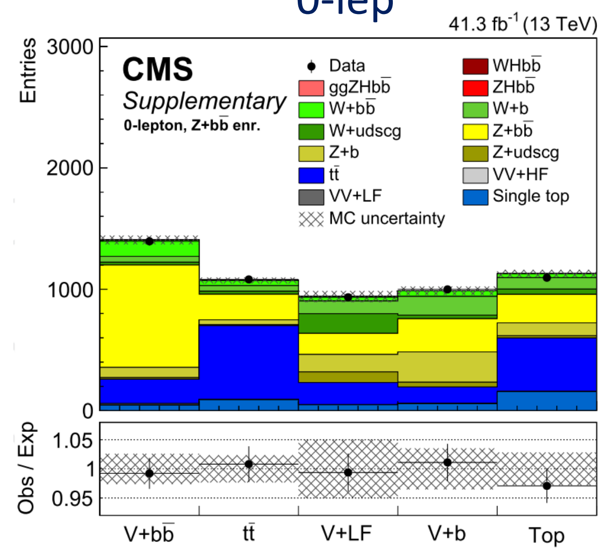
2-lep



1-lep



0-lep



Fit setup and background normalization

- **Simultaneous fit of Signal and Control Regions to extract signal and background normalizations**
  - Fitted variables: DNN or b-tagging shapes, or yields depending on the region
- **MC shapes and normalizations floated within constraints from systematic uncertainties through nuisance parameters**

Process	$Z(\nu\nu)H$	$W(\ell\nu)H$	$Z(\ell\ell)H$ low-p_T	$Z(\ell\ell)H$ high-p_T
$W + \text{udscg}$	1.04 ± 0.07	1.04 ± 0.07	–	–
$W + b$	2.09 ± 0.16	2.09 ± 0.16	–	–
$W + b\bar{b}$	1.74 ± 0.21	1.74 ± 0.21	–	–
$Z + \text{udscg}$	0.95 ± 0.09	–	0.89 ± 0.06	0.81 ± 0.05
$Z + b$	1.02 ± 0.17	–	0.94 ± 0.12	1.17 ± 0.10
$Z + b\bar{b}$	1.20 ± 0.11	–	0.81 ± 0.07	0.88 ± 0.08
$t\bar{t}$	0.99 ± 0.07	0.93 ± 0.07	0.89 ± 0.07	0.91 ± 0.07

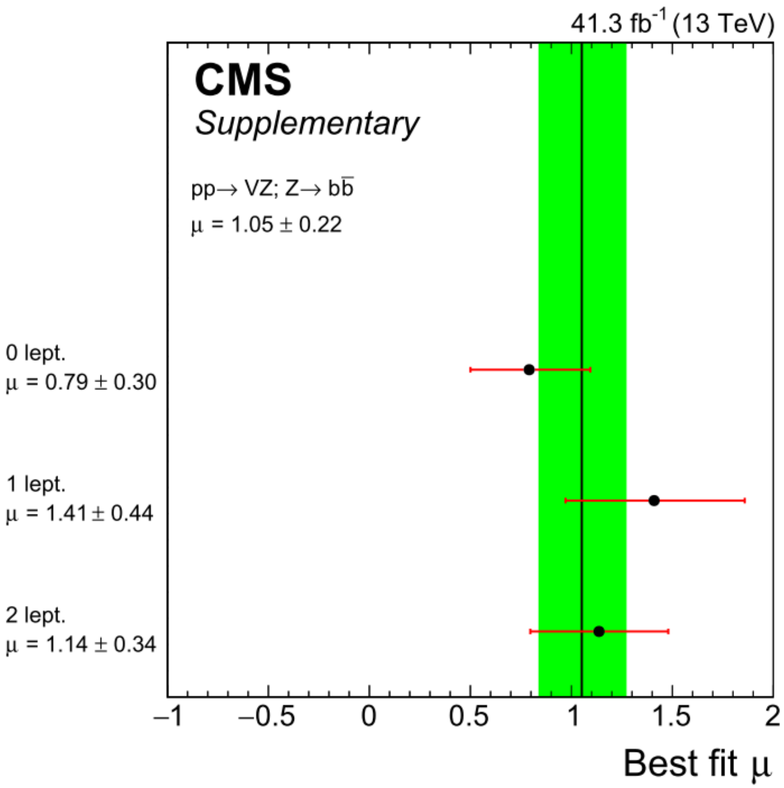
Systematic uncertainties

- **Total uncertainty ~34%, statistically dominated**
- Major sources of systematic uncertainties from **background normalization and modeling, b-tagging, MC sample size**

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

Validation: VZ(bb)

- **VZ analysis** using Z(bb) standard candle next to H(bb) peak
- **Same "technology" used for VH(bb) fit**
  - Same DNN inputs (but dedicated training), same Control Regions, VH(bb) normalized to SM and left free to float
  - Larger m(bb) window in Signal Region to fully include Z(bb) peak



Significance

5.0σ expected

5.2σ observed

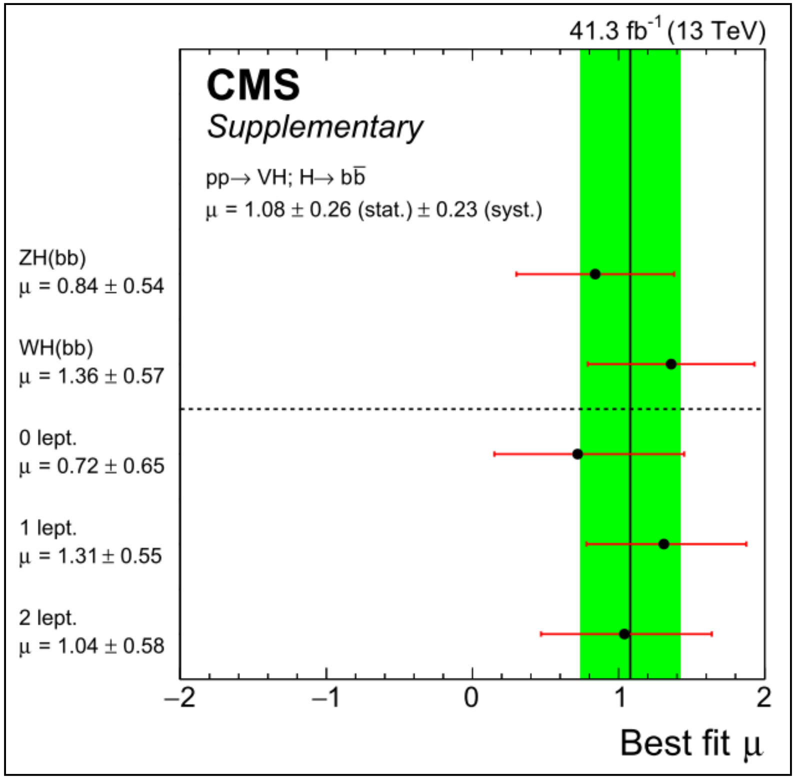
Signal strength

μ = 1.05 ± 0.22

VH(bb) Results with 2017 data

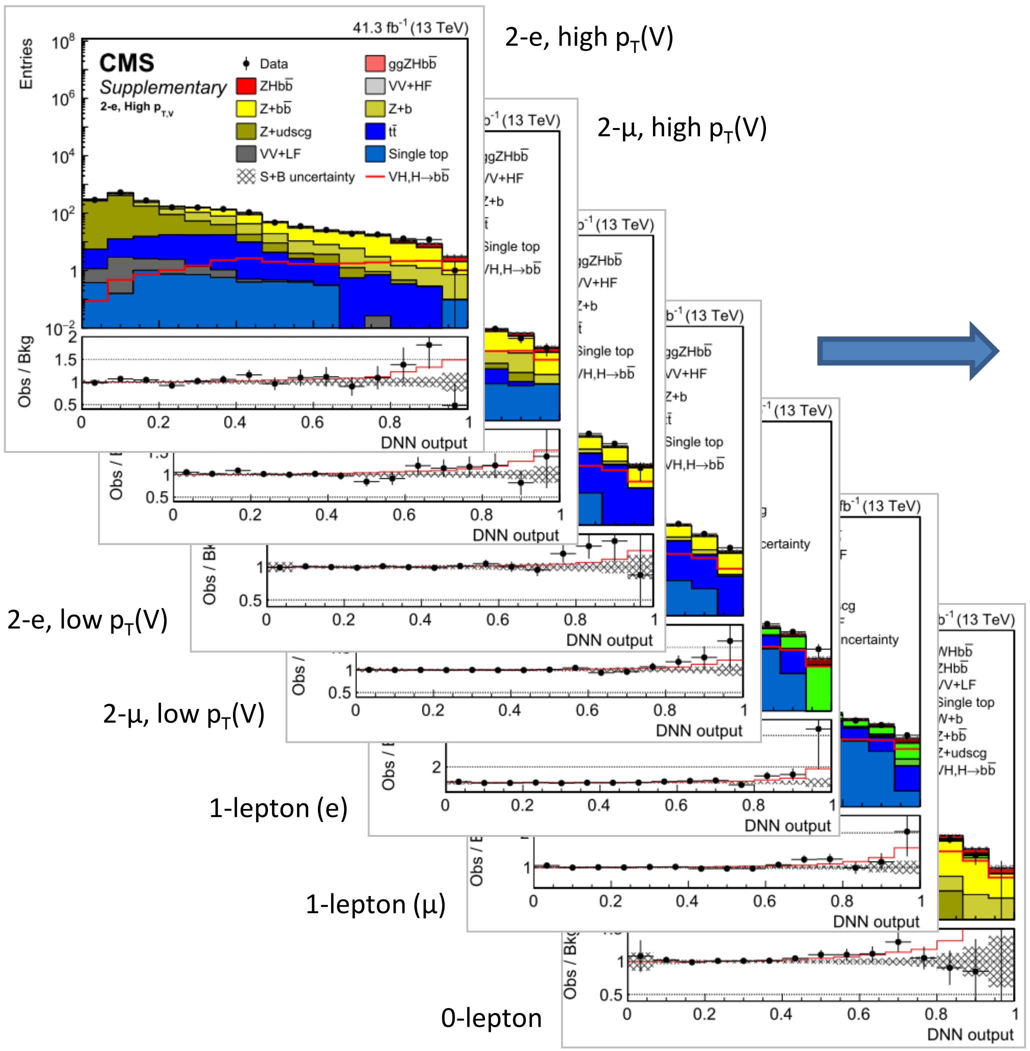
- **Results with 2017 data compatible with SM expectations**
  - **Observed significance 3.3σ, signal strength 1.08 ± 0.34**
  - O(5-10%) increase in analysis sensitivity wrt 2016, depending on channel
  - Remarkable channel compatibility

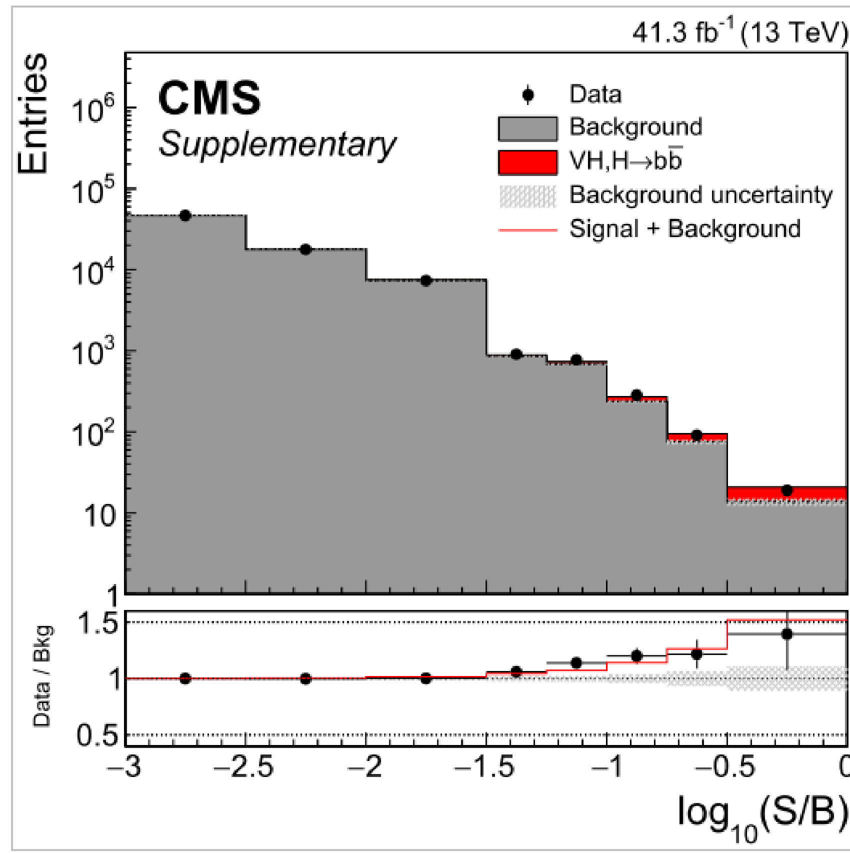
Data set	Significance (σ) Expected	Observed	Signal strength
2017			
0-lepton	1.9	1.3	0.73 ± 0.65
1-lepton	1.8	2.6	1.32 ± 0.55
2-lepton	1.9	1.9	1.05 ± 0.59
Combined	3.1	3.3	1.08 ± 0.34
2016	2.8	3.3	1.2 ± 0.4



Combining DNN distributions

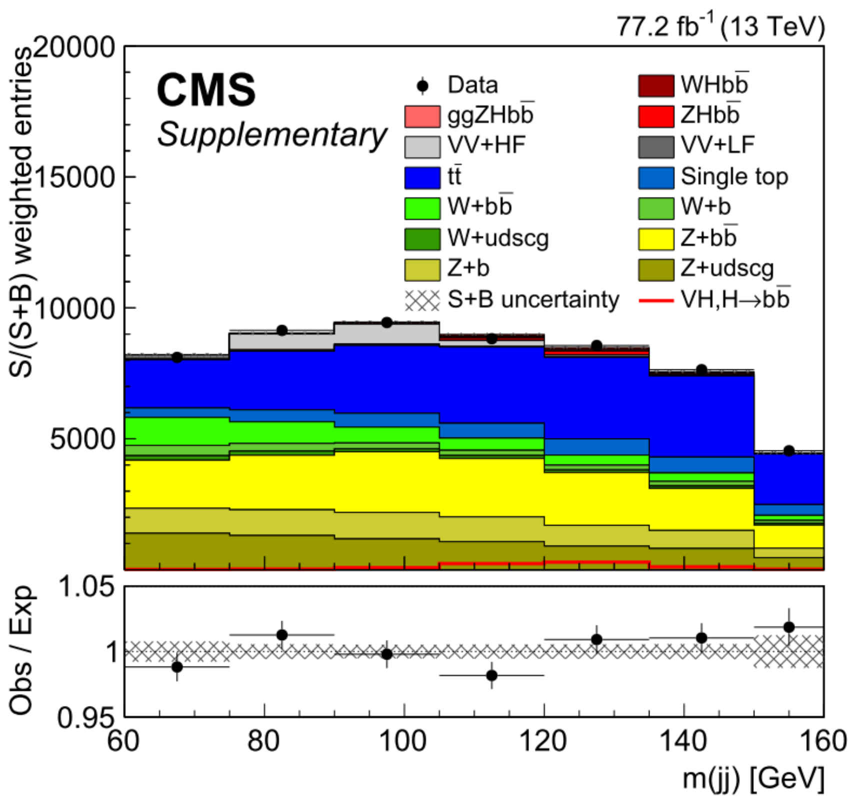
- DNN distributions sorted into bins of similar S/B ratio and combined
- Excess well compatible with SM Higgs boson signal hypothesis

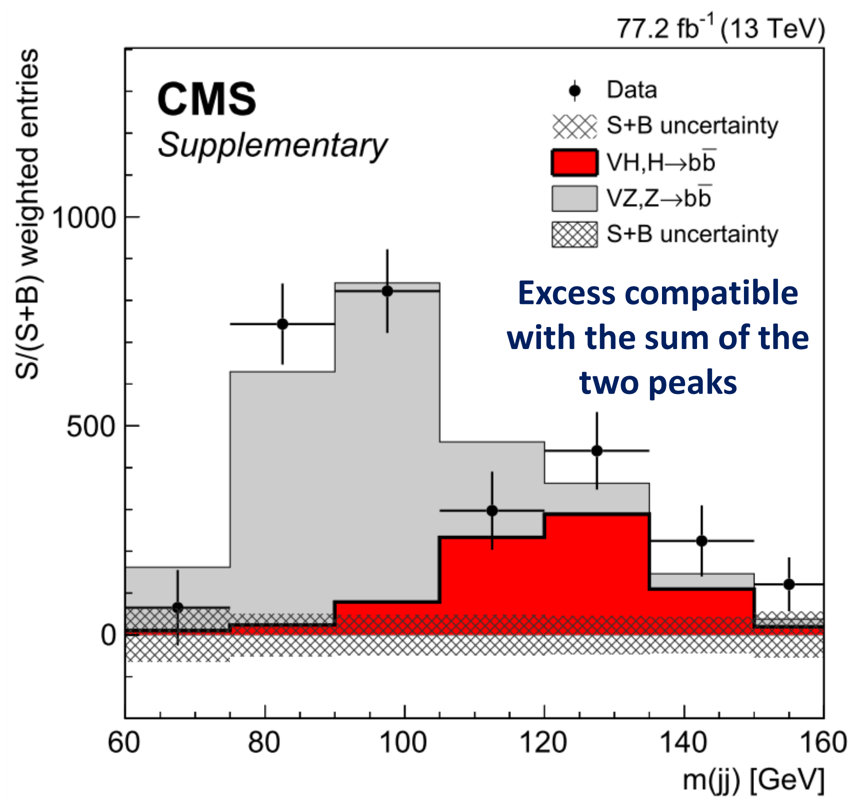




Visualizing the excess: m(jj) analysis

- **Fit to the m(jj)**: lower sensitivity but direct visualization of the Higgs boson signal
- **Events categorized in DNN sensitivity after removing correlations with m(jj)**
- m(jj) distributions combined and weighted by S/(S + B)
- Signal strengths compatible with main analysis

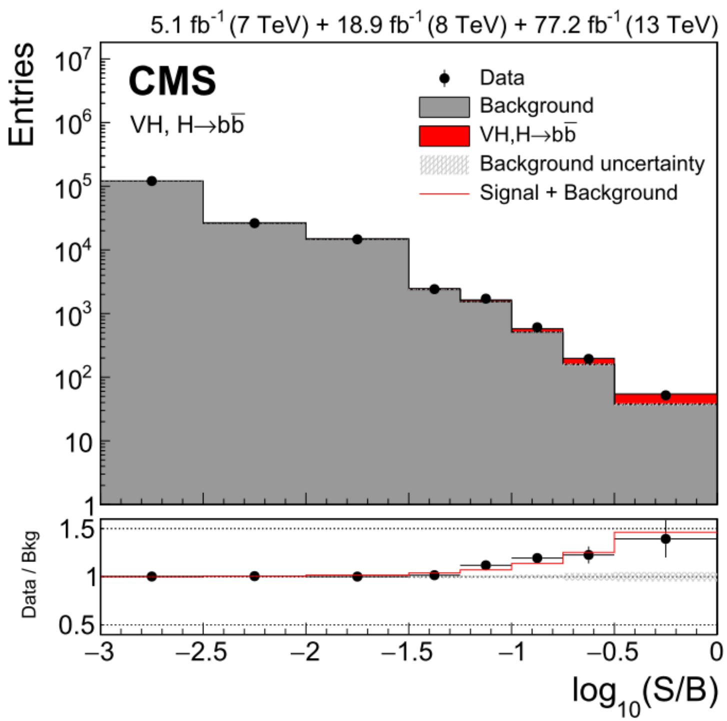


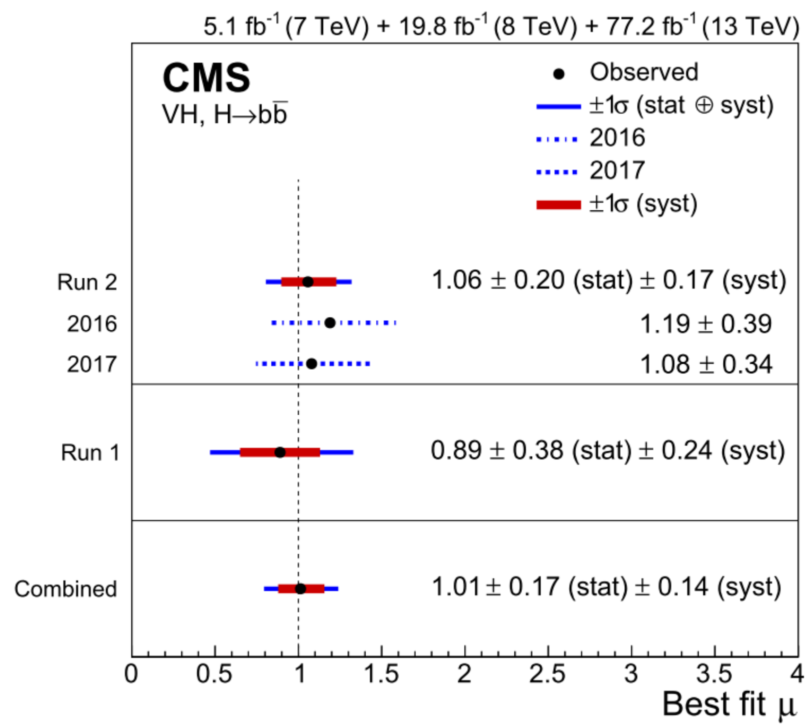


COMBINATION OF VH(bb) RESULTS

Combination of VH(bb) results

Data set	Significance (σ) Expected	Observed	Signal strength
2017	3.1	3.3	1.08 ± 0.34
Run 2 (2016+2017)	4.2	4.4	1.06 ± 0.26
Run 1 + Run 2	4.9	4.8	1.01 ± 0.23

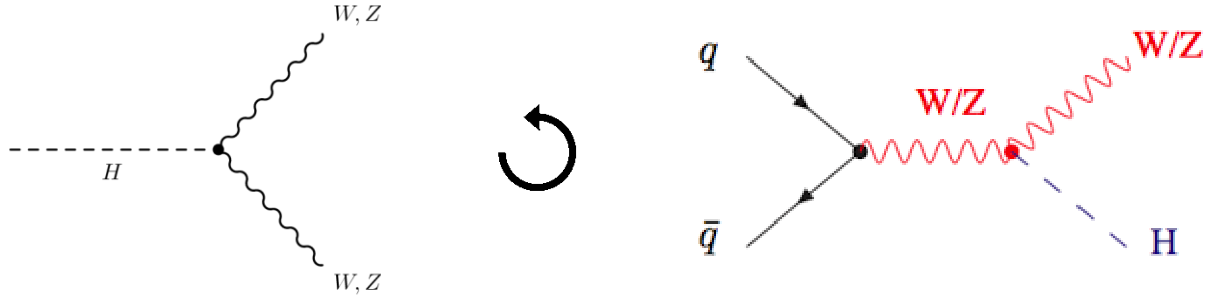




Combination of VH(bb) results

Data set	Significance (σ) Expected	Observed	Signal strength
2017	3.1	3.3	1.08 ± 0.34
Run 2 (2016+2017)	4.2	4.4	1.06 ± 0.26
Run 1 + Run 2	4.9	4.8	1.01 ± 0.23

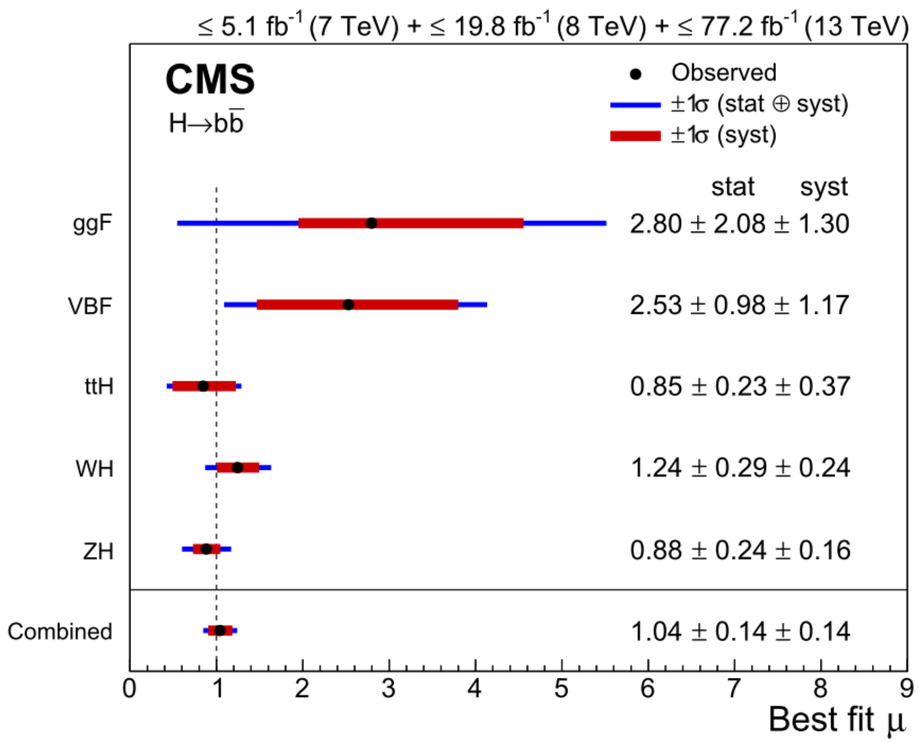
NB: 5σ observation of VH production in reach if VH(ττ) from HIG-18-007 is added.
This result is not contained in the paper as VH probes the HVV coupling, already established in Run 1



COMBINATION OF H$\rightarrow$bb RESULTS

Combination of H→bb measurements

- **Combination of CMS H→bb measurements : VH, boosted ggH, VBF, ttH**
- **Most sources of systematic uncertainty are treated as uncorrelated**
  - Theory uncertainties are correlated between all processes and data sets
- **Measured signal strength is μ = 1.04 ± 0.20**



Significance

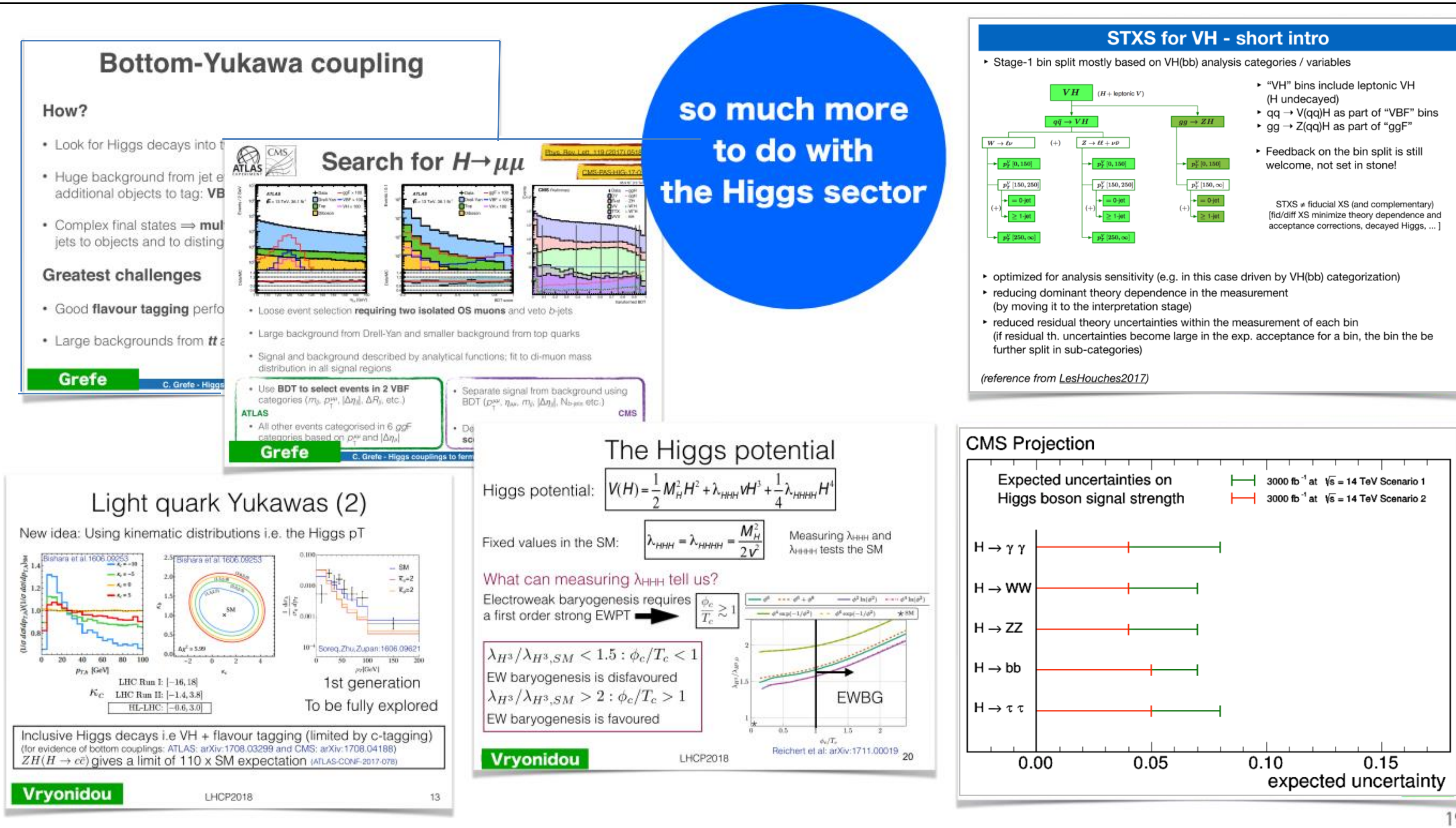
5.5σ expected

5.6σ observed

Observation of the H→bb decay by the CMS Collaboration

FUTURE PROSPECTS

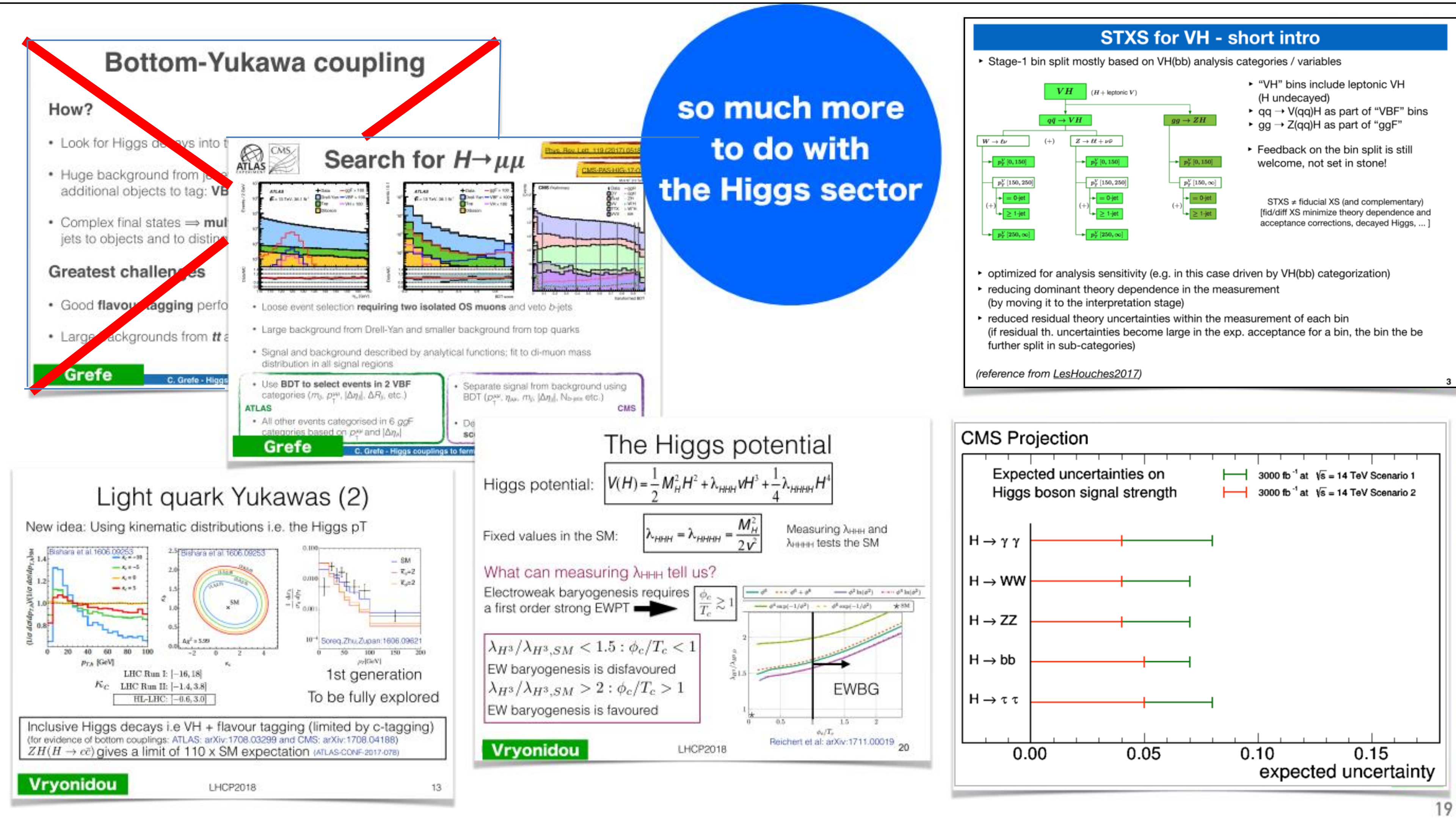
June '18



adapted from G. Salam, LHCP '18

Today

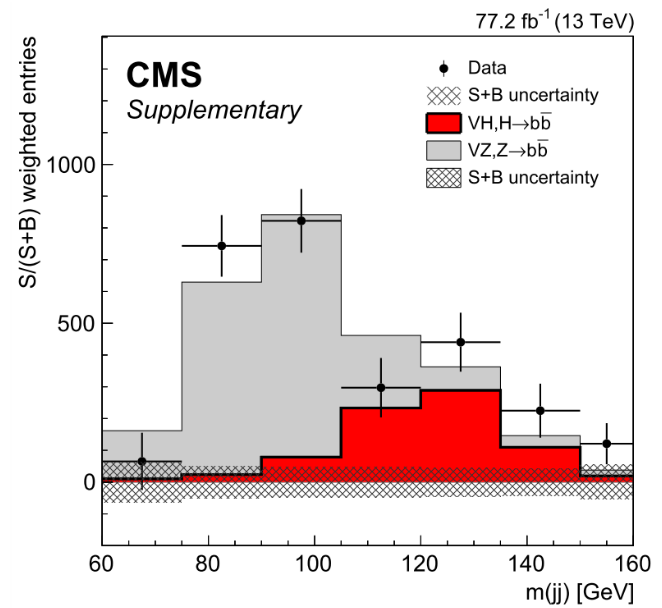
Yukawa coupling of second generation leptons and quarks challenging but already promising



adapted from G. Salam, LHCP '18

Conclusions

- **CMS has reached a 5.6σ observation of the H→bb decay, with signal strength μ = 1.04 ± 0.20**
  - Combination of several production channels, dominated by VH(bb)
  - Result contained in **arXiv:1808.08242** and **provisionally accepted for publication in PRL**
    - Thank you to PRL and its referees for their impressive turn-around in reviewing the paper!



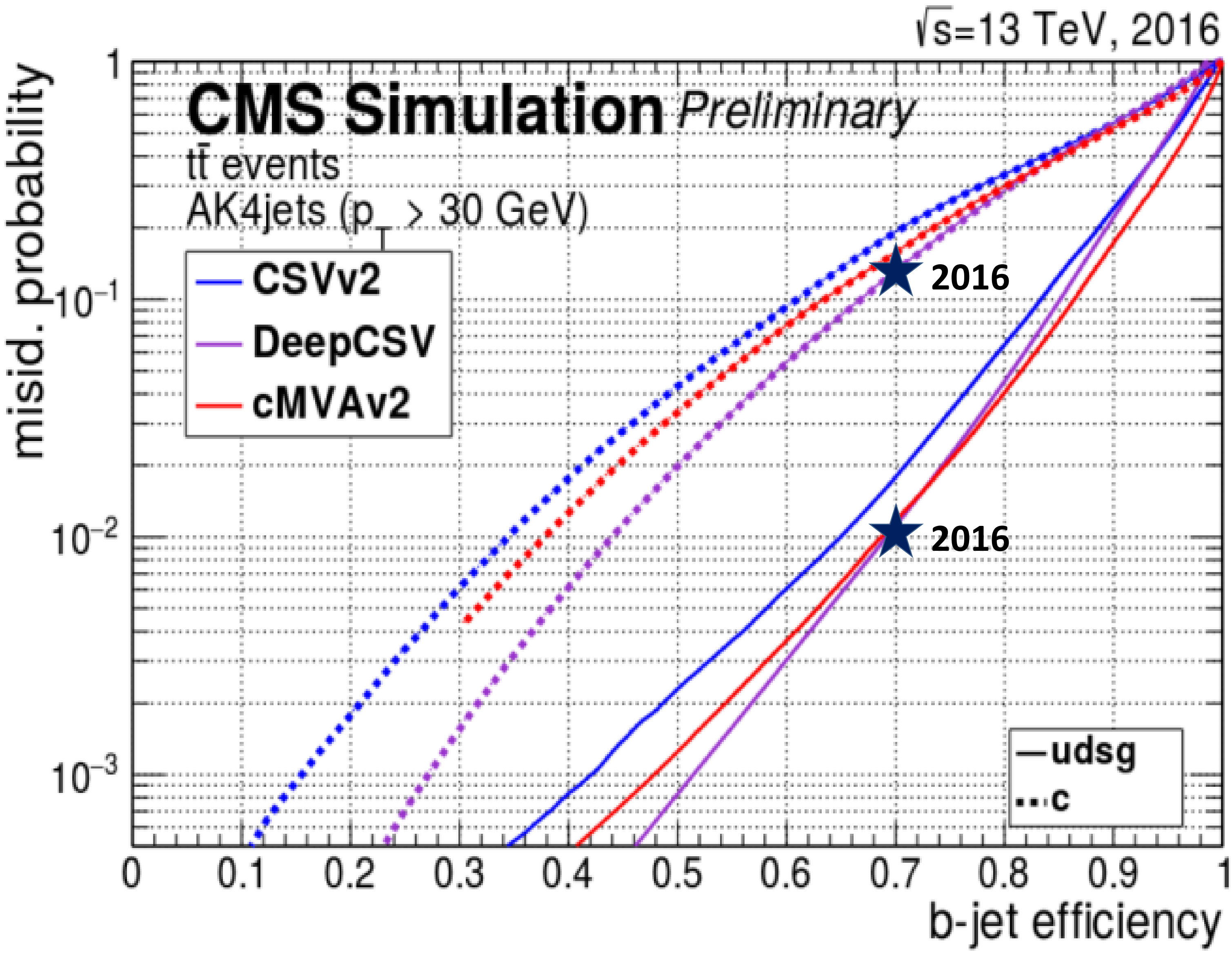
- **Standard Model** assumption on **Yukawa coupling to b's confirmed** within the present uncertainty

- This result is the **culmination of H→bb searches** that started at LEP, continued at Tevatron and at the LHC

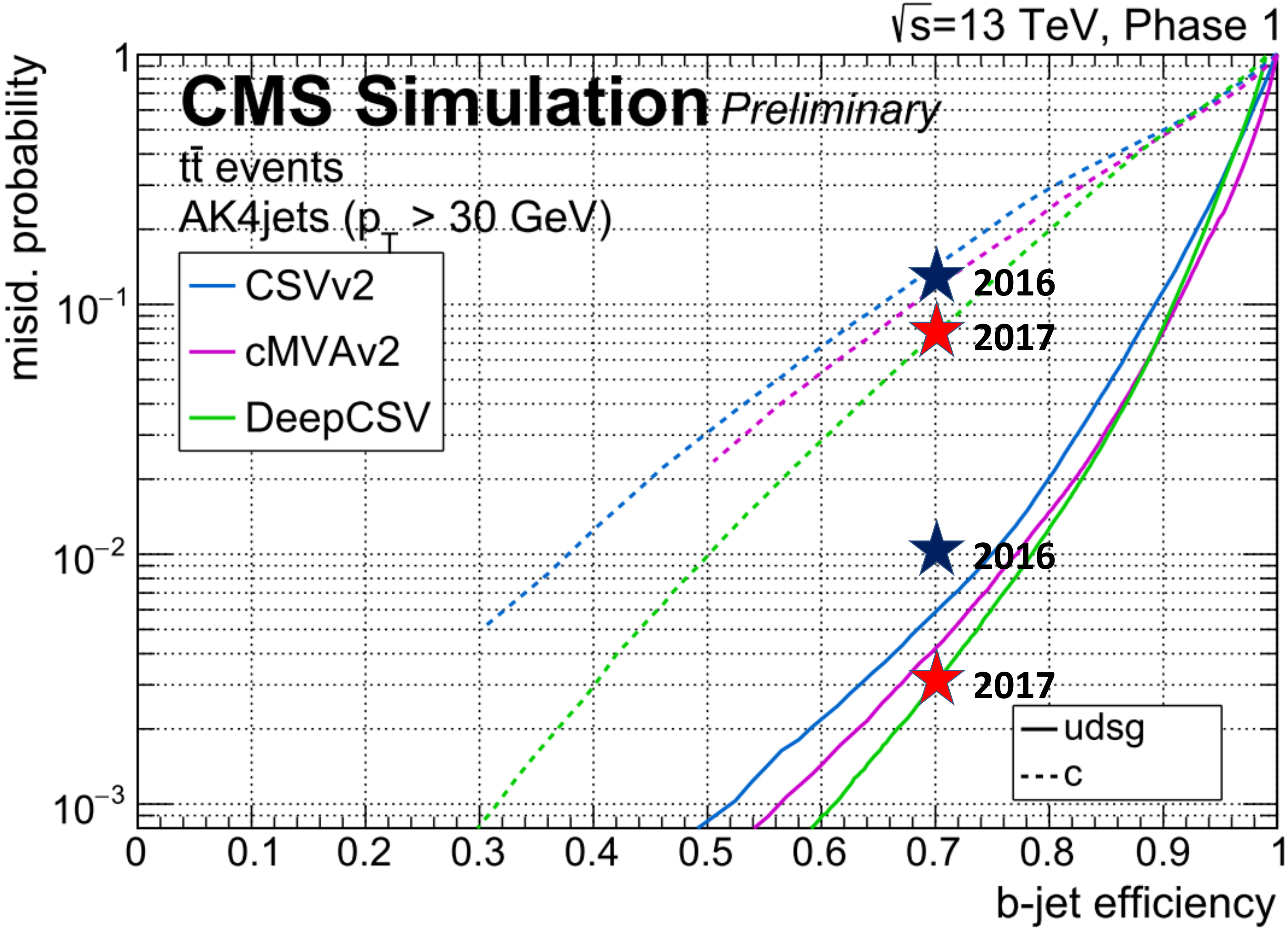
- Achievement possible only thanks to the **fantastic run of the LHC, and the CMS detector performance**
  - But is only a step towards the ultimate H→bb precision at LHC

BACKUP SLIDES

b-tagging performance in 2016



b-tagging performance in 2017



DeepCSV

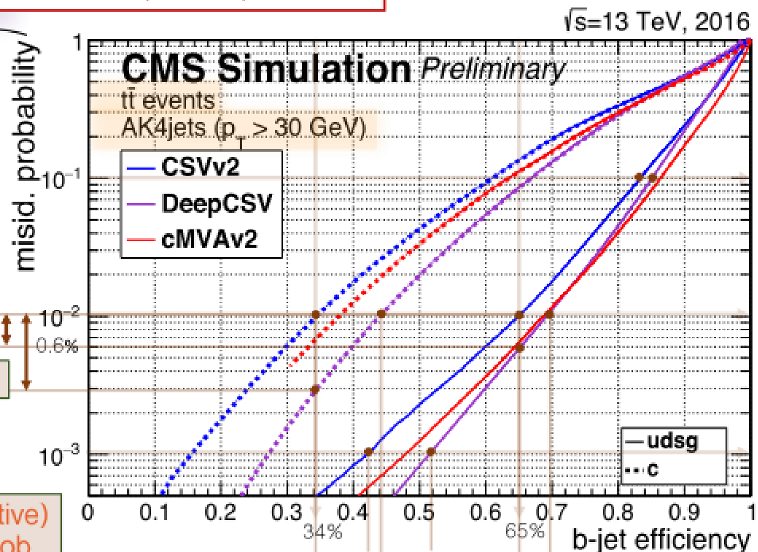
CMS DP 2017-005
cds.cern.ch/record/2255736

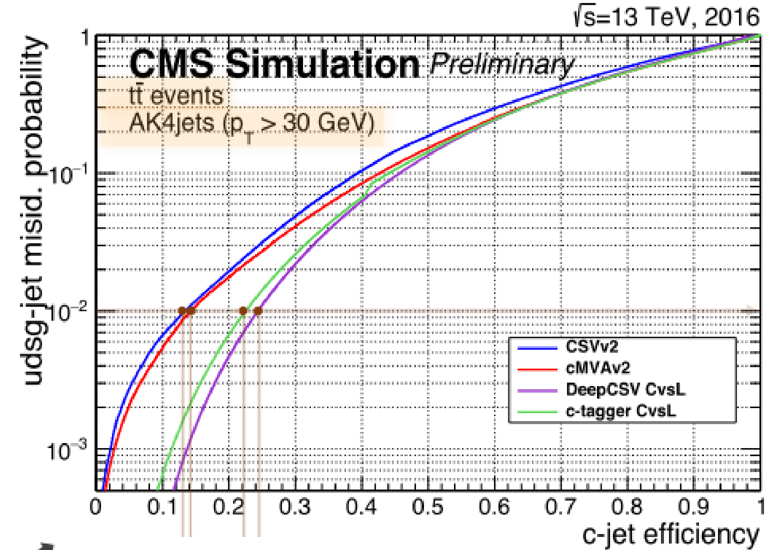
of a light (or c) jet as a b jet

0.4% (40%)

~0.7% (70%)

Absolute (relative) less misid. prob.



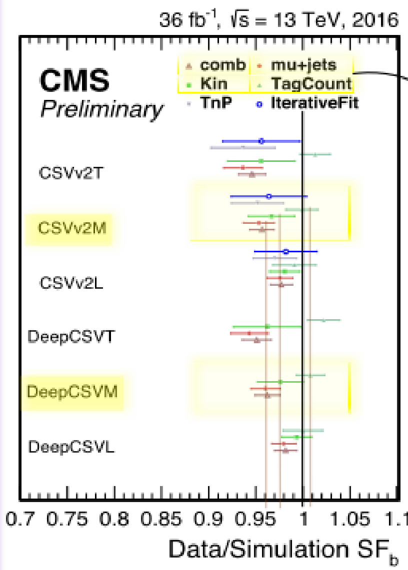


10% (30%)

5% (8%)

Absolute (relative) gain in efficiency

- DeepCSV significantly better than CSVv2 in ~every b jet efficiency value, vs both light and c jet misid. (as b jet) prob.
- DeepCSV efficiency equal or better than (i) cMVAv2 and (ii) ctagger (in most cases), despite not containing lepton info. (ii) also shows DeepCSV as a multi-classifier



Methods for the b tagging efficiency measurement

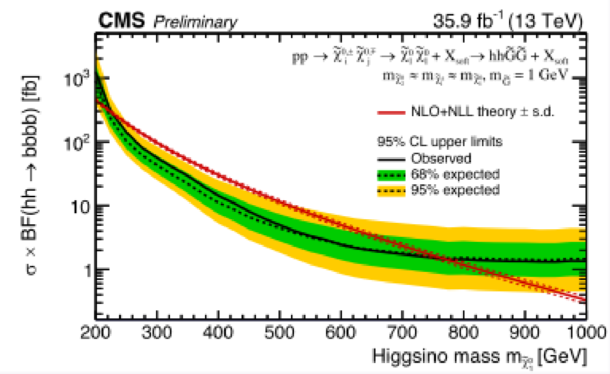
- Similar Scale Factors between CSVv2 and DeepCSV ⇒ better performance for DeepCSV in data too

- DeepCSV already been applied and improved sensitivity in the analysis on the search for:

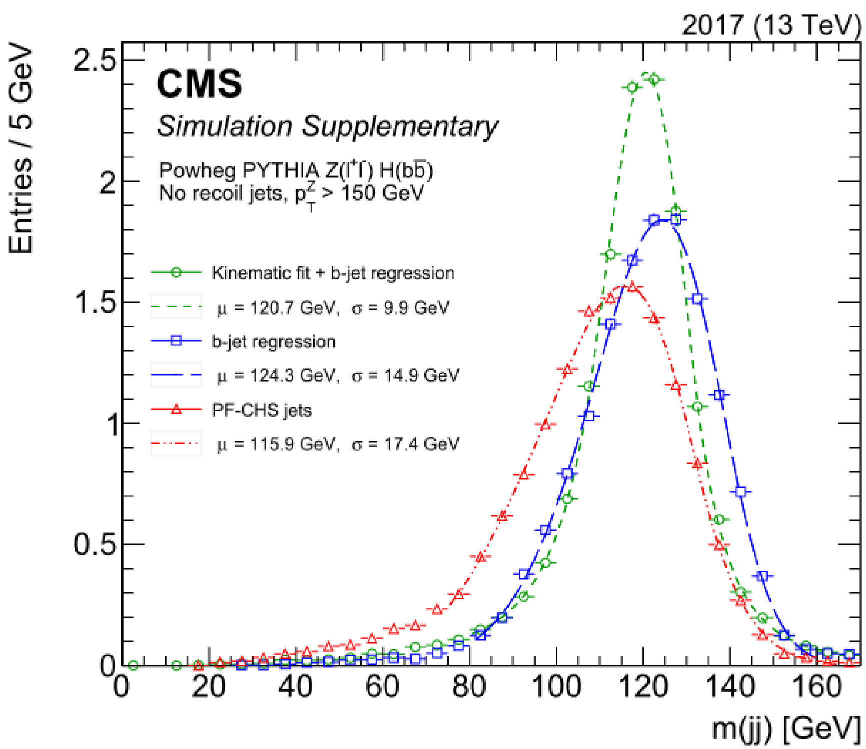
$\widetilde{H} \to h\,(\to b\bar{b}) + |\vec{p}_T^{\,miss}|$,

4 b jets in the final state

CMS PAS SUS-16-044
cds.cern.ch/record/2256648



b-jet energy regression



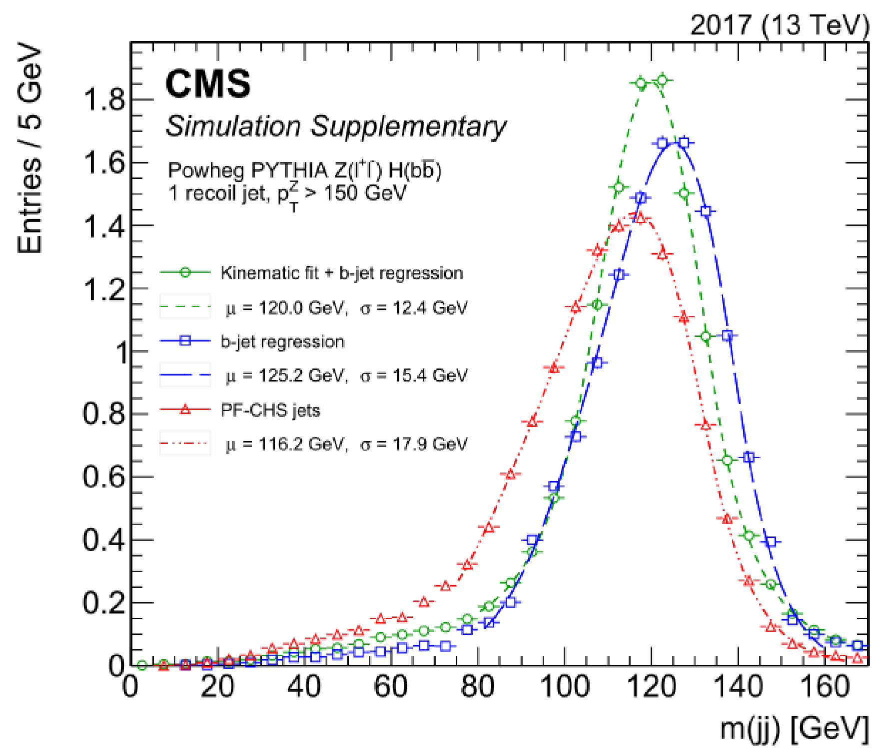


Figure 1: Dijet invariant mass distributions for simulated samples of $Z(\ell\ell)H(bb)$ events ($m_H = 125\,\text{GeV}$) without (left) and with one additional recoiling jet (right). Distributions are shown before (red) and after (blue) the energy corrections from the b-jet regression are applied, and when a kinematic fit procedure (green) is used on top of them. A Bukin function is used to fit the distribution. The fitted mean and width of the core of the distribution are displayed on the figure.

b-jet energy regression

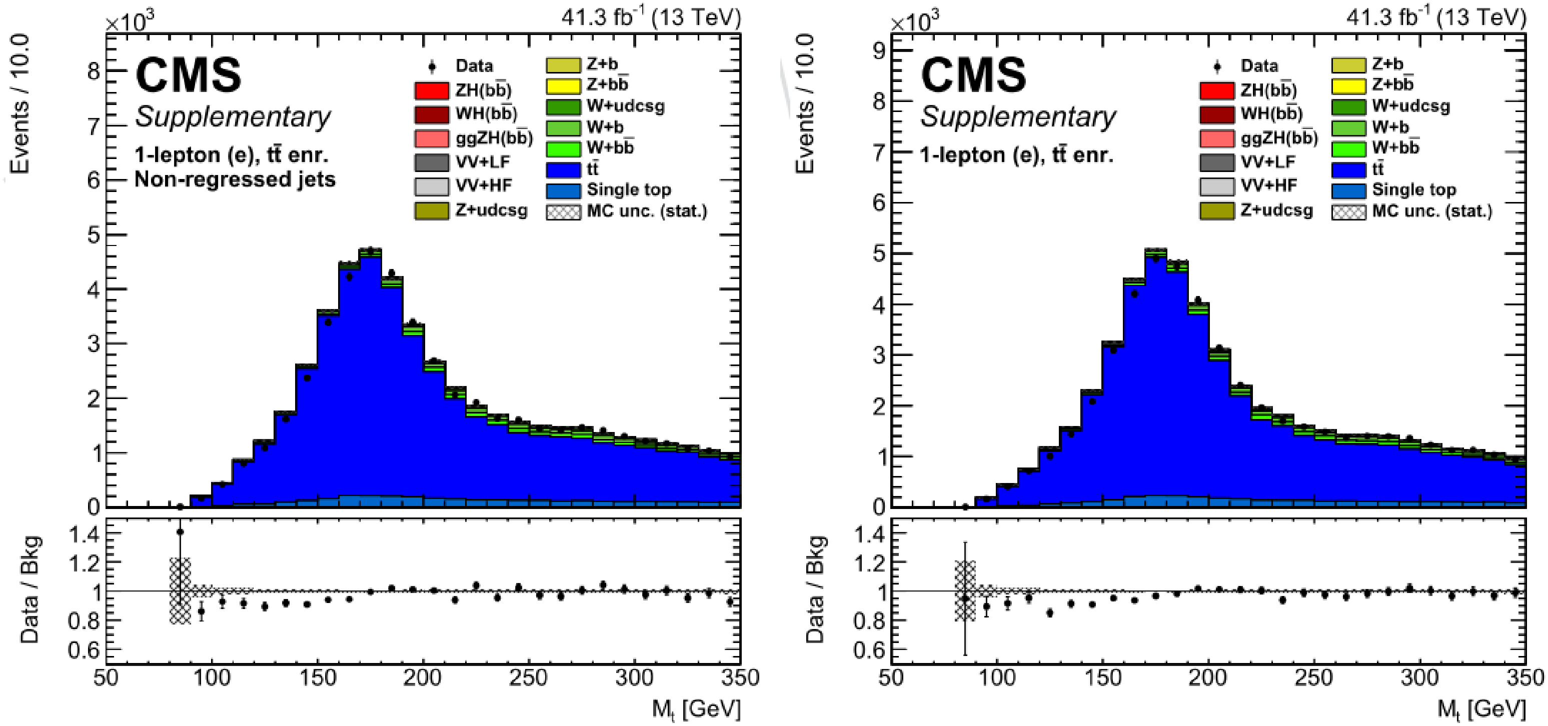


Figure 9: The two plots above show the top quark mass reconstructed in the 1-lepton $t\bar{t}$ control region using the tagged lepton, p_T^{miss}, one of the two b-jets and the constraint of the W mass to estimate the longitudinal component of the neutrino. The reconstruction on the left uses the un-regressed b-jet energy and the right uses regressed b-jet energy.

Kinematic fit in 2-lepton channel

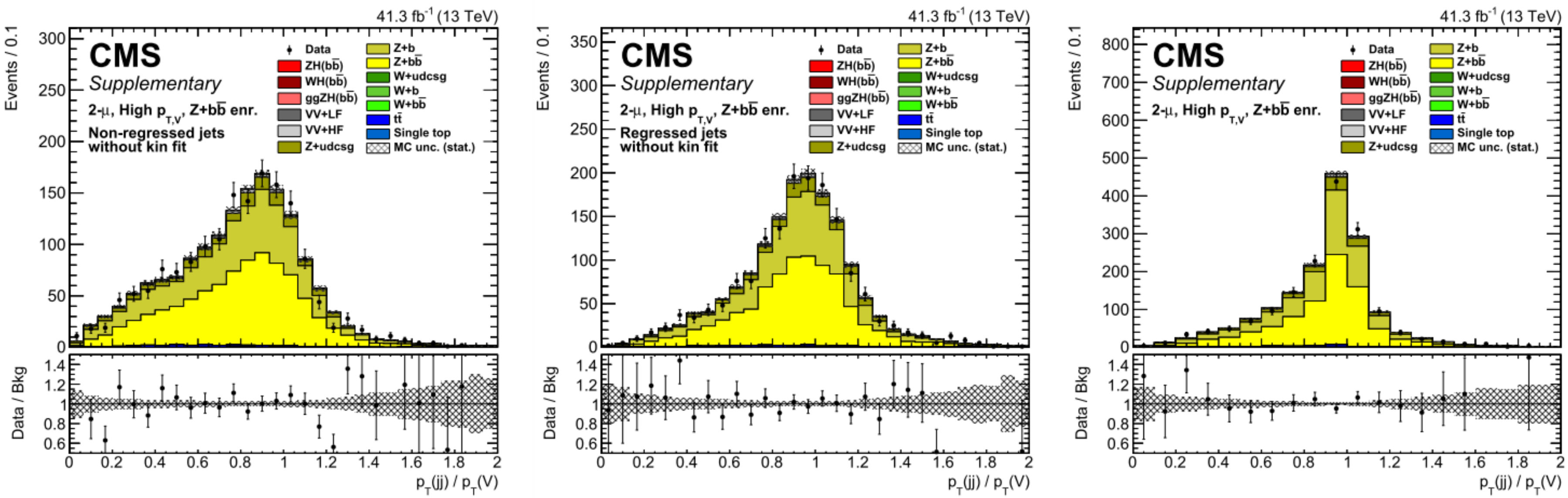


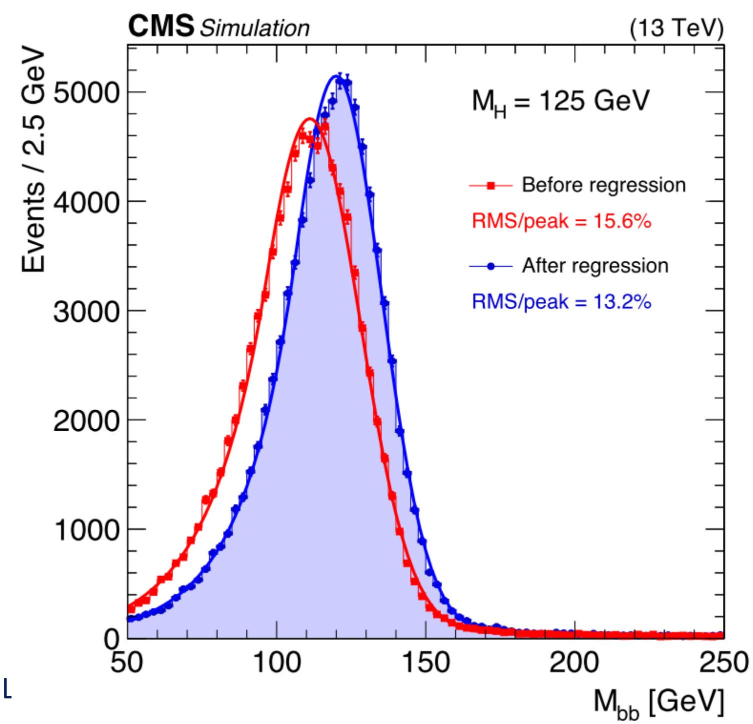
Figure 11: All three figures above show the ratio of the di-jet p_T to the di-lepton (V) p_T in the 2-lepton HF control region. The b-jets in the left plot come directly from CMS reconstruction with charged hadron subtraction applied. The b-jets in the center plot have been updated by the regression. The resolution is visible improved from left to center. On the right the b-jet energies are updated once again with a kinematic fit which constrains the b-jet energies using the lepton resolution. Again there is a visible improvement in b-jet resolution inferred by the narrowing balance of the di-jet plus di-lepton system.

m(bb) improvements: 2016 vs 2017

Table 1: Higgs boson invariant mass resolution before (including regression) and after the kinematic fit in bins of $p_T(V)$ and number of ISR jets. These estimates are made with Z bosons plus Higgs boson events where the Z boson decays to two charged leptons. This sample is simulated with POWHEG while showering is simulated with PYTHIA. The resolution listed in the table are the sigma in GeV of a Bukin function fit.

$p_T(V)$	#ISR Jets	σ_{std}	σ_{reg}	σ_{fit}
> 150	0	17.4	14.9	9.9
> 150	1	17.9	15.4	12.4
> 150	> 1	18.9	15.9	14.4

2017



2016

2-lepton heavy flavor control regions

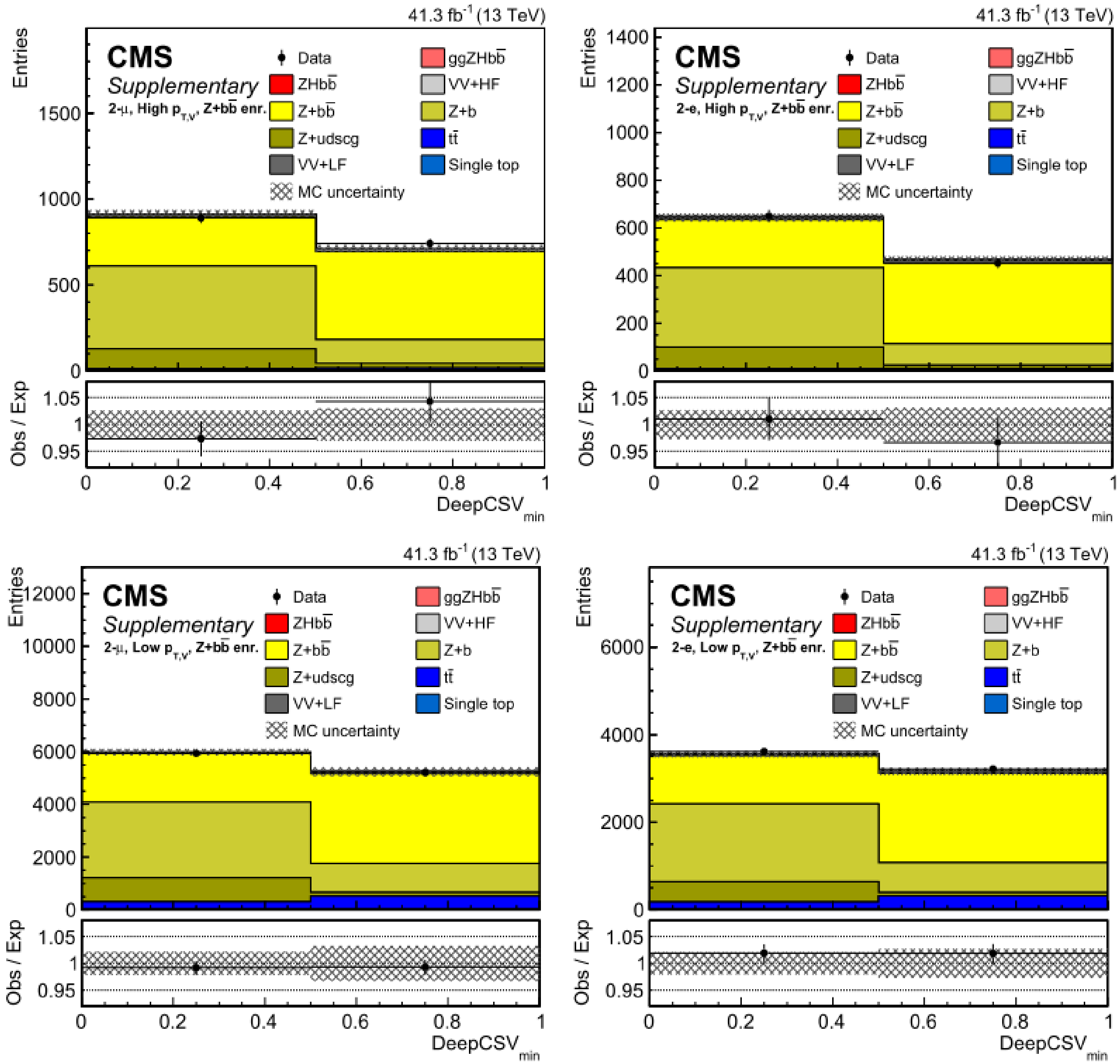


Figure 6: Post-fit distributions of the two fitted bins of $\mathrm{DeepCSV}_{\mathrm{min}}$ in the 2-lepton Z+HF control regions. Above are the high $p_{\mathrm{T}}(\mathrm{V})$ categories and below are the low $p_{\mathrm{T}}(\mathrm{V})$. The left shows the distributions for 2μ channels and on the right are the same for 2e channels.

0- and 1-lepton heavy flavor control regions

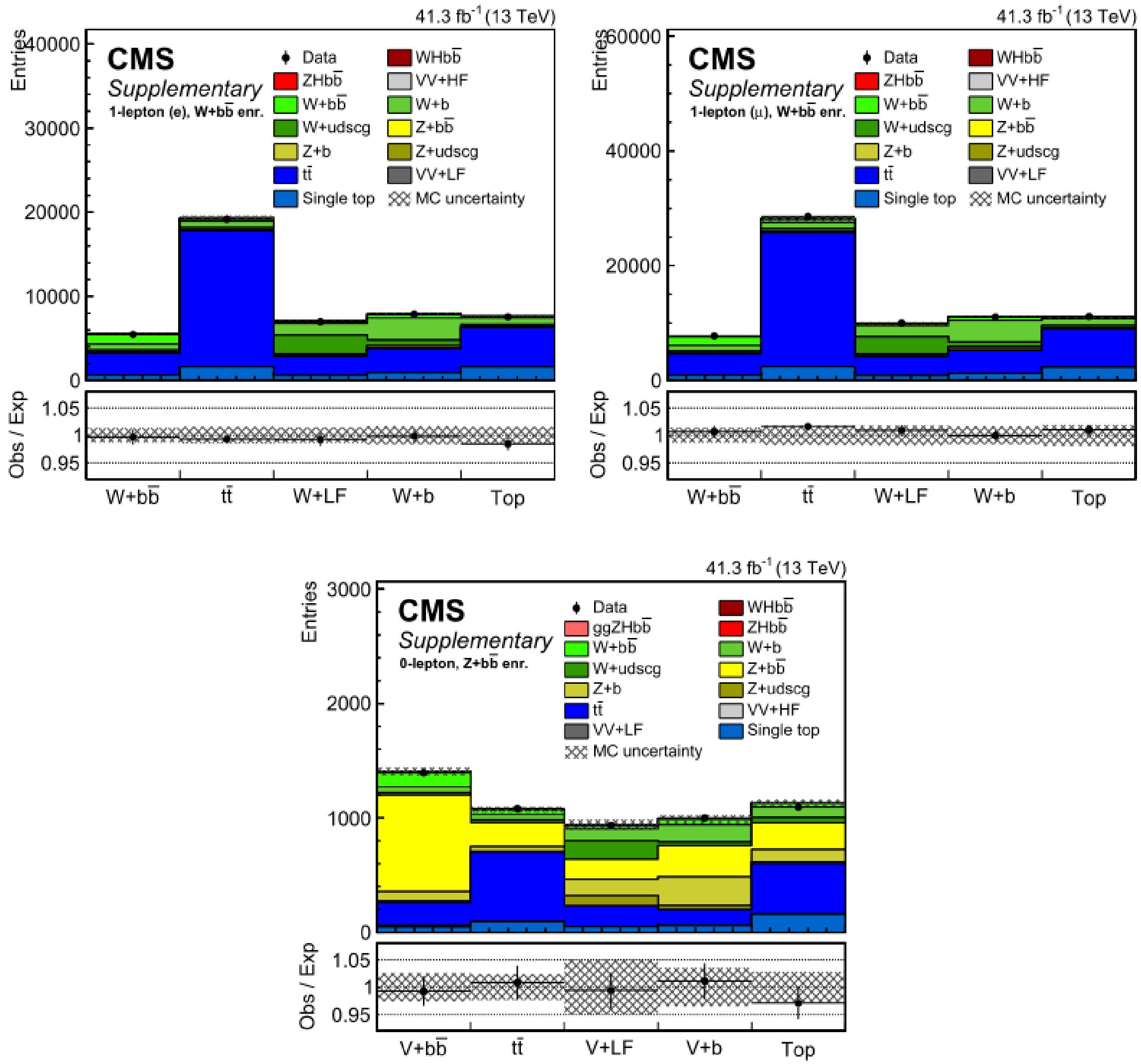
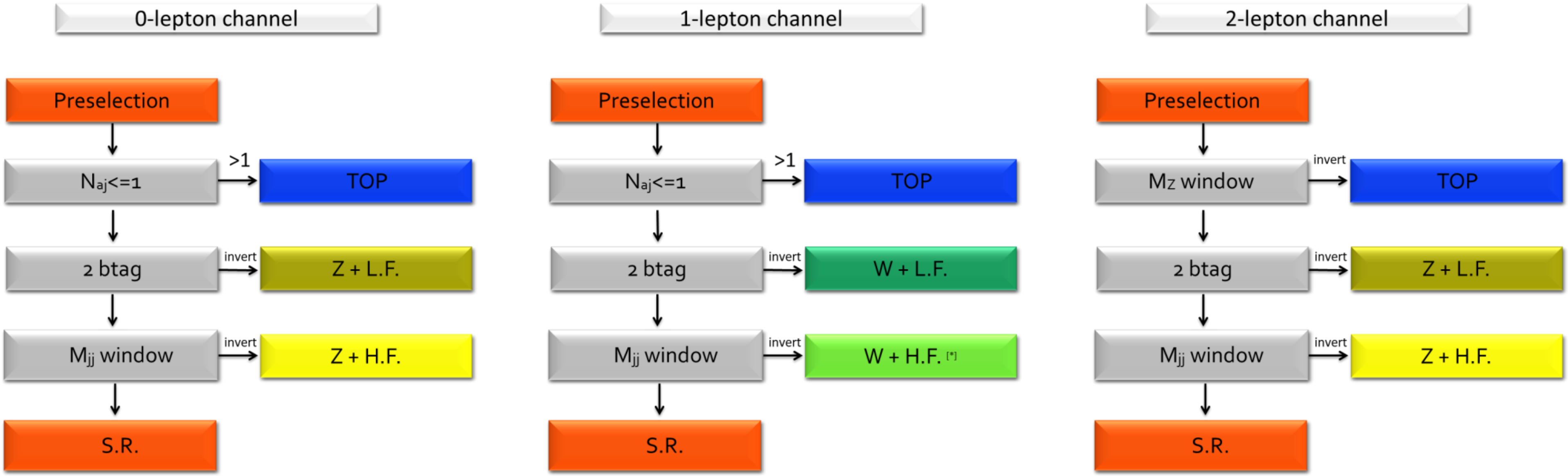


Figure 5: Post-fit distributions of the Multi-background DNN fit variable for 2017 analysis in the 1-lepton channel (top row) for muon (left) and electron (right) control regions, and for the 0-lepton channel (bottom row).

Signal and Control region definitions

Table 14: Signal region pre-selection cuts for each channel. The values listed for kinematical variables are in units of GeV.

Variable	$Z(\nu\nu)$H	$W(\ell\nu)$H	$Z(\ell\ell)$H
$p_T(V)$	> 170	> 150 (**)	$[50-150], > 150$
$m_{\ell\ell}$	–	–	$[75-105]$
p_T^{ℓ}	–	$(> 25, > 30)$	> 20
$p_T(j_1)$	> 60	> 25	> 20
$p_T(j_2)$	> 35	> 25	> 20
$p_T(jj)$	> 120	> 100	–
$M(jj)$	$[60-160]$	$[90-150]$	$[90-150]$
$btag_{max}$	>Tight	>Tight	>Loose
$btag_{min}$	>Loose	>Loose	>Loose
N_{aj}	–	< 2	–
N_{al}	$= 0$	$= 0$	–
E_T^{miss}	> 170	–	–
Anti-QCD	Yes	–	–
$\Delta\phi(V,H)$(rad)	> 2.0	> 2.5	> 2.5
$\Delta\phi$(pfMET, trkMET)(rad)	< 0.5	–	–
$\Delta\phi$(pfMET, lep)(rad)	–	< 2.0	–
Tightened Lepton Iso.	–	(0.06, 0.06)	–



[*] W+L.F and W+H.F taken from 1-lepton analysis

[*] W+H.F splitted in high and low mass

SF comparison: 2016 vs 2017

- Note: change in PDF, UE tune, generator versions, b-tagging algorithm, fit binning and 1-lepton pT(V) increase from 100 to 150 GeV between 2016 and 2017: no direct comparison of SF possible

2016

Process	0-lepton	1-lepton	2-lepton low-p_T(V)	2-lepton high-p_T(V)
W0b	1.14 ± 0.07	1.14 ± 0.07	—	—
W1b	1.66 ± 0.12	1.66 ± 0.12	—	—
W2b	1.49 ± 0.12	1.49 ± 0.12	—	—
Z0b	1.03 ± 0.07	—	1.01 ± 0.06	1.02 ± 0.06
Z1b	1.28 ± 0.17	—	0.98 ± 0.06	1.02 ± 0.11
Z2b	1.61 ± 0.10	—	1.09 ± 0.07	1.28 ± 0.09
$t\bar{t}$	0.78 ± 0.05	0.91 ± 0.03	1.00 ± 0.03	1.04 ± 0.05

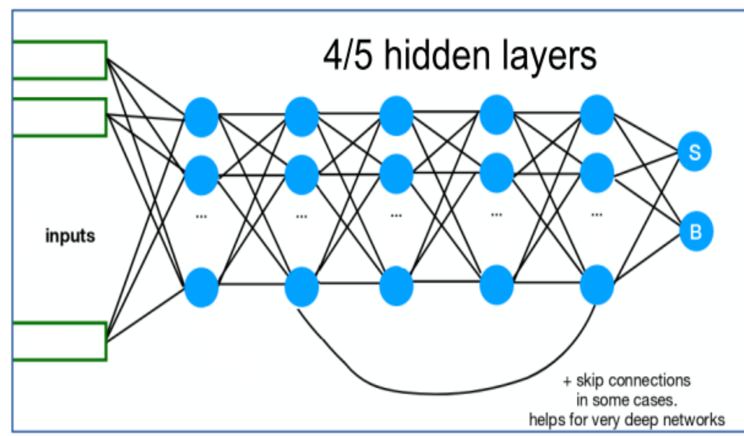
2017

Process	$Z(\nu\nu)H$	$W(\ell\nu)H$	$Z(\ell\ell)H$ low-p_T	$Z(\ell\ell)H$ high-p_T
W + udscg	1.04 ± 0.07	1.04 ± 0.07	–	–
W + b	2.09 ± 0.16	2.09 ± 0.16	–	–
W + $b\bar{b}$	1.74 ± 0.21	1.74 ± 0.21	–	–
Z + udscg	0.95 ± 0.09	–	0.89 ± 0.06	0.81 ± 0.05
Z + b	1.02 ± 0.17	–	0.94 ± 0.12	1.17 ± 0.10
Z + $b\bar{b}$	1.20 ± 0.11	–	0.81 ± 0.07	0.88 ± 0.08
$t\bar{t}$	0.99 ± 0.07	0.93 ± 0.07	0.89 ± 0.07	0.91 ± 0.07

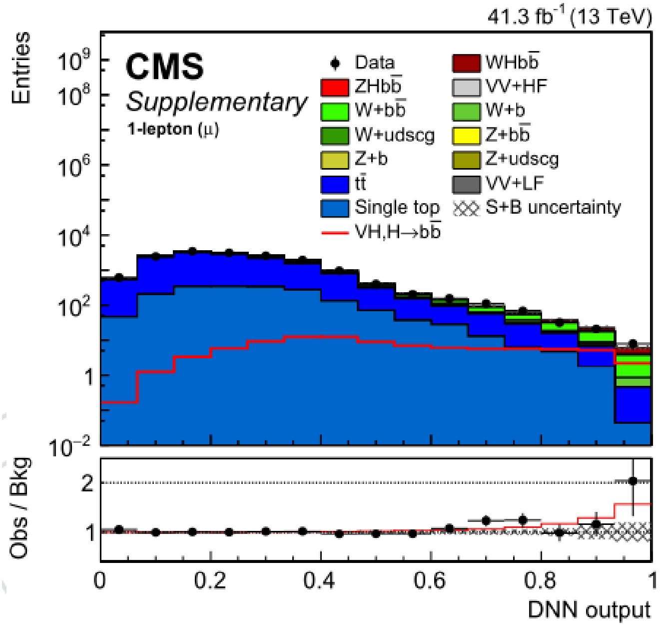
Signal vs Background discriminator

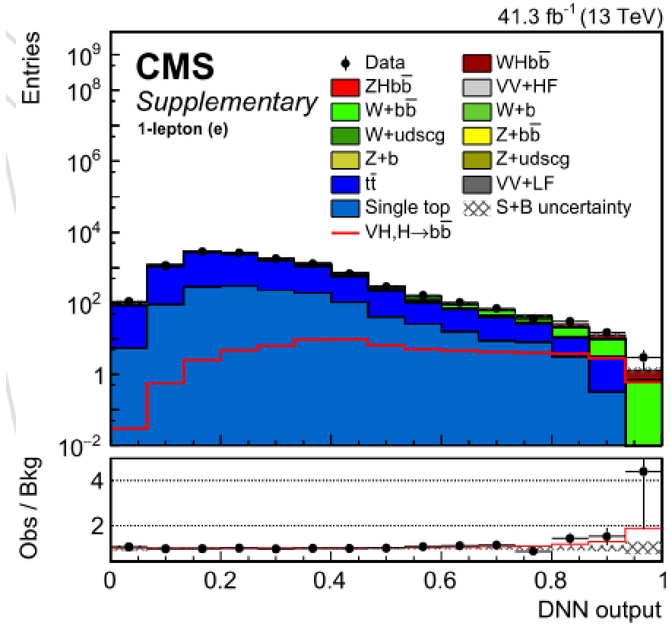
- To increase sensitivity, use **DNN discriminator to extract signal**
  - DNN outperforms BDT due to network depth
    - Same input variables as 2016 (b-jet properties, di-jet kinematics, event topology)
    - Validated through data/MC comparison
  - Trained separately in each channel to discriminate VH($b\bar{b}$) from the weighted sum of all backgrounds
  - Parameters optimized to maximize sensitivity

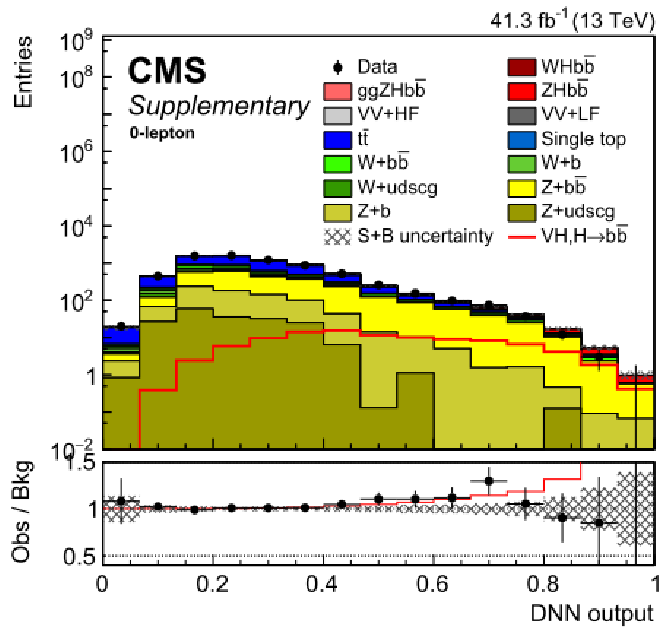
Variable	Description	0-lepton	1-lepton	2-lepton
$M(\mathrm{jj})$	dijet invariant mass	✓	✓	✓
$p_{\mathrm{T}}(\mathrm{jj})$	dijet transverse momentum	✓	✓	✓
$p_{\mathrm{T}}(\mathrm{j}_1)$, $p_{\mathrm{T}}(\mathrm{j}_2)$	transverse momentum of each jet	✓		✓
$\Delta R(\mathrm{jj})$	distance in η–ϕ between jets			✓
$\Delta\eta(\mathrm{jj})$	difference in η between jets	✓		✓
$\Delta\varphi(\mathrm{jj})$	azimuthal angle between jets	✓		
$p_{\mathrm{T}}(\mathrm{V})$	vector boson transverse momentum		✓	✓
$\Delta\phi(\mathrm{V},\mathrm{H})$	azimuthal angle between vector boson and dijet directions	✓	✓	✓
$p_{\mathrm{T}}(\mathrm{jj})\,/\,p_{\mathrm{T}}(\mathrm{V})$	p_{T} ratio between dijet and vector boson			✓
M_{Z}	reconstructed Z boson mass			✓
$btag_{\mathrm{max}}$	value of the b-tagging discriminant (DeepCSV) for the jet with highest score	✓		✓
$btag_{\mathrm{min}}$	value of the b-tagging discriminant (DeepCSV) for the jet with second highest score	✓	✓	✓
$btag_{\mathrm{add}}$	value of b-tagging discriminant for the additional jet with highest value	✓		
$E_{\mathrm{T}}^{\mathrm{miss}}$	missing transverse momentum	✓	✓	✓
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}},\mathrm{j})$	azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and closest jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$	✓		
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}},\ell)$	azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and lepton		✓	
m_{T}	mass of lepton $\vec{p}_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$		✓	
M_{t}	reconstructed top quark mass		✓	
N_{aj}	number of additional jets		✓	✓
$p_{\mathrm{T}}(\mathrm{add})$	transverse momentum of leading additional jet	✓		
SA5	number of soft-track jets with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$	✓	✓	✓



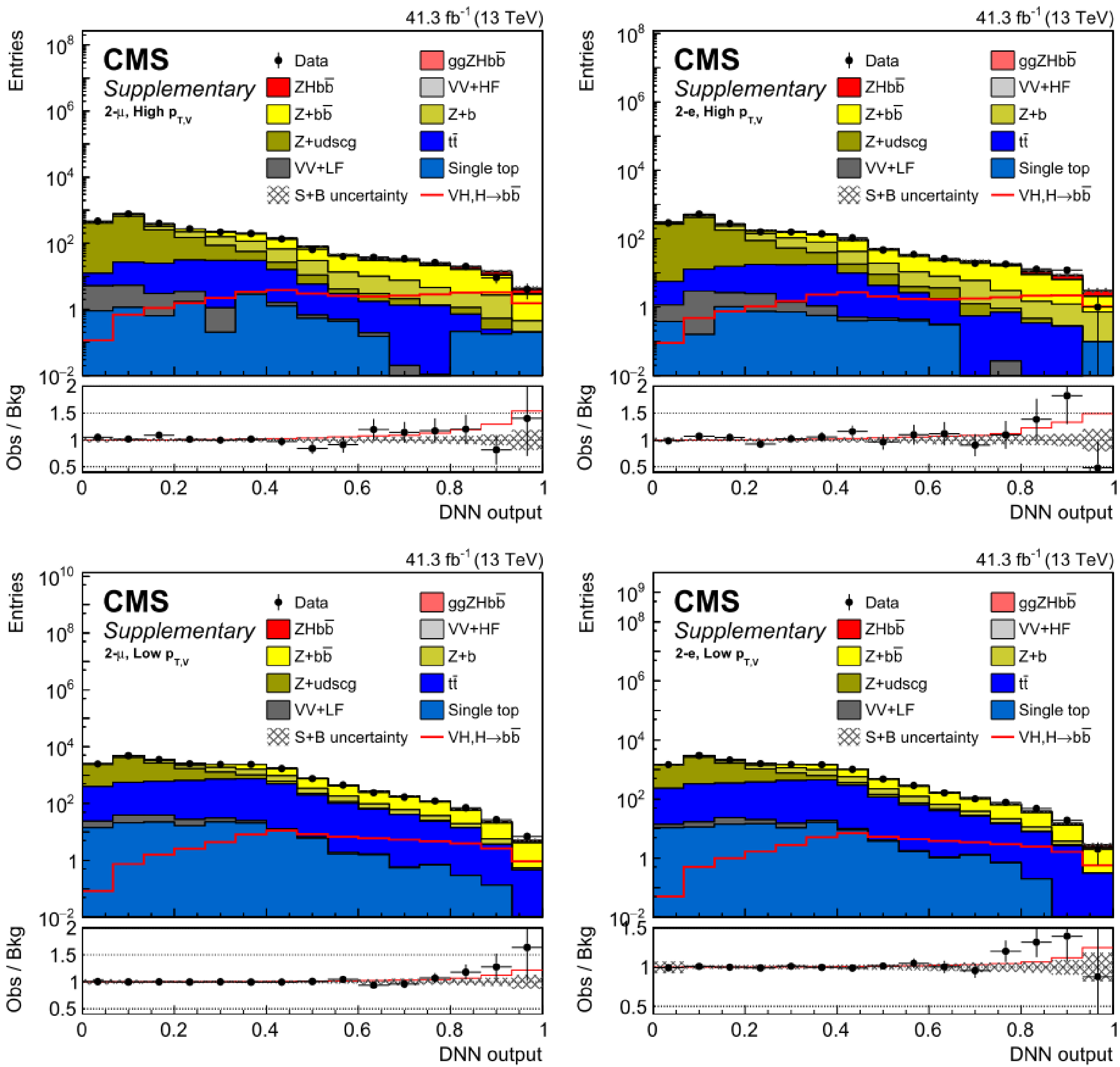
0- and 1-lepton signal regions’ DNN







2-lepton signal regions' DNN



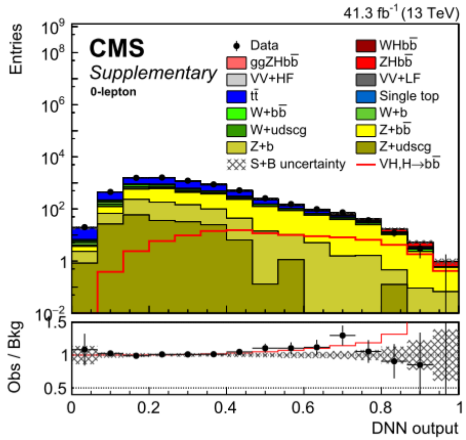
Systematic uncertainties

- **Jet energy scale:**
  - Split into 27 independent uncertainty sources
- **Jet energy resolution:**
  - 10% uncertainty on regressed b-jets from dedicated study
    - Decorrelated for signal to avoid any possible constraining, covers any uncertainties from PS.
  - Standard JER uncertainty for additional jets.
- **B-tagging:**
  - Split into independent uncertainty sources
  - Further de-correlated based on jet pT/η, as in 2016 analysis
- **Background normalizations:**
  - Derived from fit to data for backgrounds with floating normalisation (V+udcsg, V+b, V+bb, tt)
  - 15% uncertainty on VV and single top cross section.
- **Monte Carlo statistics**
- **QCD scales and PDF variations**
  - Acceptance as well as overall cross section
- **Lepton efficiency, pile-up re-weighting, luminosity**
- **Residual data/MC discrepancies**
  - Δη(jj) LO to NLO re-weighting in V+jets
    - Full correction taken as uncertainty.
  - p_T(W) linear re-weighting for tt (all channels) and W+jets, single top (1-lepton channel only)
    - Statistical uncertainty band from fit to derive corrections

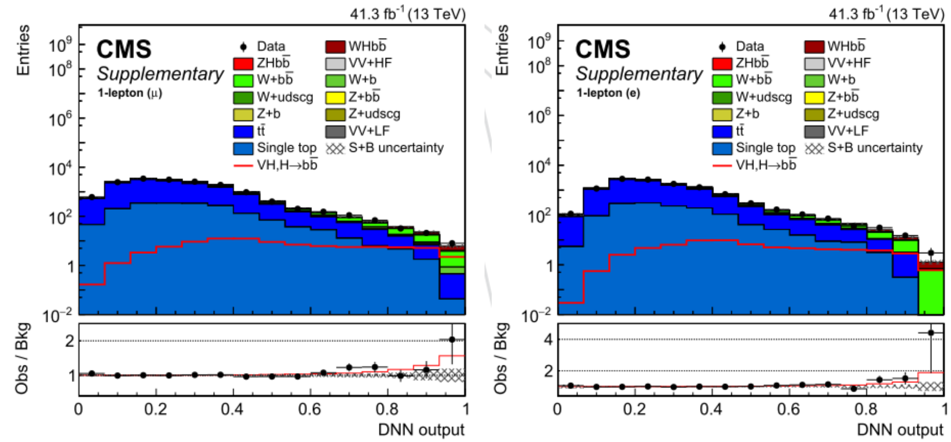
Uncertainty source	$\Delta\mu$	
Statistical	$+0.26$	-0.26
Normalization of backgrounds	$+0.12$	-0.12
Experimental	$+0.16$	-0.15
b-tagging efficiency and misid	$+0.09$	-0.08
V+jets modeling	$+0.08$	-0.07
Jet energy scale and resolution	$+0.05$	-0.05
Lepton identification	$+0.02$	-0.01
Luminosity	$+0.03$	-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

VH(bb) DNN distributions

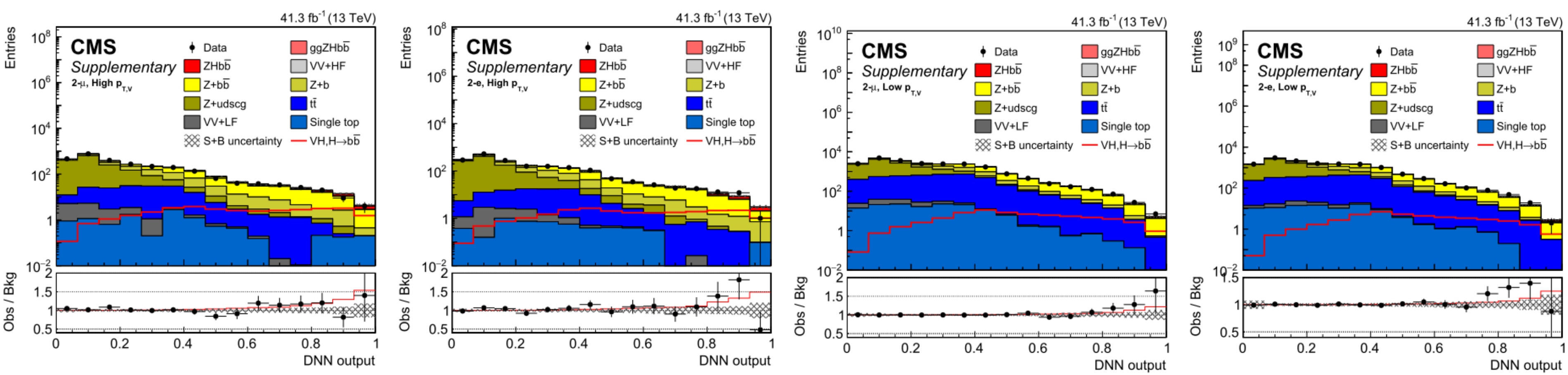
0-lepton



1-lepton

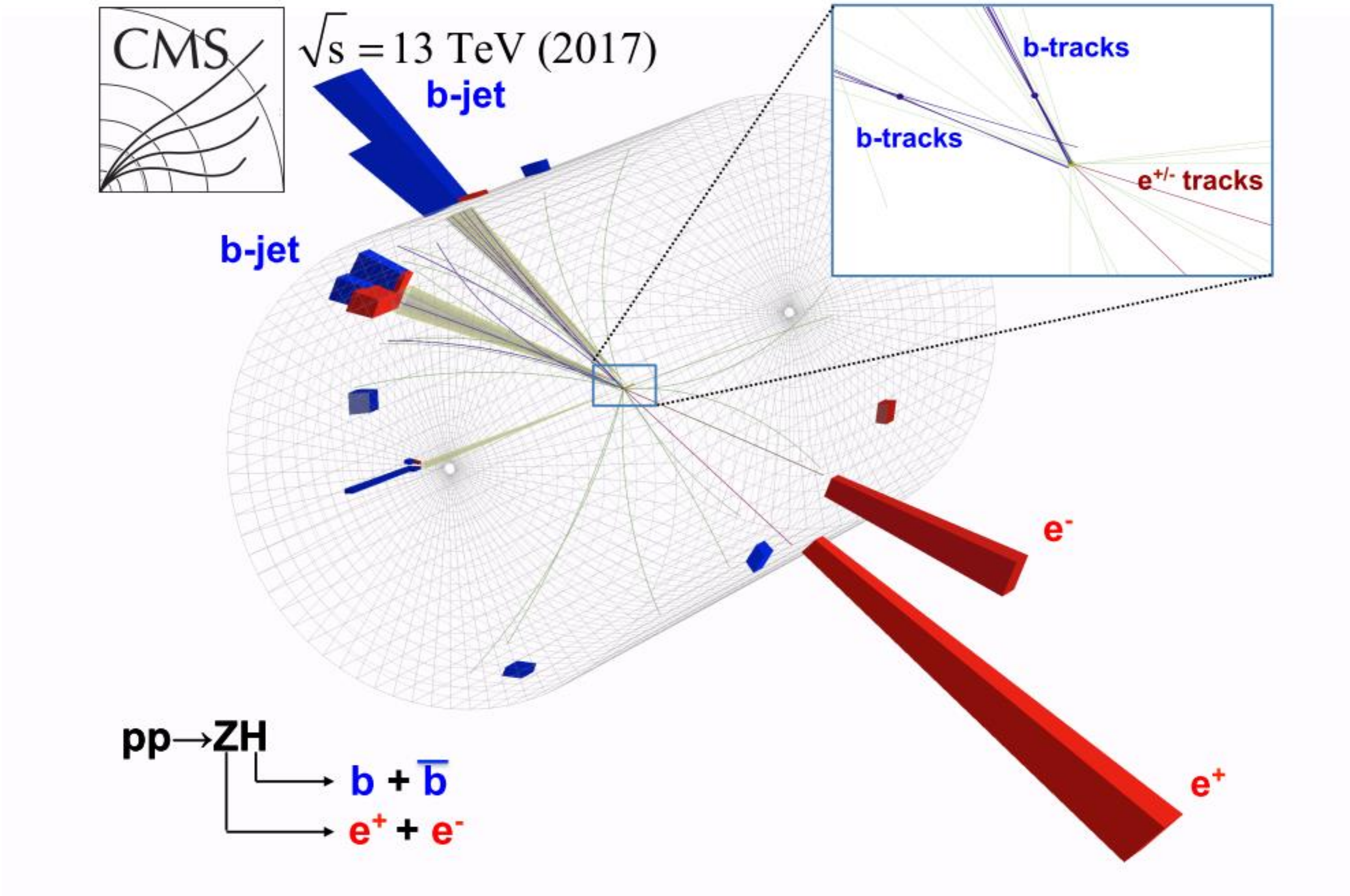


2-lepton



DNN distributions can also be sorted into bins of similar signal-to-background ratio, and combined

Candidate event for Z(ee)H(bb)

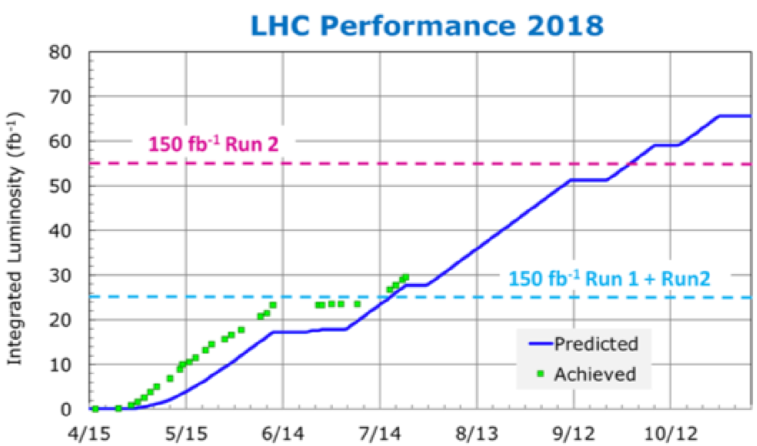


H$\rightarrow$bb combination: syst unc correlations

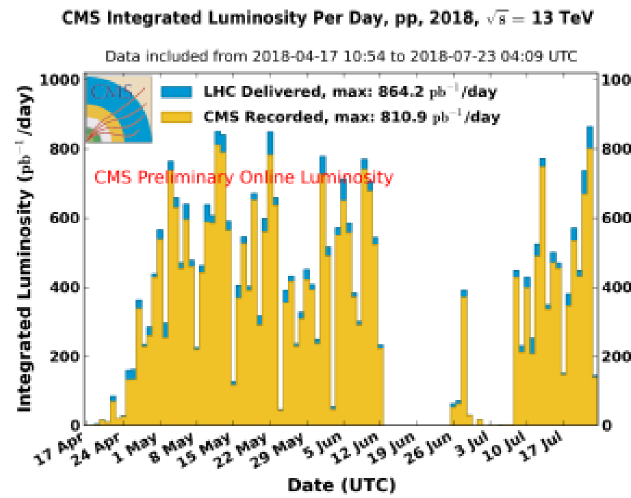
- Combination of published Run1 and Run2 CMS measurements on H$\rightarrow$bb: **VH, boosted ggH, VBF, ttH**
- Most sources of systematic uncertainty are treated as uncorrelated
  - The dominant jet energy scale uncertainties correlated between processes at the same energy
  - Theory uncertainties are correlated between all processes and data sets

Luminosity accumulation in CMS for 2018

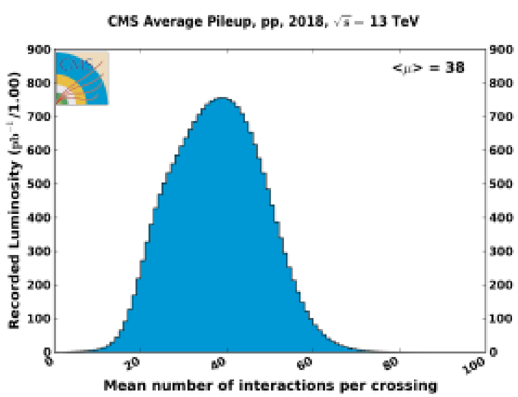
J. Butler - 25th Rencontres du Vietnam ‘18



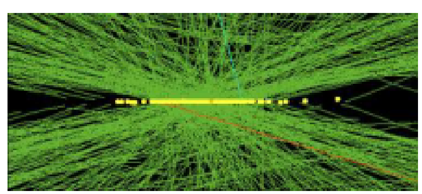
Recording Efficiency 94.3%

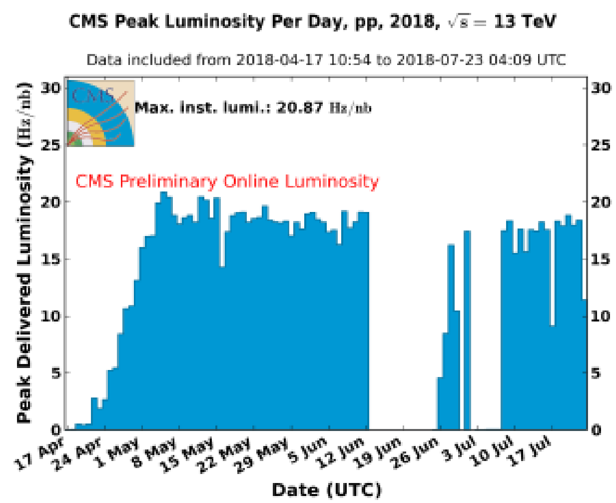


Can get 800 pb^{-1} In a day!



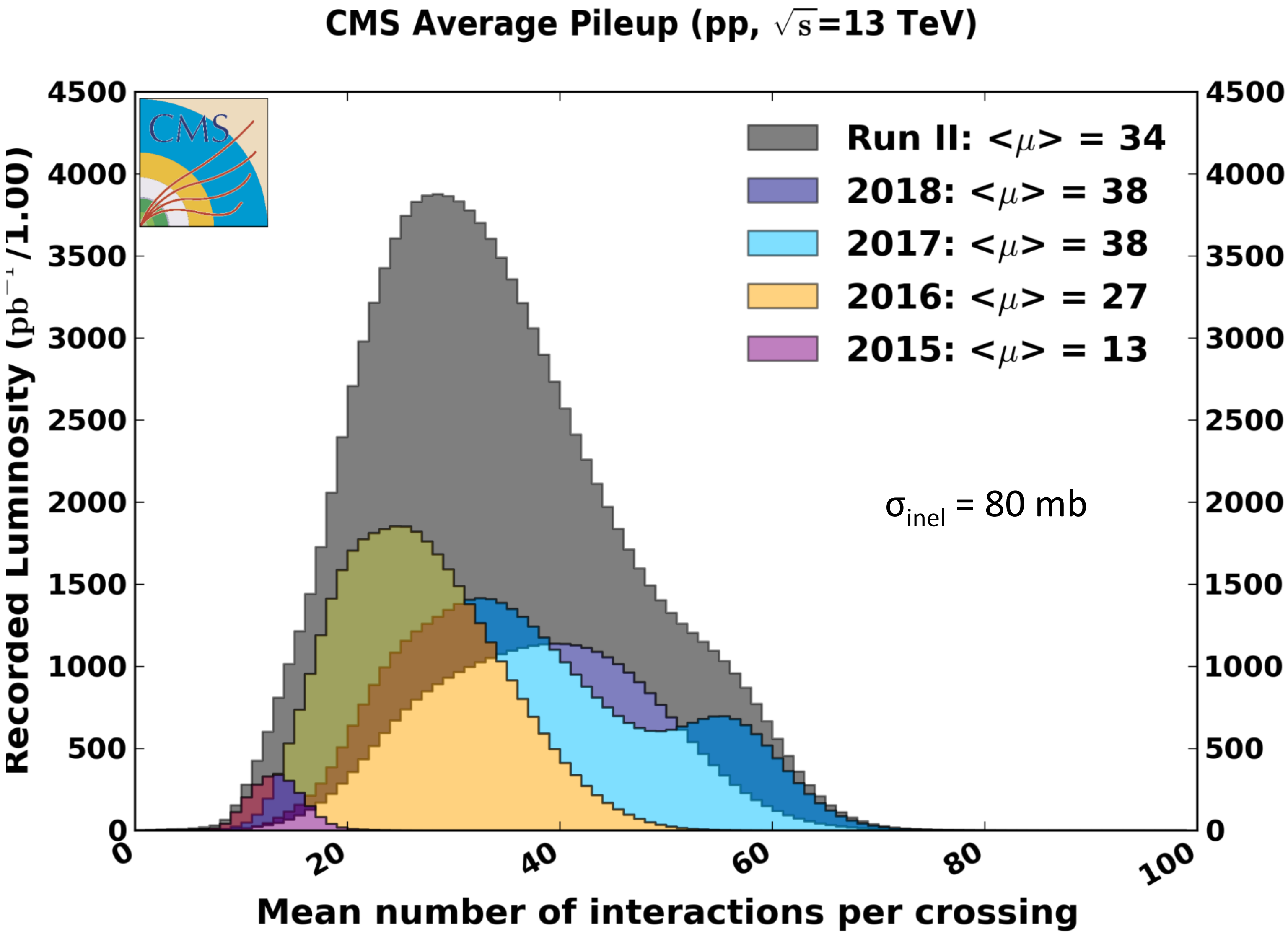
Mean Pileup 38 int/Xing





Peak lumi ~1.8-1.9 x 10^{34} $cm^{-2}s^{-1}$

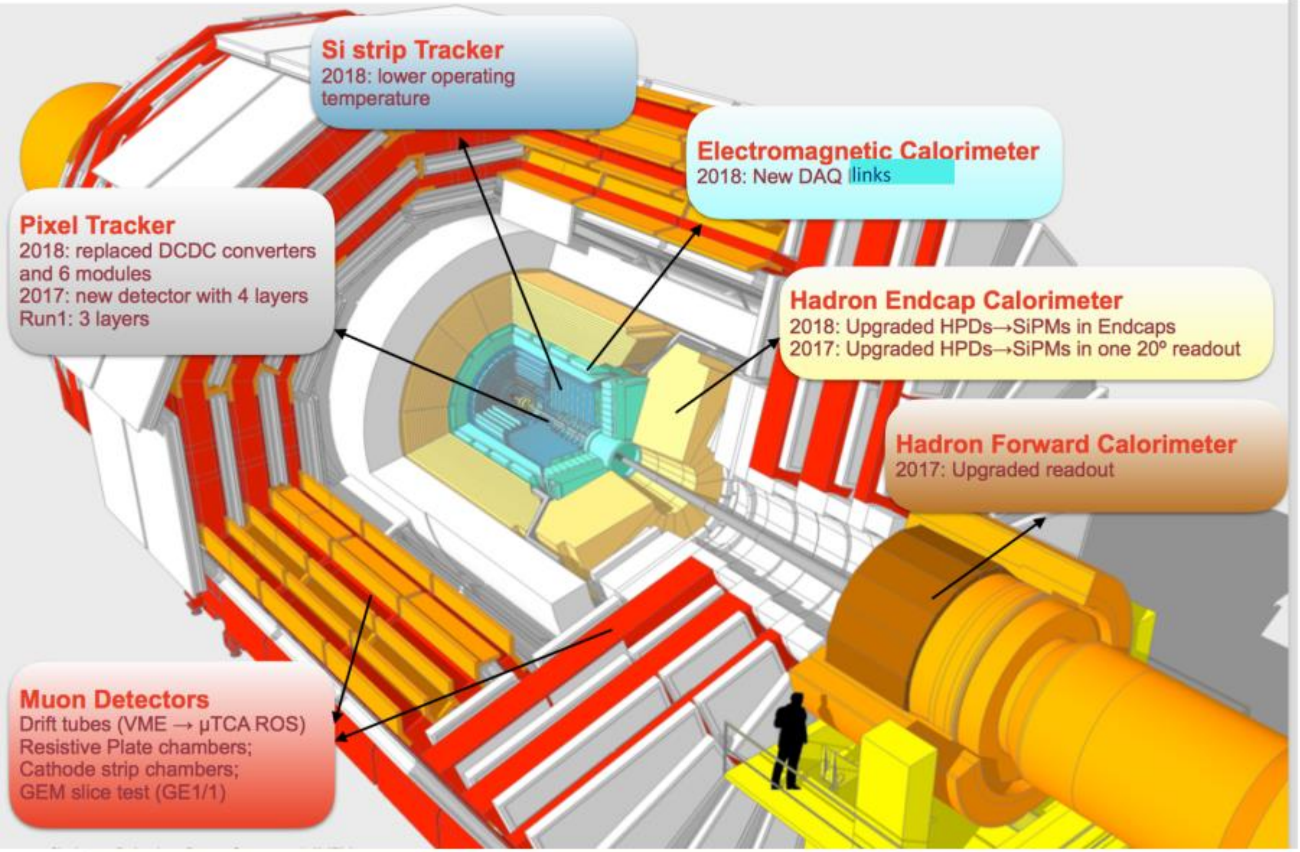
CMS pileup profiles during Run 2



CMS design and 2017/18 Evolution

CMS Design

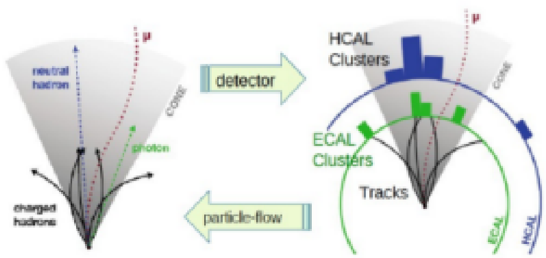
- **Large solenoid** – 6m diameter x 13 m long
  - Tracking and calorimetry fit inside
- **Very strong field** – 3.8T
  - Excellent momentum resolution
- **Chambers in the return iron** track and identify muons, leading to a very compact system
- A **lead tunstenate crystal calorimeter** (~76K crystaks) for photon and electron reconstruction
- **Hadron calorimeters for jet and missing E_T** reconstruction up to η~5
- **Charged Particle Tracking with all-silicon** components
  - A silicon pixel detector out to radius ~20 cm
  - A silicon microstrip detector from there out to 1.1m
- **Weigh, dominated by steel, is 14'000 Tonnes**

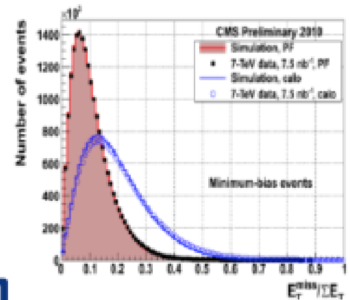


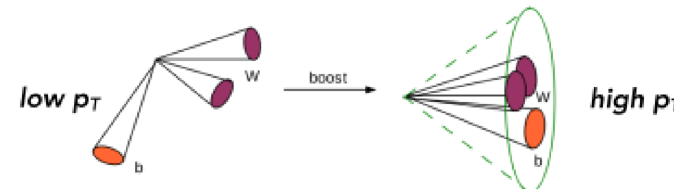
CMS is continuously upgraded to handle higher luminosity and do better physics

Evolution of Analysis Techniques

J. Butler - 25th Rencontres du Vietnam ‘18







Particle Flow

Boosted Jets, Jet Substructure

- Particle Flow uses all available information to reconstruct physics objects, e.g. charged track momenta in jets
  - produces a big improvement in jet energy resolution, tau-lepton identification, and helps with high pileup
- PUPPI (PileUp Per Particle Identification) is a special tool to deal with high pileup
- Use of multivariate analysis techniques to maximize power of available statistics
- Boosted jet topologies and jet substructure analysis
- Use of Deep Neural Nets/Machine Learning

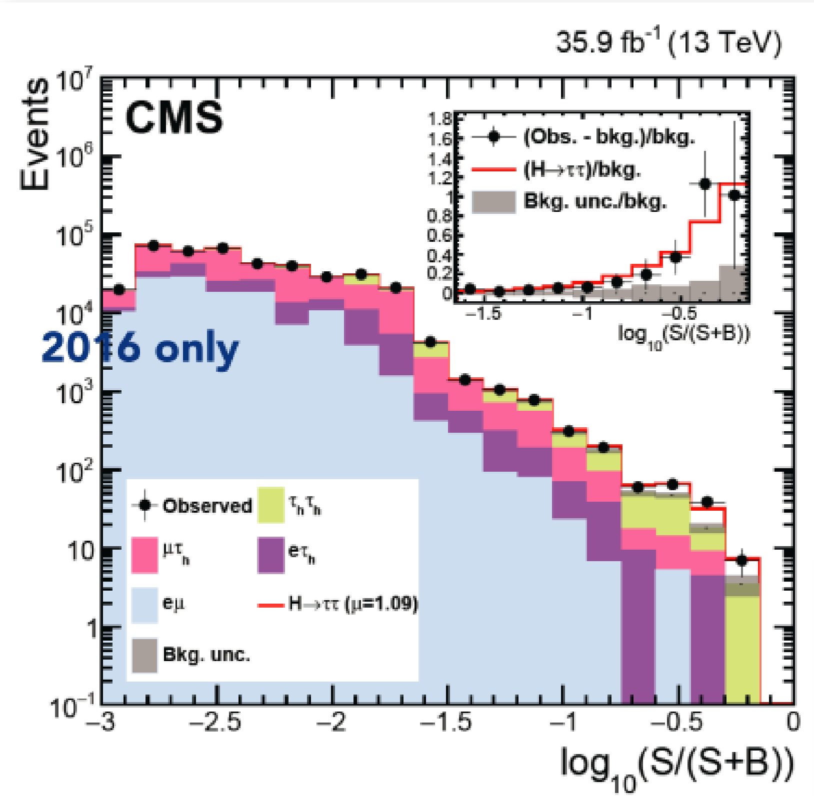
Rapid growth in 2017/18

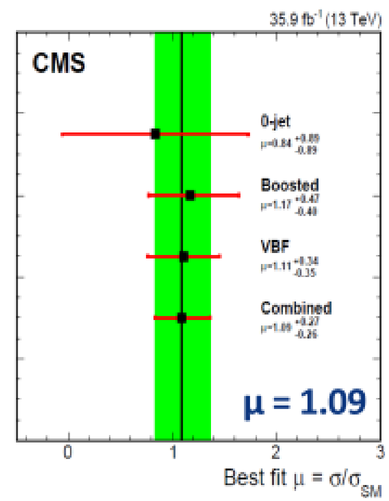
Observation of $H \rightarrow \tau^+\tau^-$ using 7, 8, and 13 (2016 only) TeV data

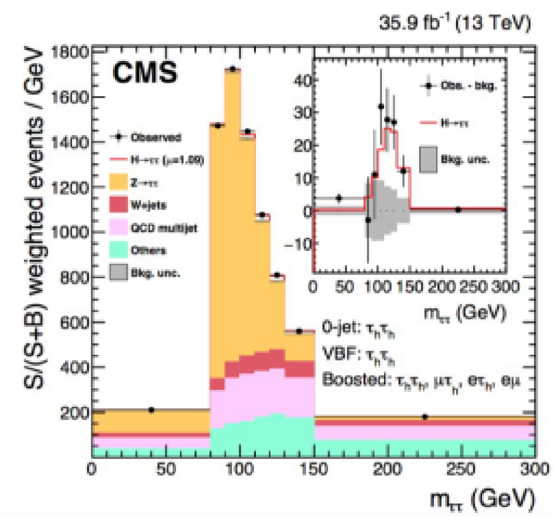
PLB 779 (2018) 283

arXiv:1708.00373

- Branching ratio ~ 6.3%, best channel to establish coupling of Higgs boson to fermions
- Final states: $\tau_h\tau_h$; $e\tau_h$; $\mu\tau_h$; $e\mu$ → Significance of 4.9σ observed (4.7σ expected) with13 TeV data
- **Combination with 7, 8 TeV data: 5.9σ obs. (5.9σ exp.) and μ = 0.98 ± 0.18**







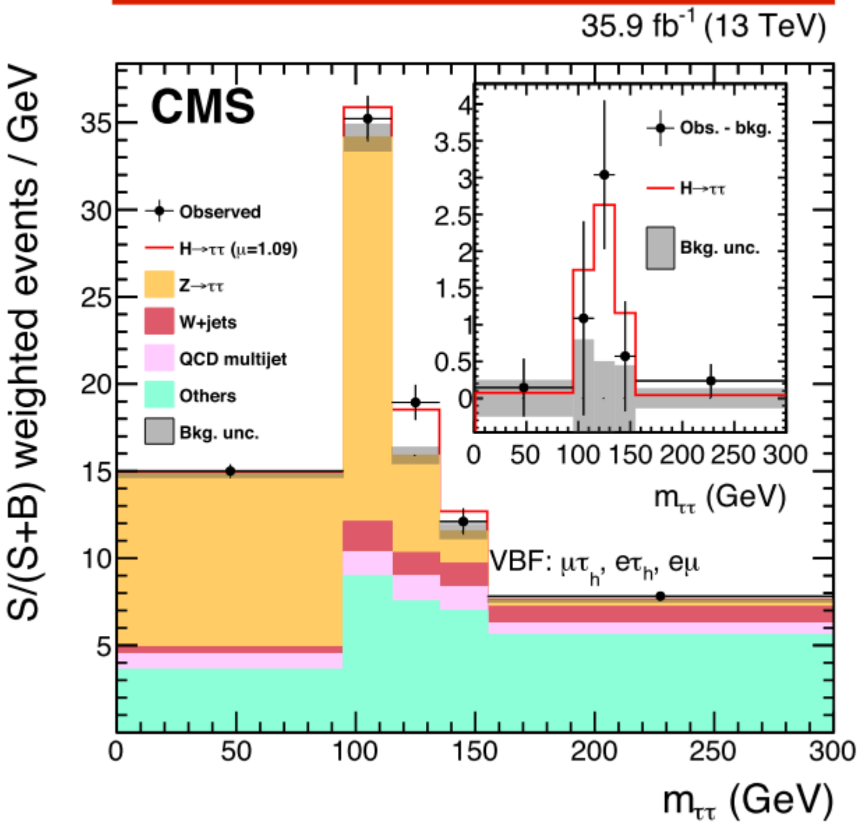
First direct observation by a single experiment of Higgs coupling to fermions!

– **Observed before in CMS+ATLAS combination**

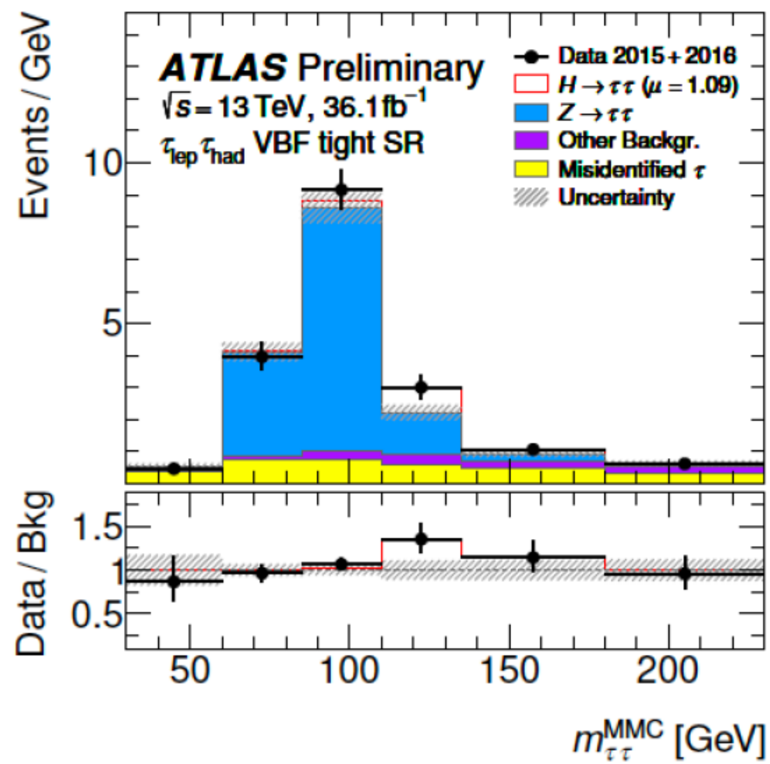
First direct observation of H coupling to leptons and to fermions of the 3rd generation!

the news of the past 12 months

A year ago:
CMS >5-sigma H → τ τ



This week:
ATLAS >5-sigma H → τ τ

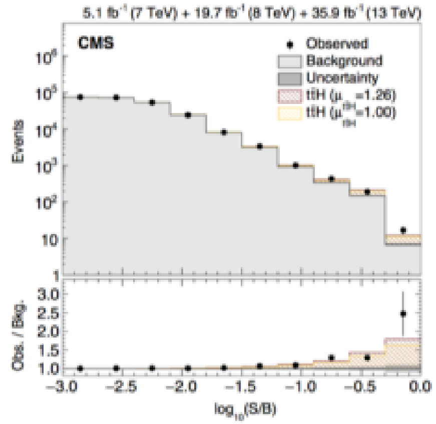


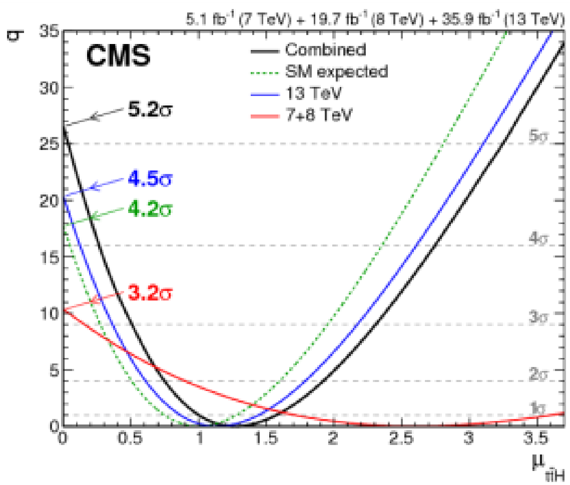
13

G. Salam, LHCP '18

ttH: 7,8, and13 TeV Combined

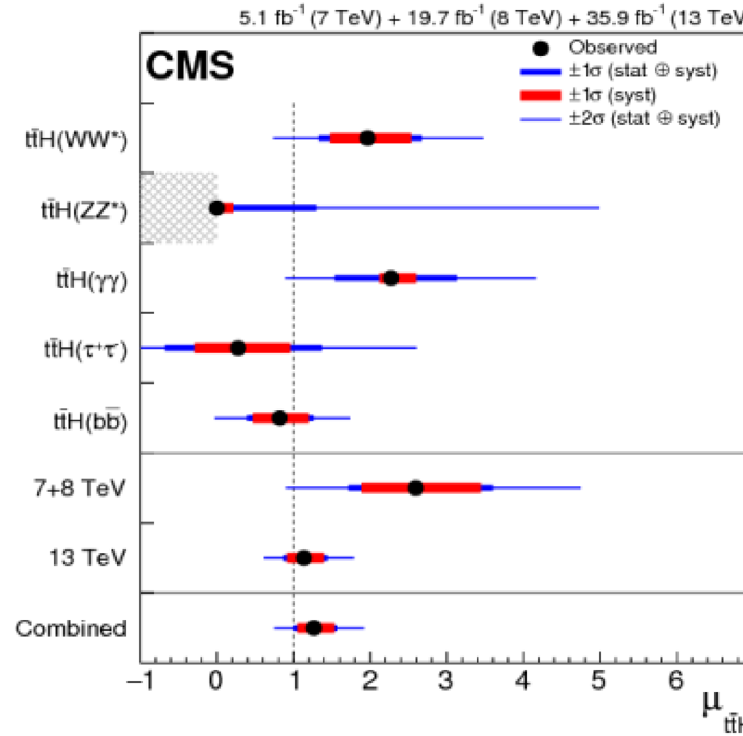
5.1 fb^{-1} (7 TeV)+19.7 fb^{-1} (8 TeV) + 35.9 fb^{-1} (13 TeV)





Test statistic vs coupling strength modifier The horizontal dashed lines indicate the *p*-values for the background-only hypothesis obtained from the asymptotic distribution of *q*,

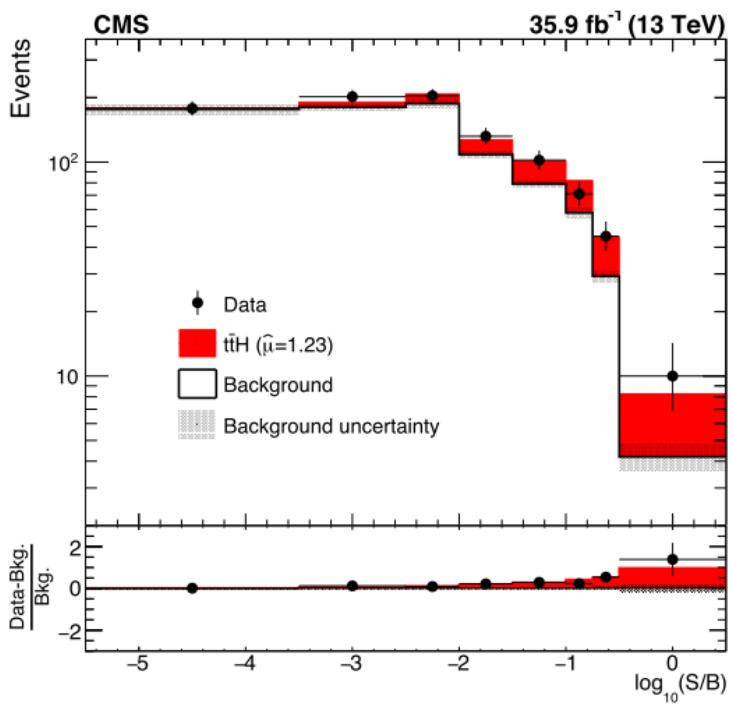
$$\mu = 1.26\,^{+0.31}_{-0.26}$$



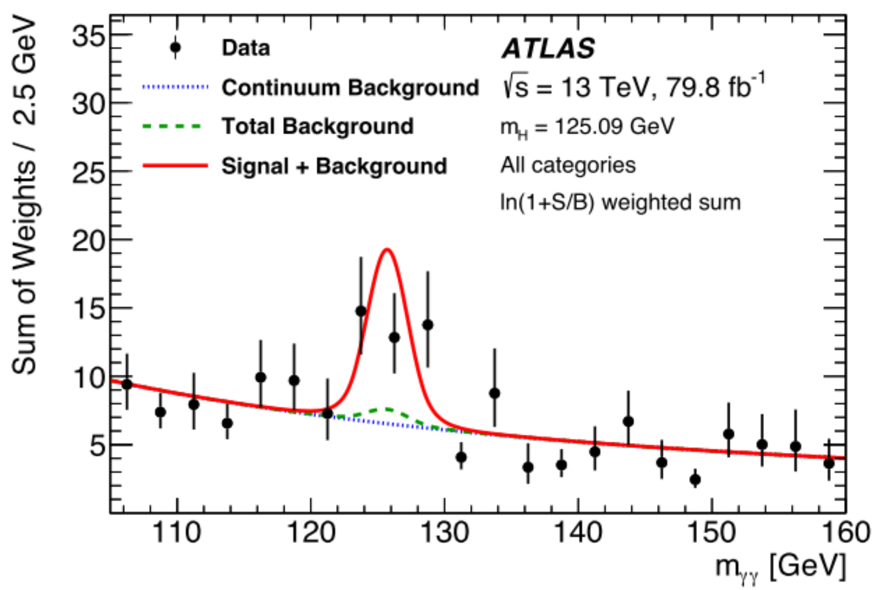
Best fit value of the signal strength modifier for (upper section) the five individual decay channels considered, (middle section) the combined result for 7+8 TeV alone and for 13TeV alone, and (lower section) the overall combined result.

the news of the past 12 months

A few weeks ago: CMS >5-sigma ttH



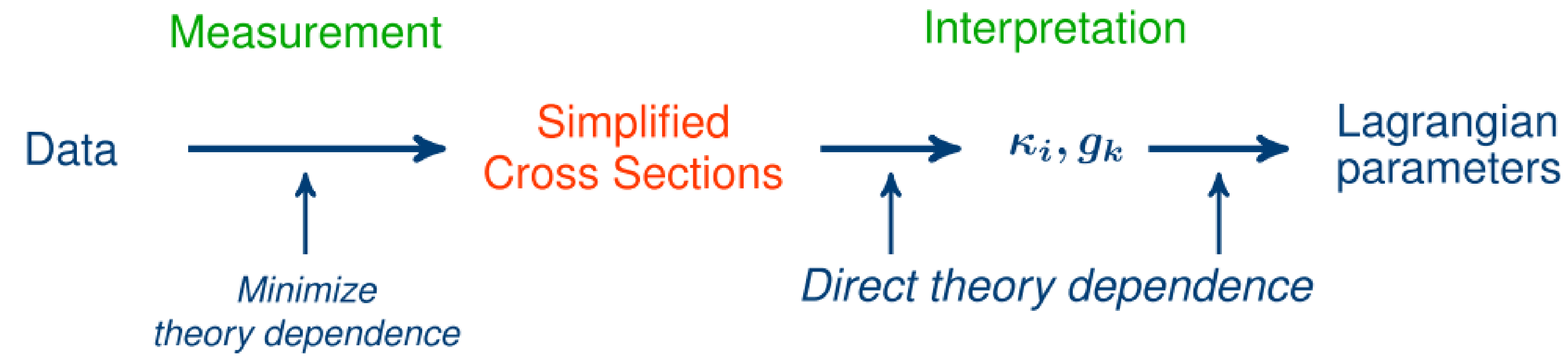
This week: ATLAS >5-sigma ttH



14

G. Salam, LHCP '18

Separating Measurement from Interpretation.



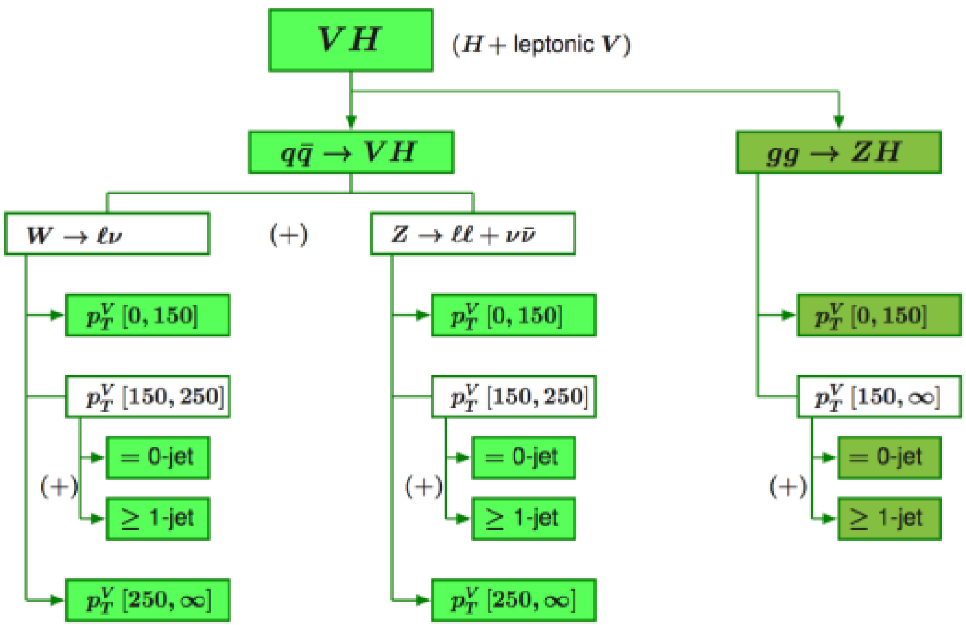
[Talk at the 12th LHC Higgs XSWG](#)
Oct. 2016

Goals

- Minimize theory systematics in measurements
  - Clearer and systematically improvable treatment at interpretation level
- Minimize model dependence in measurements
  - Decouples measurements from assumption of underlying physics model (SM, (non)linear EFT, BSM models)
- Measurements stay long-term useful
- Allows easy further (re)interpretation with different theory inputs/assumptions
  - Improved theory predictions/uncertainties
  - μ_i, κ_i, anomalous couplings, EFT coefficients, specific BSM scenarios

STXS for VH - short intro

- Stage-1 bin split mostly based on VH(bb) analysis categories / variables



- "VH" bins include leptonic VH (H undecayed)
- qq → V(qq)H as part of "VBF" bins
- gg → Z(qq)H as part of "ggF"

- Feedback on the bin split is still welcome, not set in stone!

STXS ≠ fiducial XS (and complementary)
[fid/diff XS minimize theory dependence and acceptance corrections, decayed Higgs, ...]

- optimized for analysis sensitivity (e.g. in this case driven by VH(bb) categorization)
- reducing dominant theory dependence in the measurement (by moving it to the interpretation stage)
- reduced residual theory uncertainties within the measurement of each bin (if residual th. uncertainties become large in the exp. acceptance for a bin, the bin the be further split in sub-categories)

(reference from LesHouches2017)

[Talk at the VH LHC Higgs XSWG soubgroup](#)
Apr. 2018

Prospects

Talk at the Hbb Evidence Oct. 2017

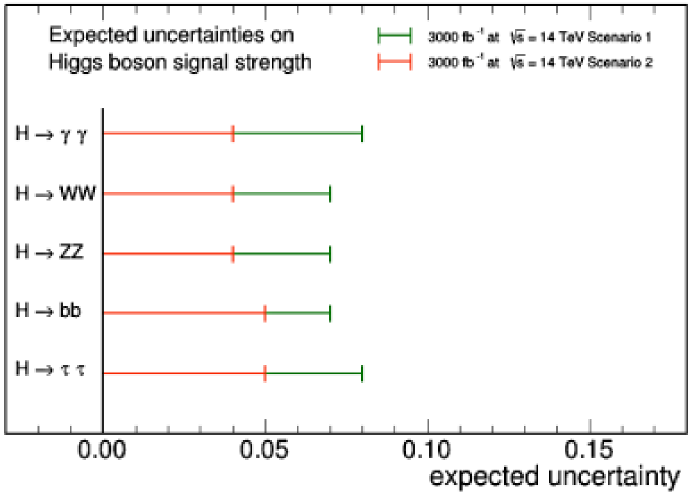
- Goal is decrease uncertainty as much as we can
- Currently both ATLAS and CMS have total $\delta\mu$ ~30%
  - but start to be systematic dominated!

$$\mu^{run2}_{CMS} = 1.19^{+0.21}_{-0.20}(stat)^{+0.34}_{-0.32}(syst)$$

$$\mu^{run2}_{ATLAS} = 1.20^{+0.24}_{-0.23}(stat)^{+0.34}_{-0.28}(syst)$$

- Projections (done in 2013) for HL-LHC

CMS Projection



L (fb^{-1})	bb
300	[11, 14]
3000	[5, 7]

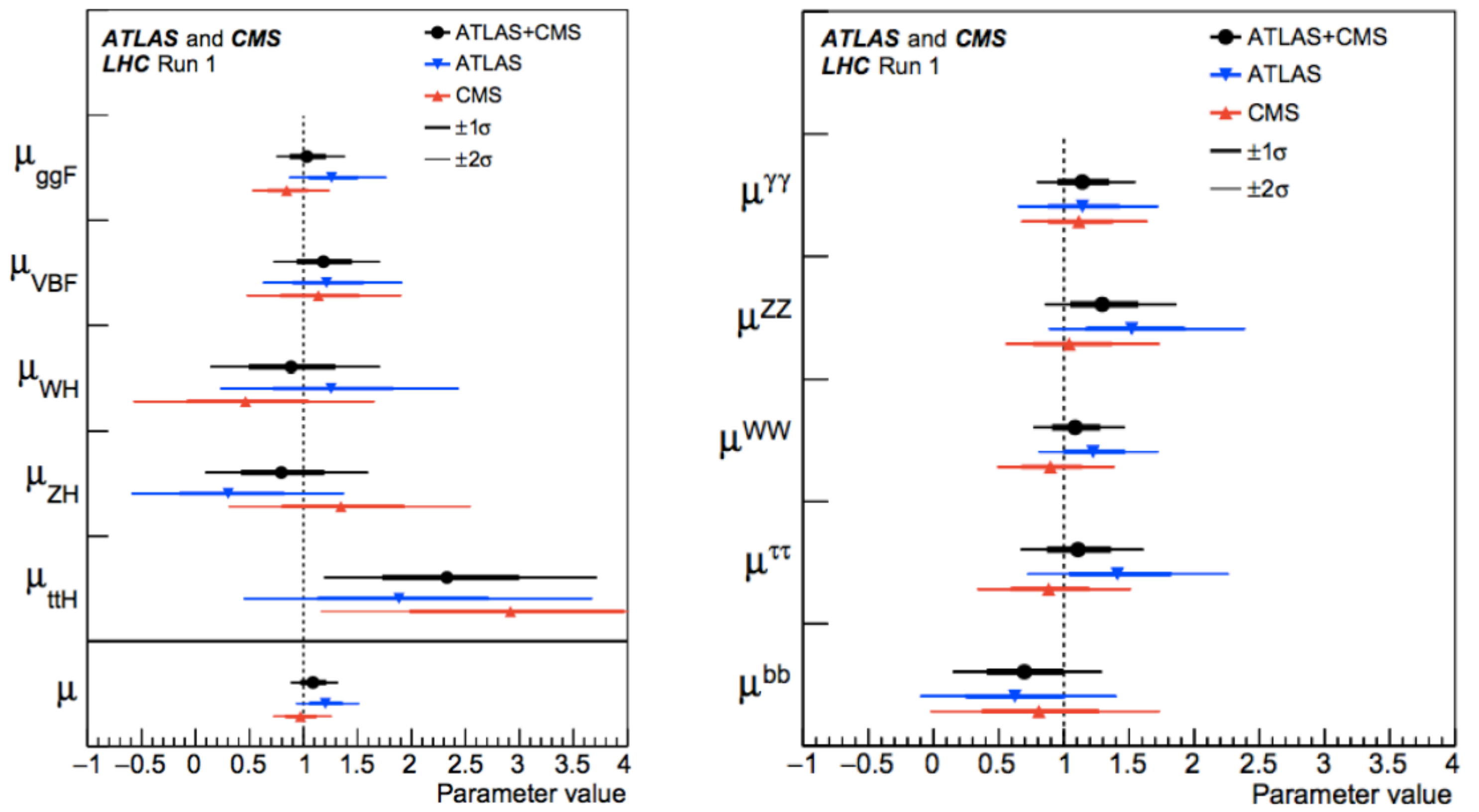
Scenario 1: no change

Scenario 2: Δ theory/2, rest -> $1/\sqrt{L}$

- our main systematics still partially scale with luminosity
  - statistics in CRs
  - important aspects are the modeling and MC statistics for V+jets background
- now with an analysis with SRs+CRs simultaneous fit we can revise these coupling projection in a more realistic way

Higgs at the LHC – Run1 legacy

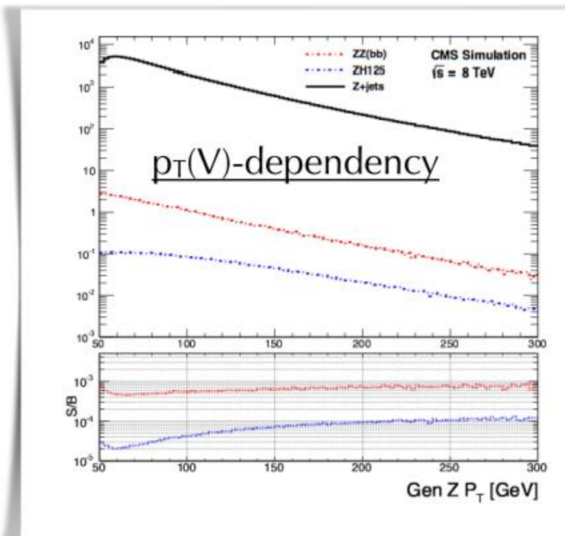
JHEP08(2016)045

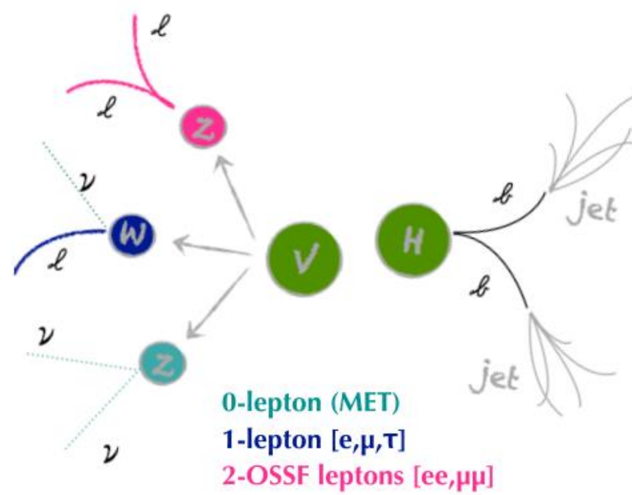


Phys. Rev. D 89 (2014)

VH, Event Topology

- H→ $b\bar{b}$ at LHC is searched in events where H is produced in association with a W or Z boson with high boost (~ 100 GeV)
  - events are triggered by the leptonic decay of the W/Z (e, μ, MET)
  - multi-jet QCD background is highly suppressed





Caterina Vernieri (FNAL)

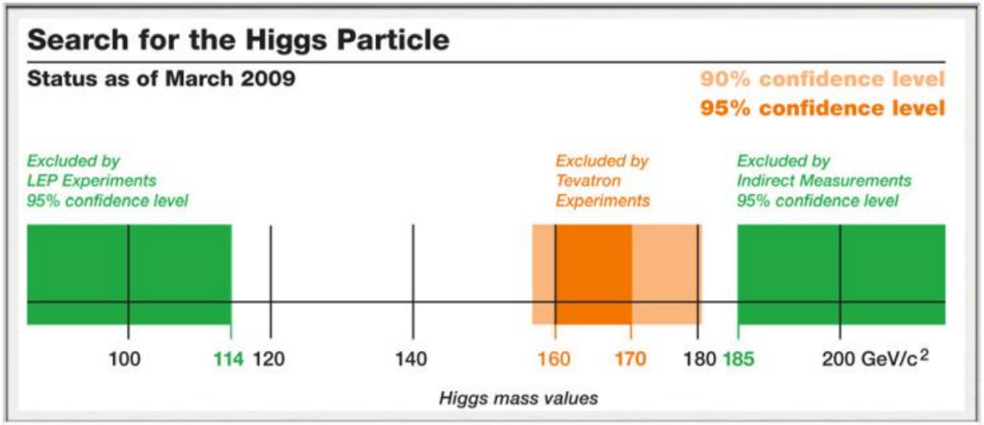
C. Vernieri - SSI 2018

Standard Model Higgs Hunting: Basics

LEP+Tevatron legacy:

low-mass range [114,158] GeV

2009



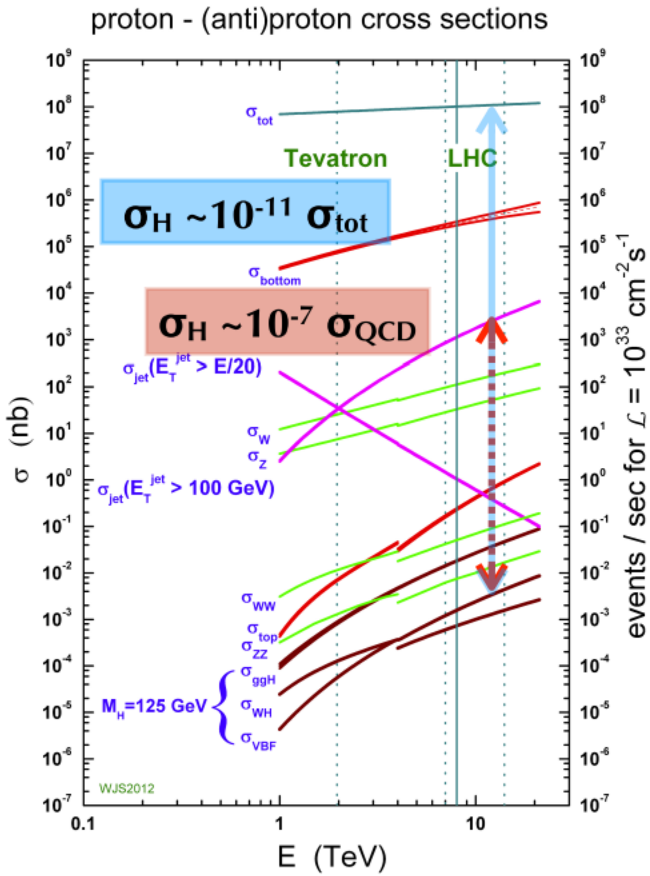
The natural width is less than 100 MeV

observed peak dominated by instrumental mass resolution

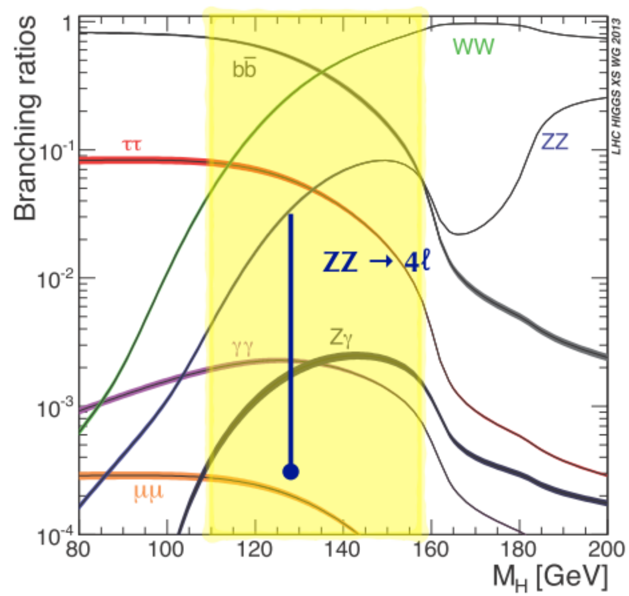
Caterina Vernieri (FNAL)

C. Vernieri - SSI 2018

Standard Model Higgs Hunting: Strategy



H needs
- high energy
- optimal invariant mass resolution
- luminosity



lowest BR decay modes ↔
Excellent mass resolution

C. Vernieri - SSI 2018