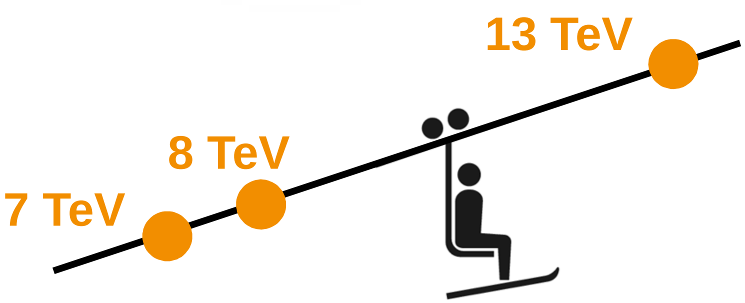
Results and Prospects for ttH at CMS



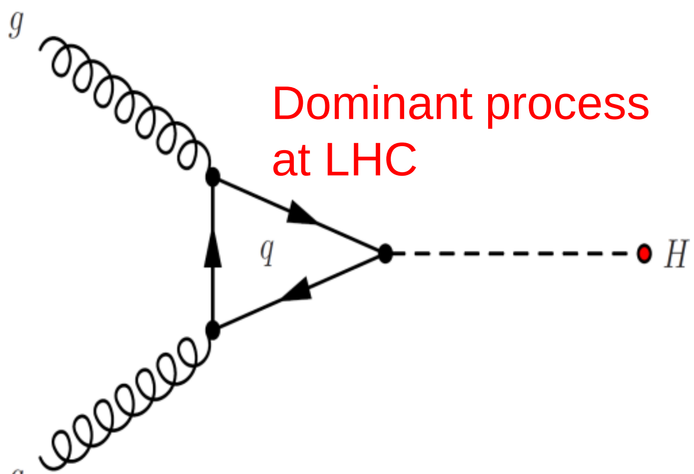
Johannes Hauk
DESY

on behalf of the CMS Collaboration

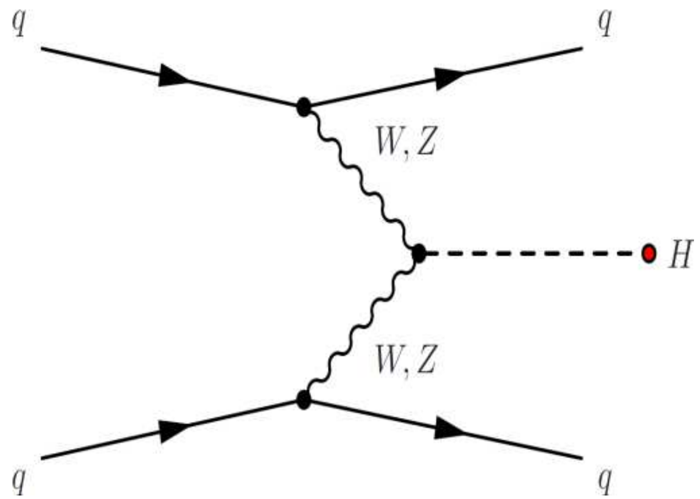
51st Rencontres de Moriond EW 2016
16.03.2016

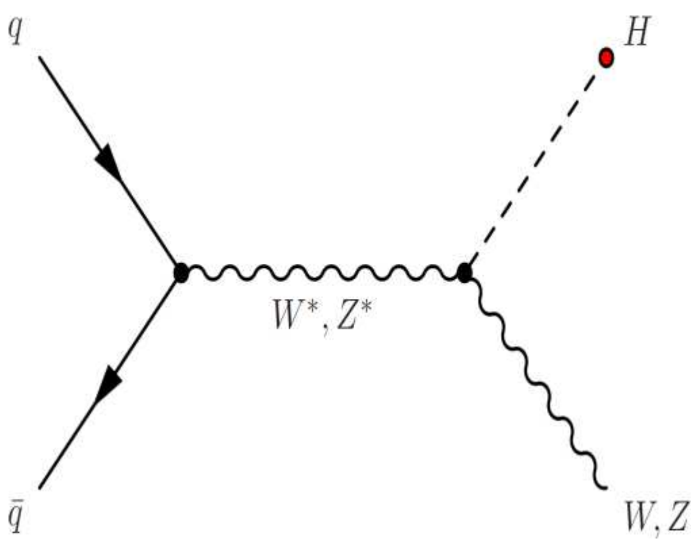
Higgs Boson Production at LHC

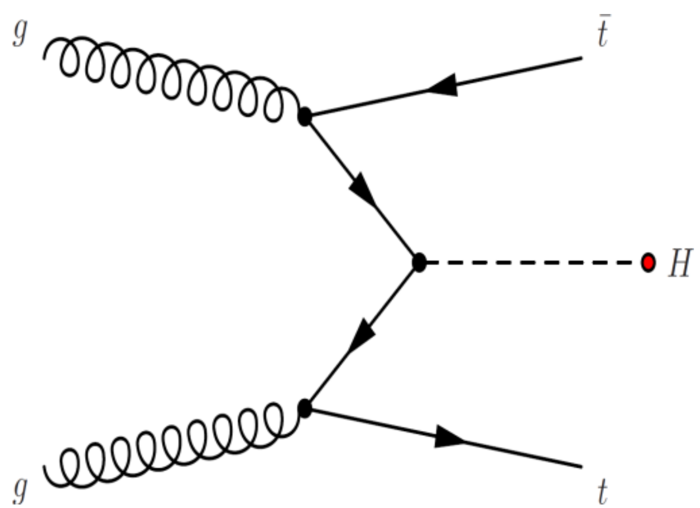
> Production mechanims with very different topologies and cross sections



Dominant process
at LHC

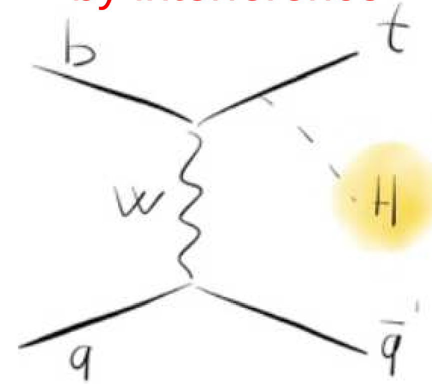






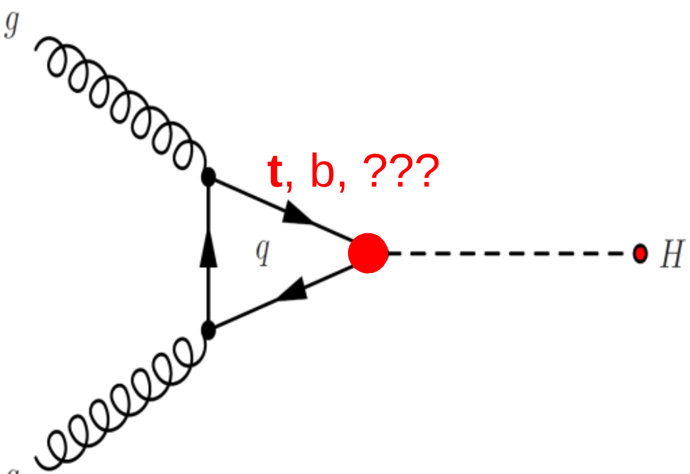
factor ~100 smaller than
inclusive cross section (13 TeV)

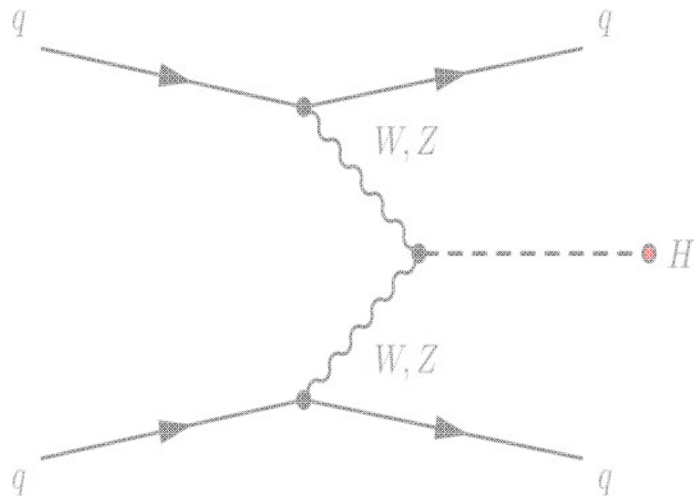
SM: suppressed
by interference

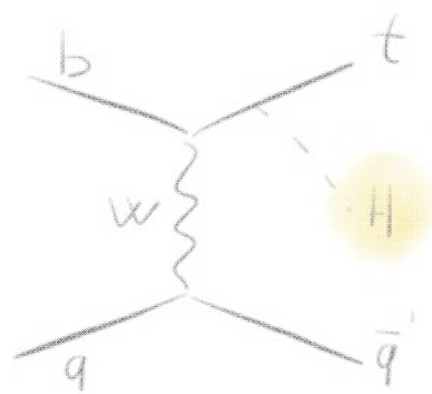


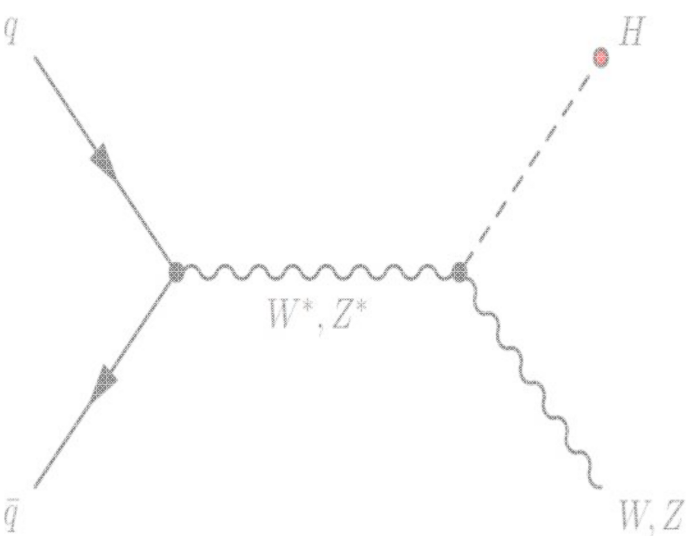
Higgs Boson Production at LHC

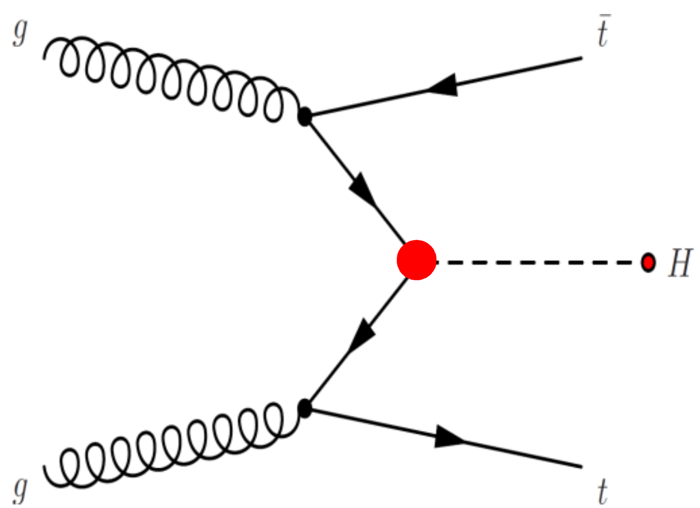
- In SM, top-Higgs Yukawa coupling strongest ($Y_T \approx 1$)





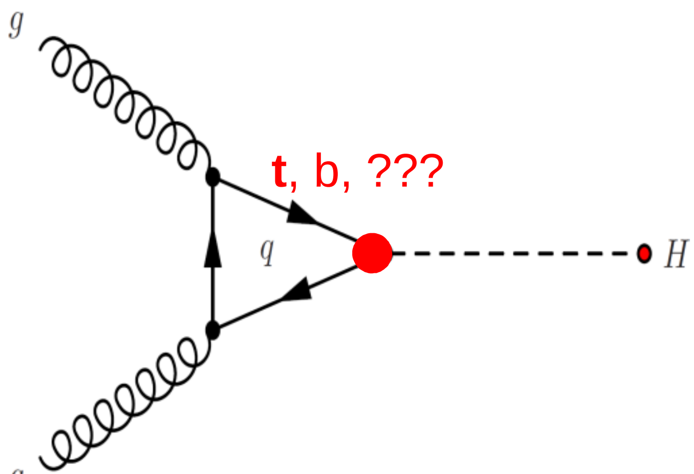


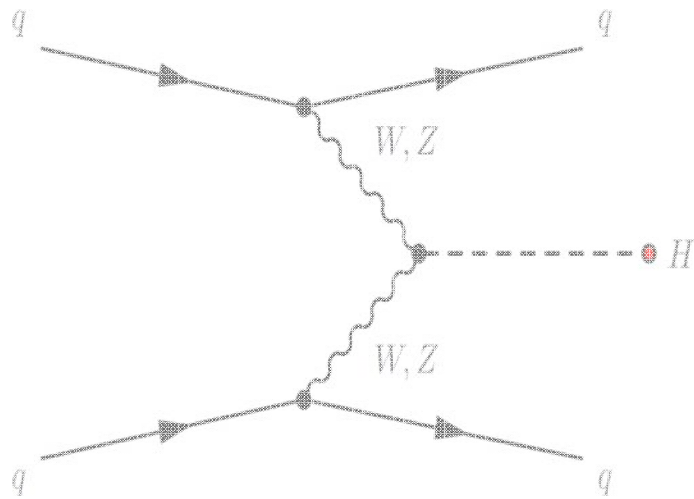


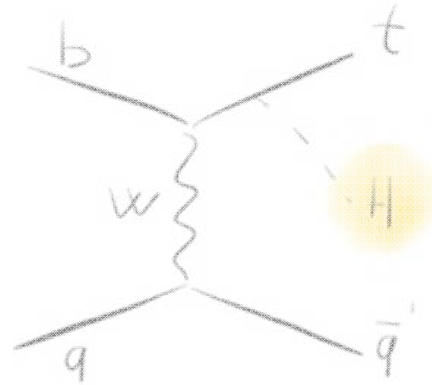


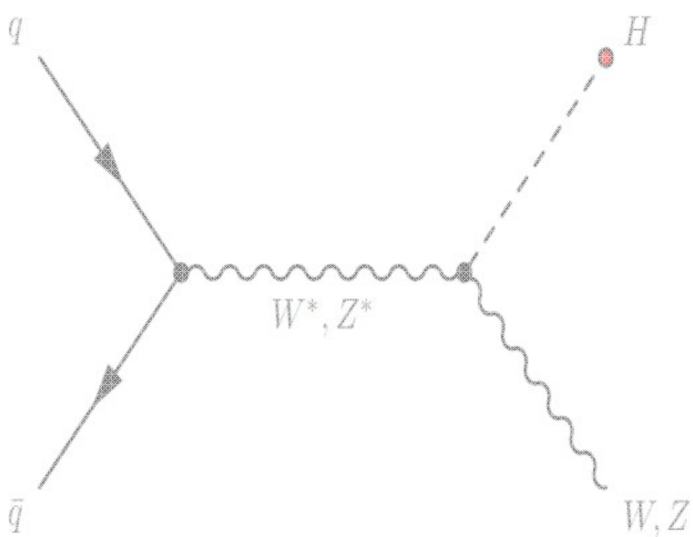
Higgs Boson Production at LHC

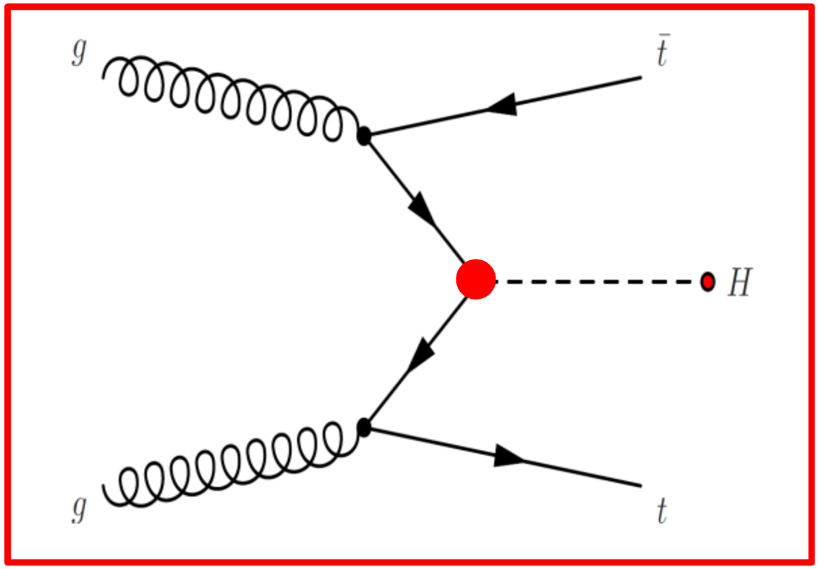
- Indirect constraints from loops, ttH only possibility of direct measurement









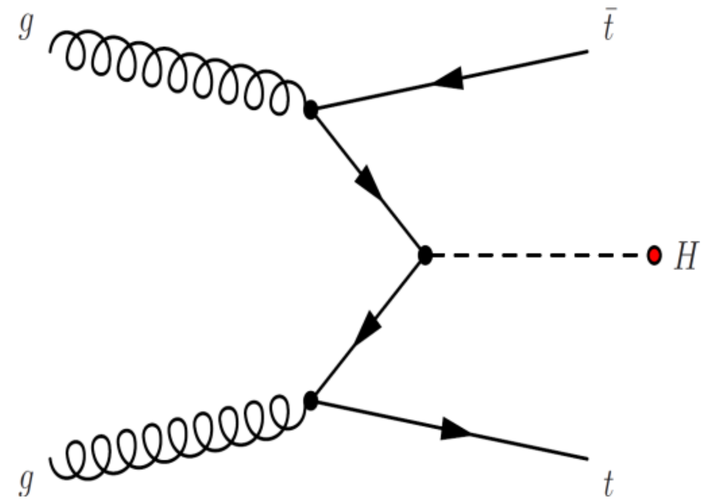


ttH Production

- Strong increase of cross section with center-of-mass energy (m_H = 125 GeV)

ttH (NLO)	Cross section
7 TeV	89 fb
8 TeV	133 fb
13 TeV	507 fb

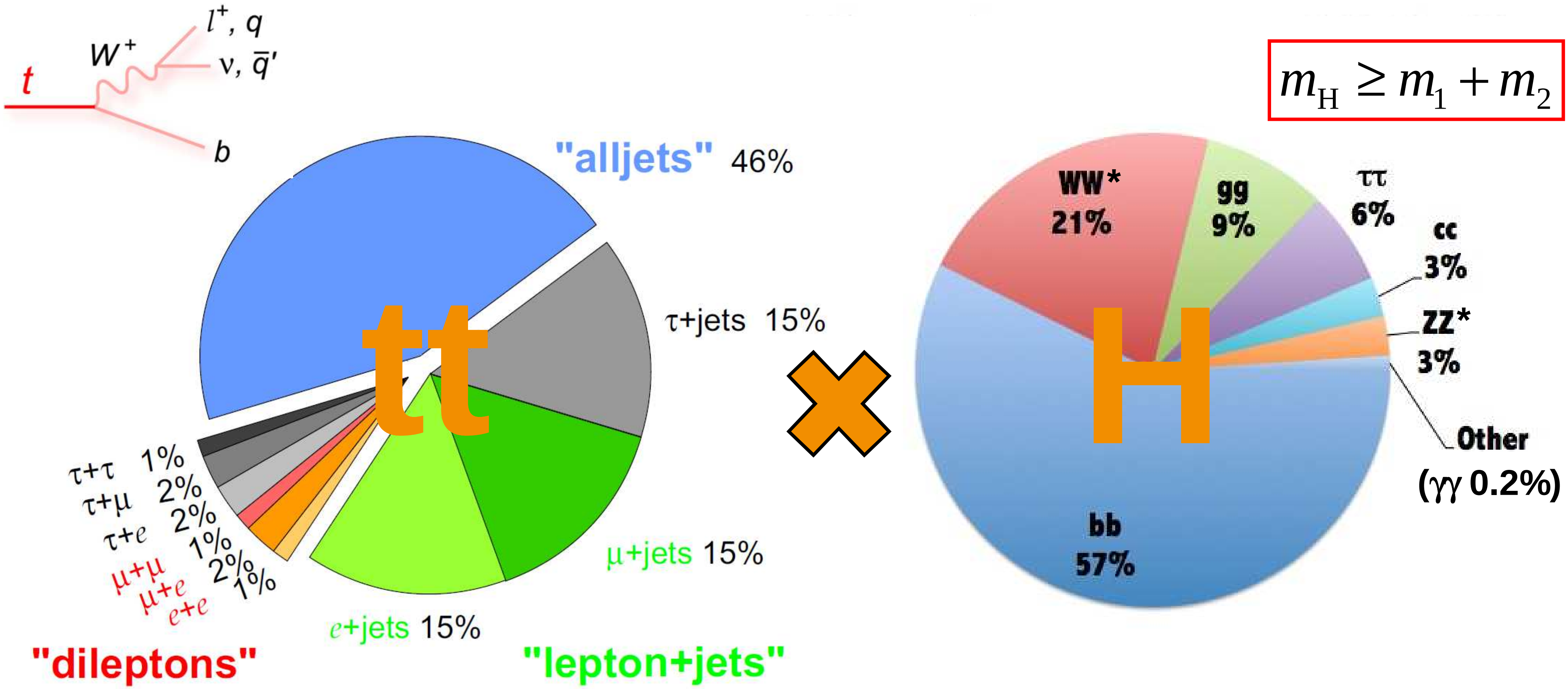
x3.8



- Luminosity of 2015 dataset 2.3 – 2.7 fb^{-1}
  - Equivalent to ≈ 50% of 8 TeV statistics

- Dominant background tt+X
  - Similar increase in cross sections

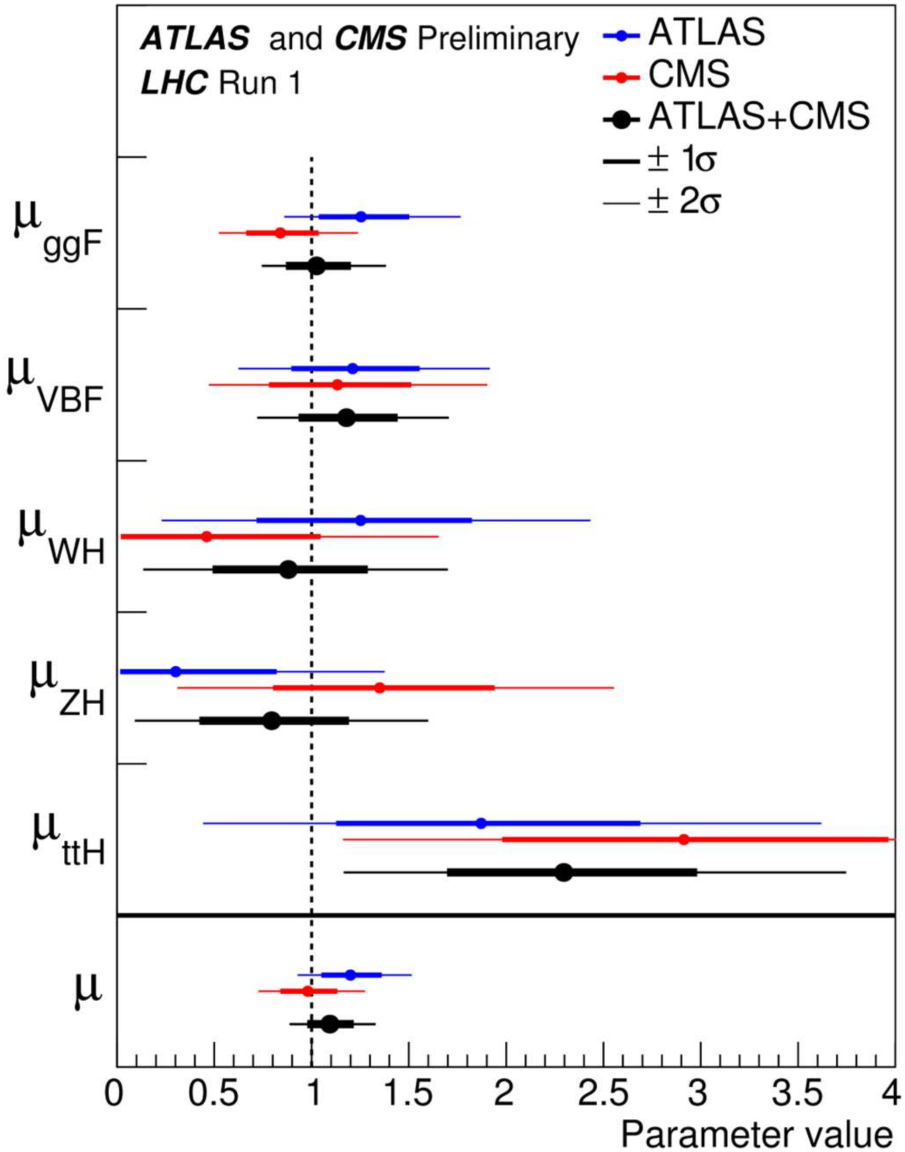
ttH Decays – Very complex Final States



- ttH(γγ): leptonic (dileptons, l+jets), hadronic
- ttH(multileptons): dileptons, l+jets – categorisation via lepton multiplicity
  - multileptons = leptonic decays of H→WW*, ZZ*, ττ
- ttH(bb): dileptons, l+jets

ttH – Knowledge from Run 1

- Combination of all Higgs analysis channels
- μ_{ttH} dominated by: ttH($\gamma\gamma$), ttH(multilepton), ttH(bb)

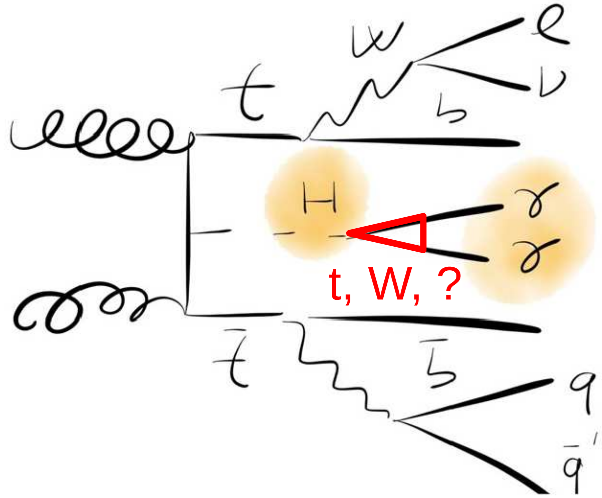


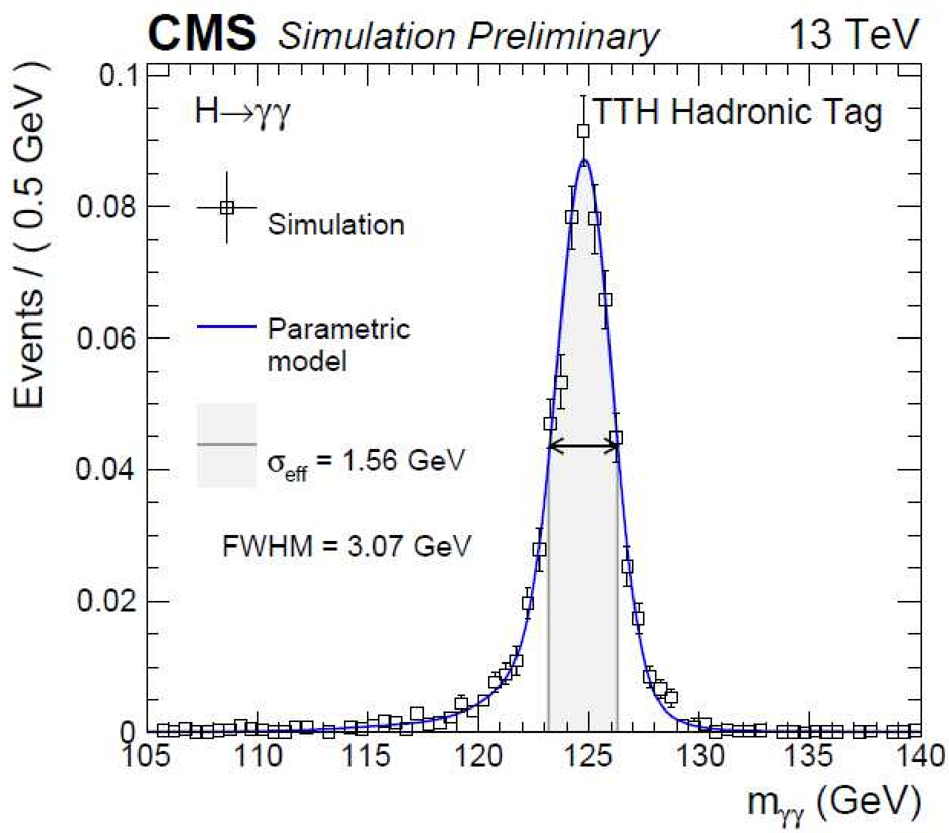
	μ (ttH)
ATLAS	1.9 +0.8 -0.7
CMS	2.9 +1.0 -0.9
Combined	2.3 +0.7 -0.6

Observed (expected) significance
4.4σ (2.0σ)

ttH(γγ)

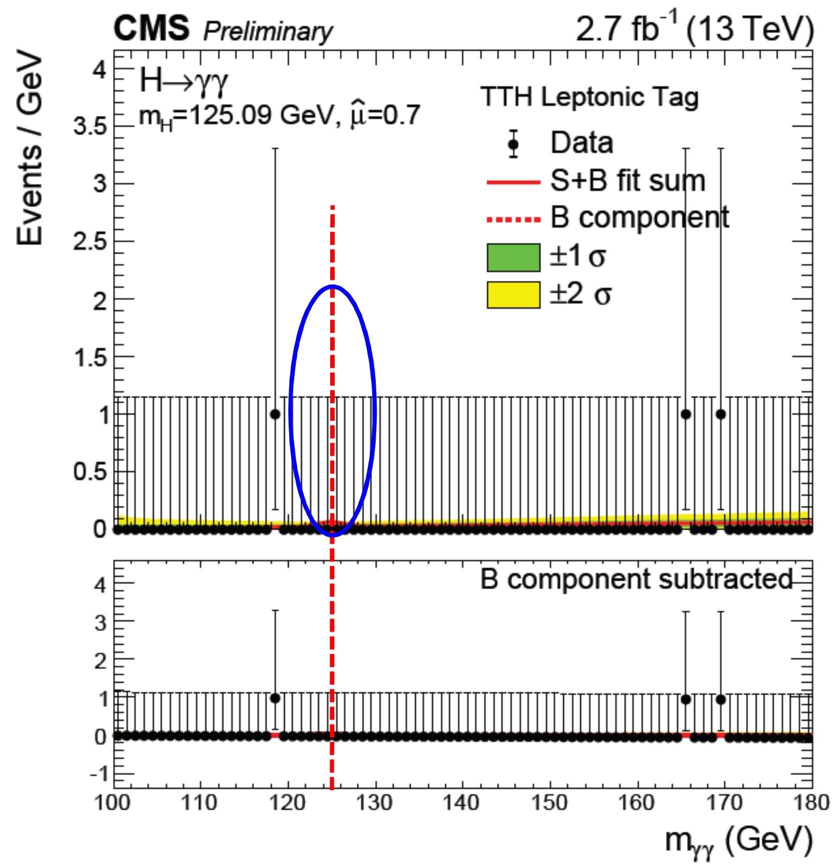
- Tiny branching ratio, but clean resonant signature
- Main backgrounds
  - tt+γγ, tt+jets (→fake photons)
- Integral part of inclusive H→γγ
  - Suppression of fake photons and backgrounds
  - Excellent diphoton mass resolution
- Categorise via leptonic, hadronic
  - Diphoton triggers and offline selection
  - ≥1, 0 leptons
  - ≥2, ≥5 jets
  - ≥1 b-tag

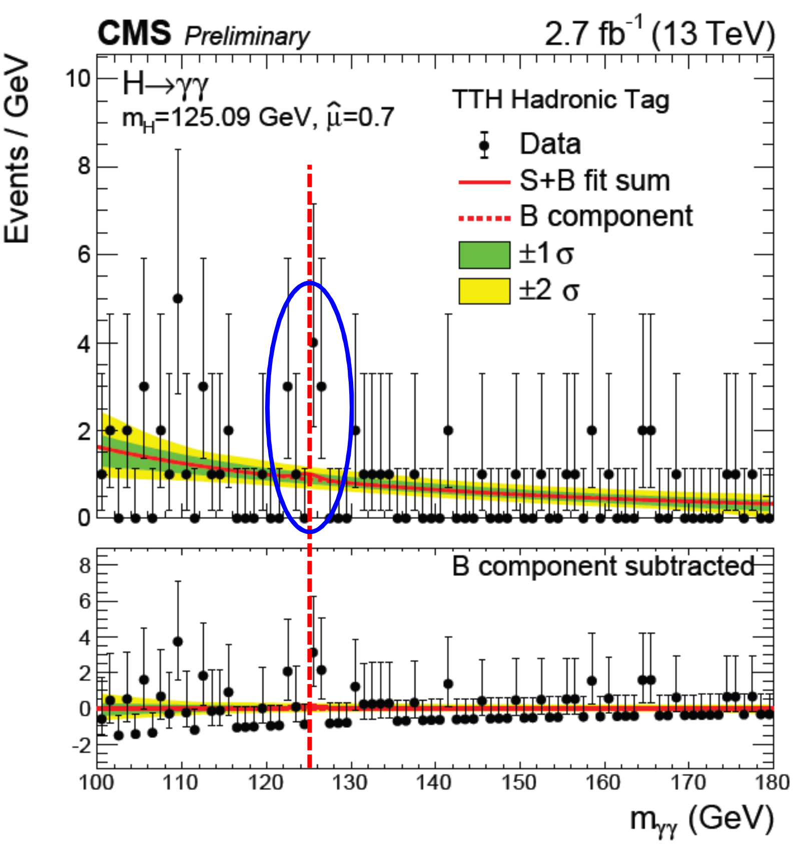




ttH(γγ) – Signal Separation

- Same strategy as for inclusive H→γγ
  - Search for resonance in $m_{\gamma\gamma}$
- Smooth fit functions,
  several functional forms
  - Control regions by
    inverting photon ID + loosened event selection



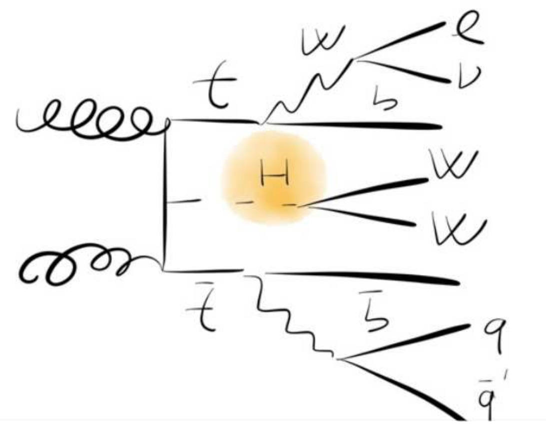


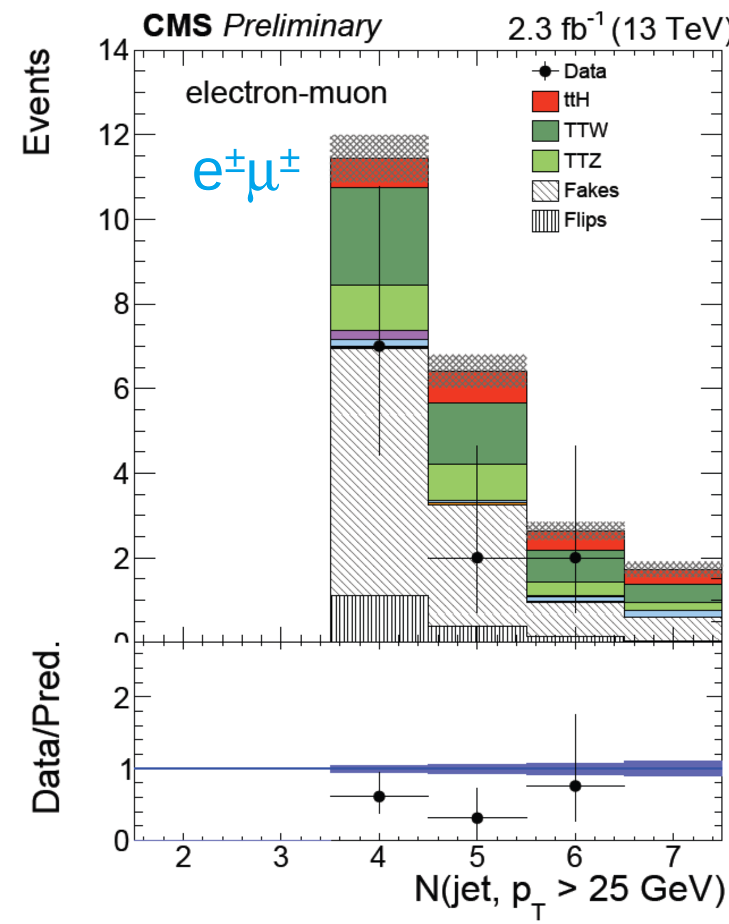
- High-purity ttH selection
  - Statistically limited, small impact of systematics

$$\hat{\mu}_{\text{obs}} = 3.8^{+4.5}_{-3.6}$$

ttH(multileptons)

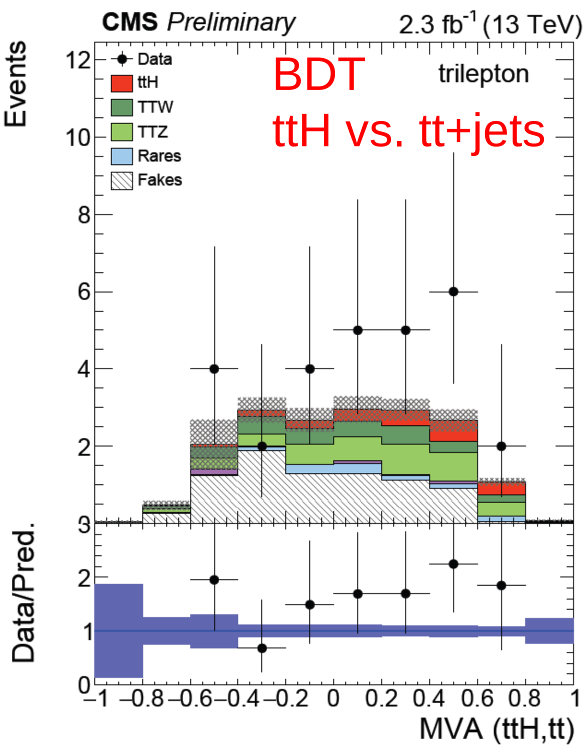
- Smallest irreducible background, focus on reducible
  - tt+V, tt+jets (→fake leptons)
- Categorise 2 same-sign (SS) leptons, ≥3 leptons
  - Lepton triggers and offline selections
  - ≥4, ≥2 jets
  - ≥1 b-tag
  - Sub-categories: lepton flavour, lepton charge, presence of τ_h, presence of 2 b-tags
- Separation of prompt leptons from fakes via Boosted Decision Tree (BDT)
- Modelling of fake lepton backgrounds from control region relaxing lepton selection
  - Mis-identification (fakes)
  - Charge mis-reconstruction of electrons (flips)

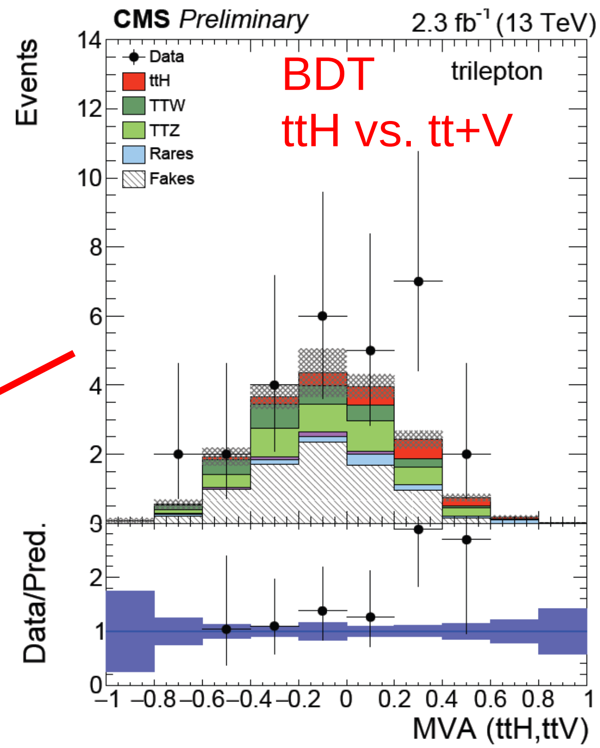


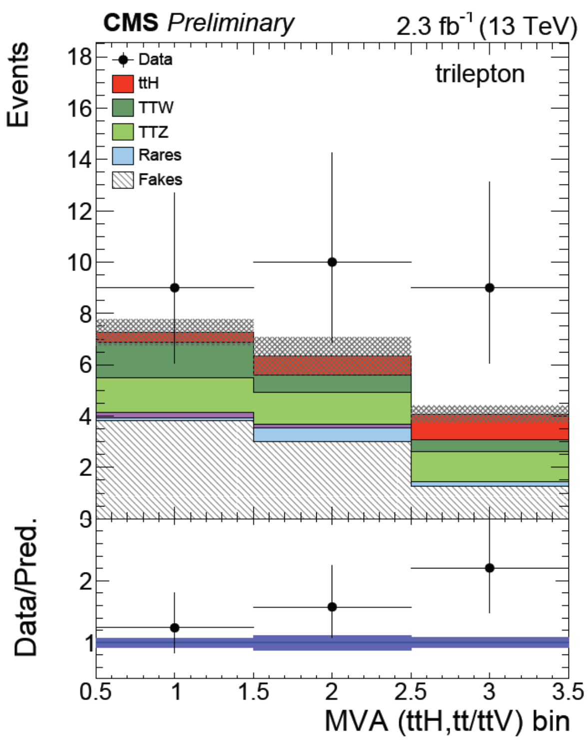


ttH(multileptons) – Signal Separation

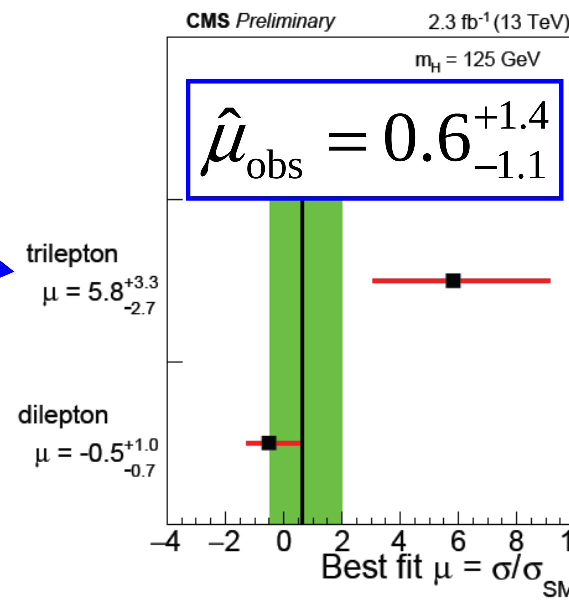
trilepton





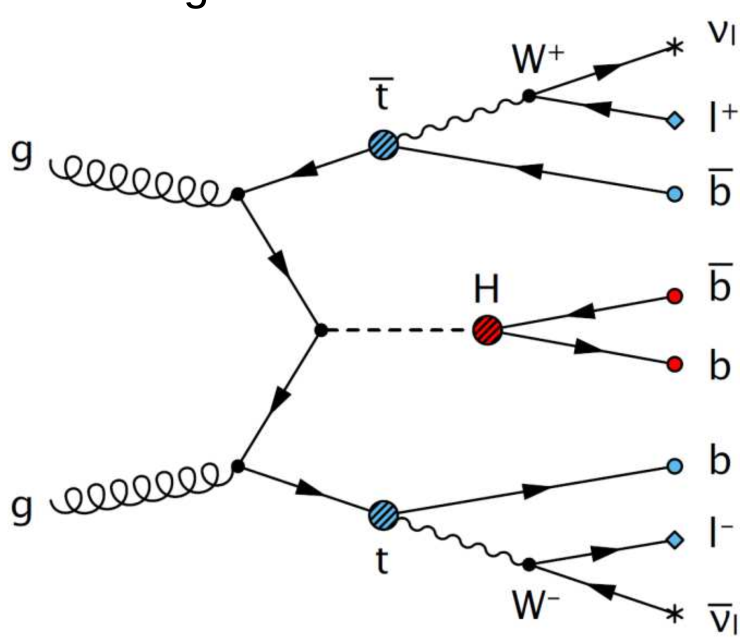


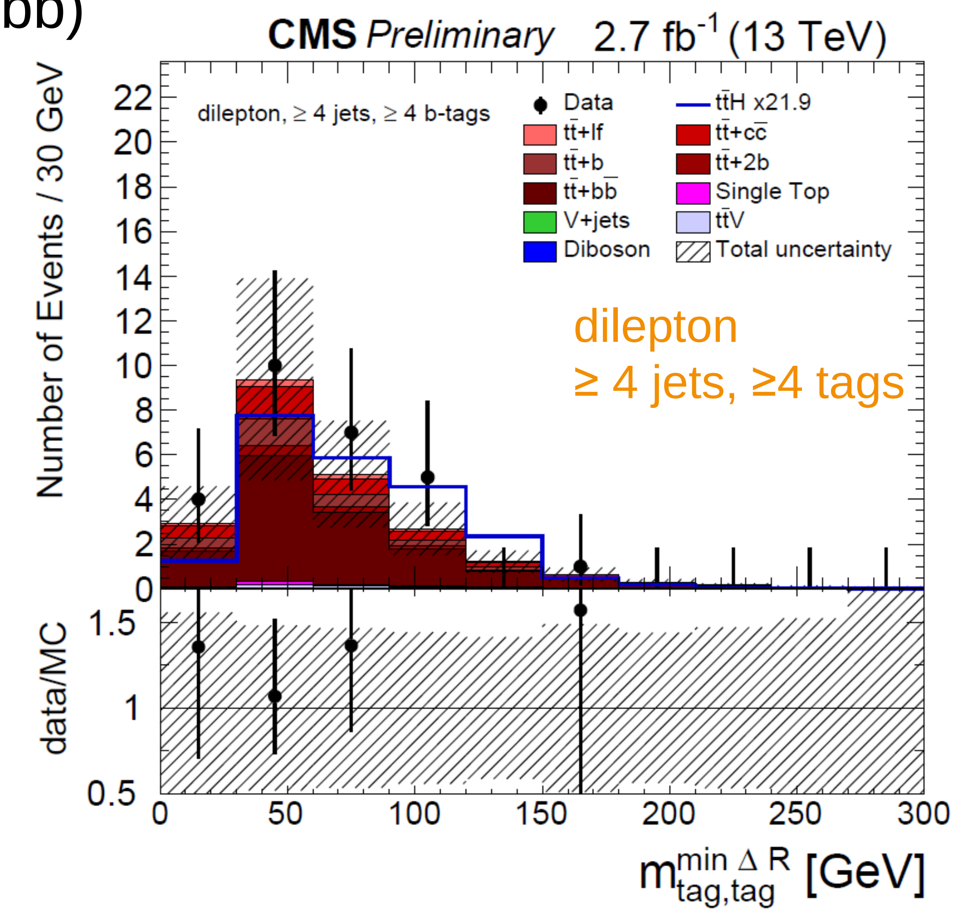
- Combined fit of all sub-categories



ttH(bb)

- tt+jets overwhelming background for ttH(bb)
  - Especially tt+bb irreducible, theoretically challenging
- Categorise via l+jets, dilepton
  - Lepton triggers and offline selections
  - =1, =2 opposite-sign (OS) leptons
  - ≥4, ≥3 jets
  - ≥2 b-tags

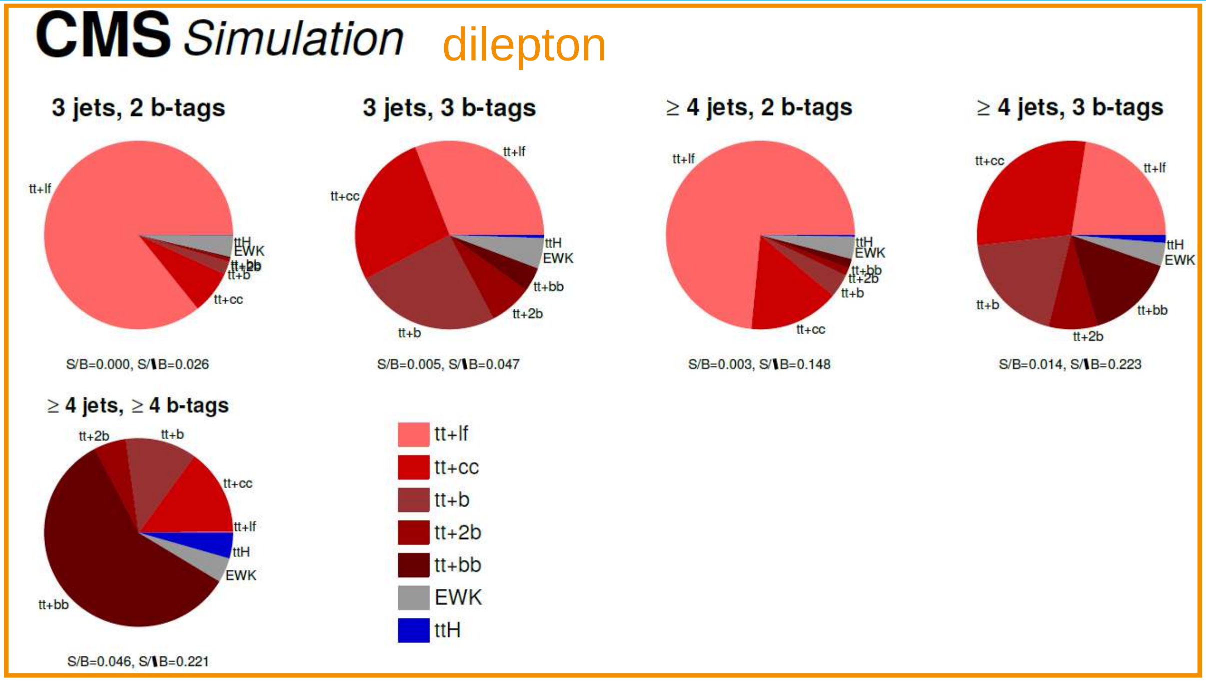


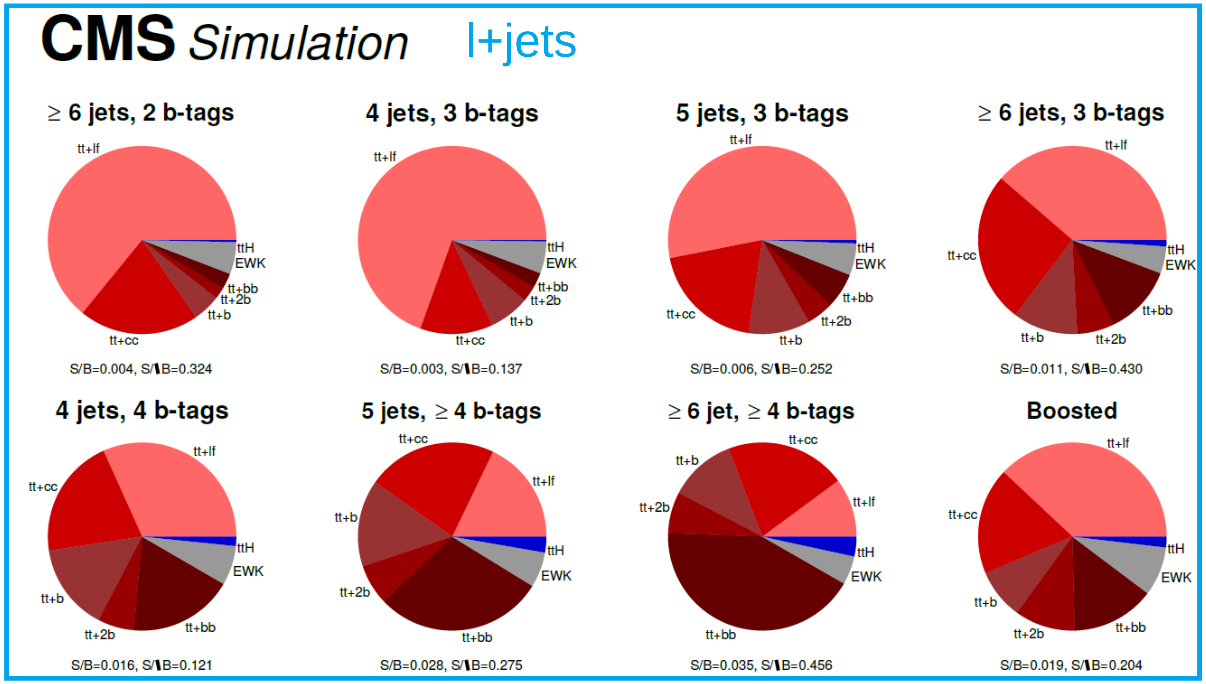


- Limited mass resolution for H→bb, jet combinatorics
  - Dilepton: minimal non-tt backgrounds, minimal jet combinatorics
  - l+jets: high statistics

ttH(bb) – Event Classification

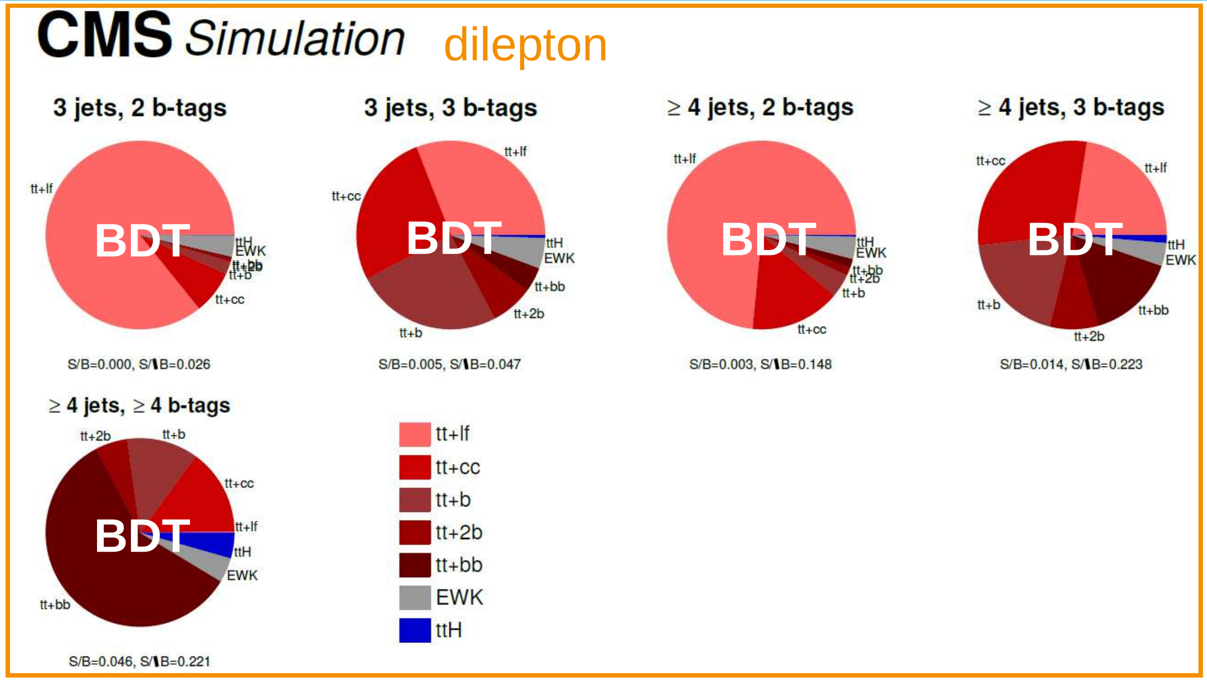
- Classify by number of jets, number of b-tags
  - Background-like: constrain systematic uncertainties
  - Signal-like: (close to) topology of ttH
- Boosted category for first time (l+jets)
  - Fat-jet algorithm
  - Identify hadronic top and Higgs using substructure information
- 13 orthogonal categories

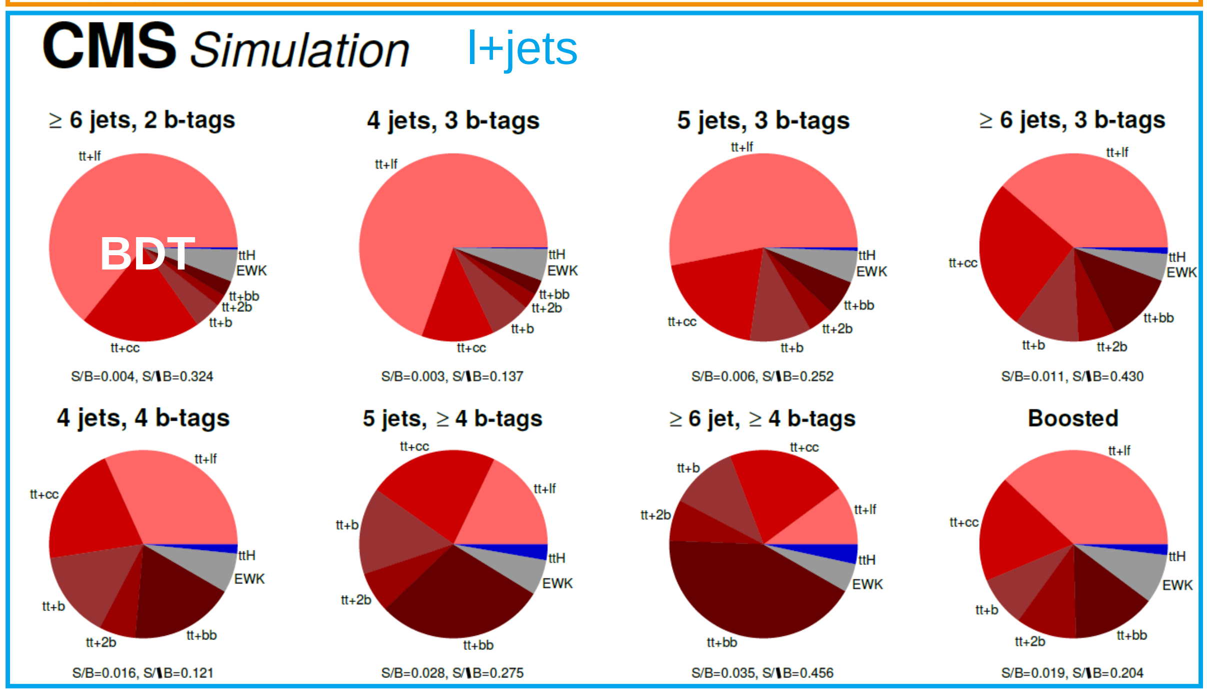




ttH(bb) – Signal Separation

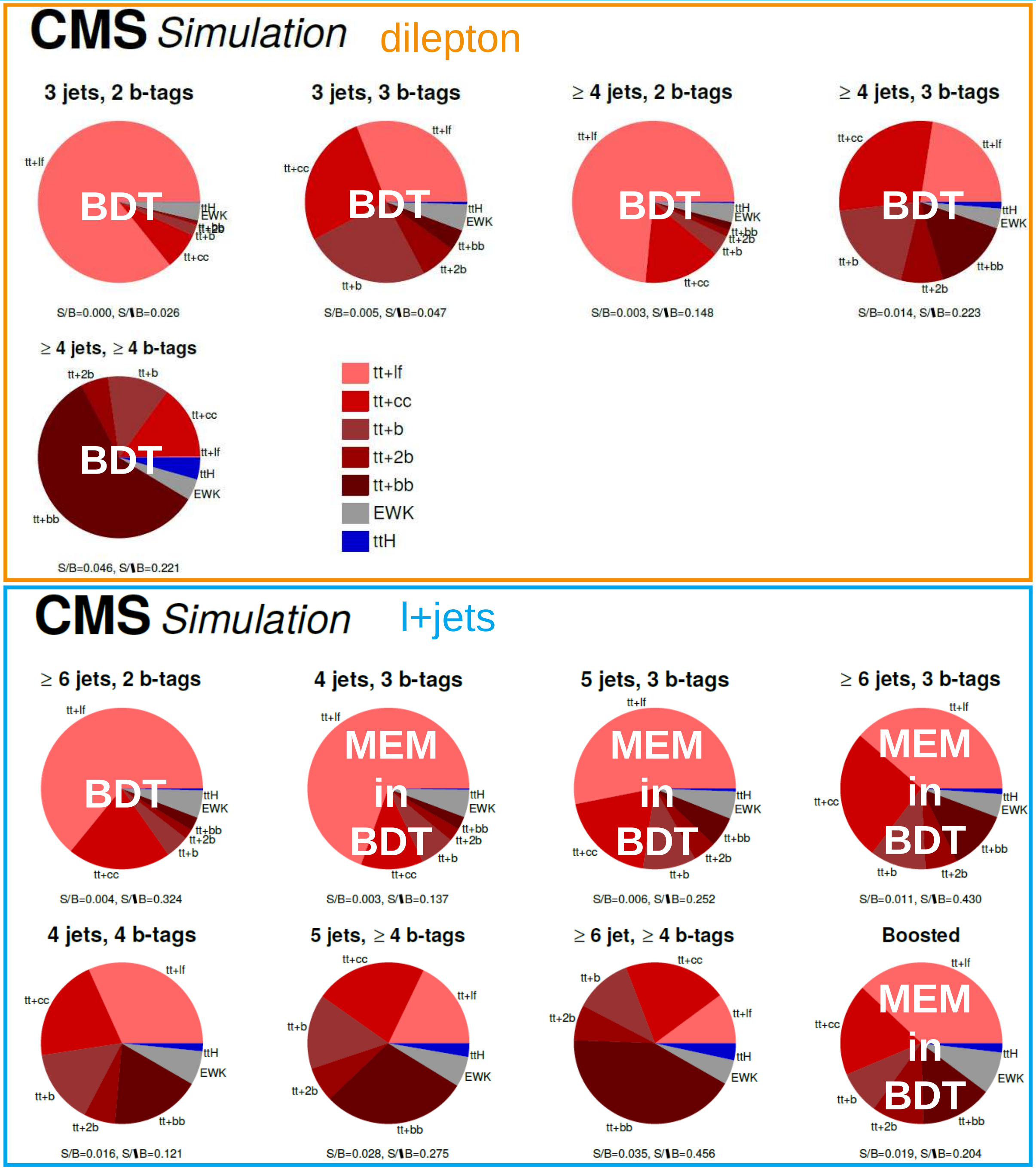
- In each category, BDT with different variables





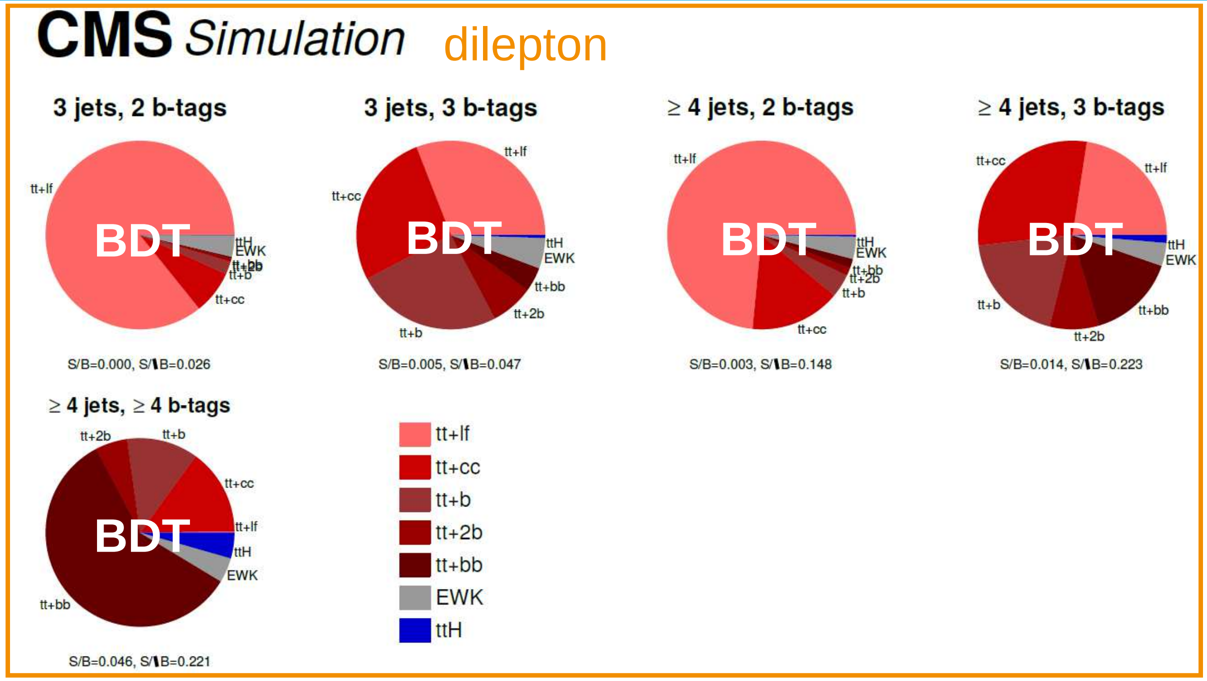
ttH(bb) – Signal Separation

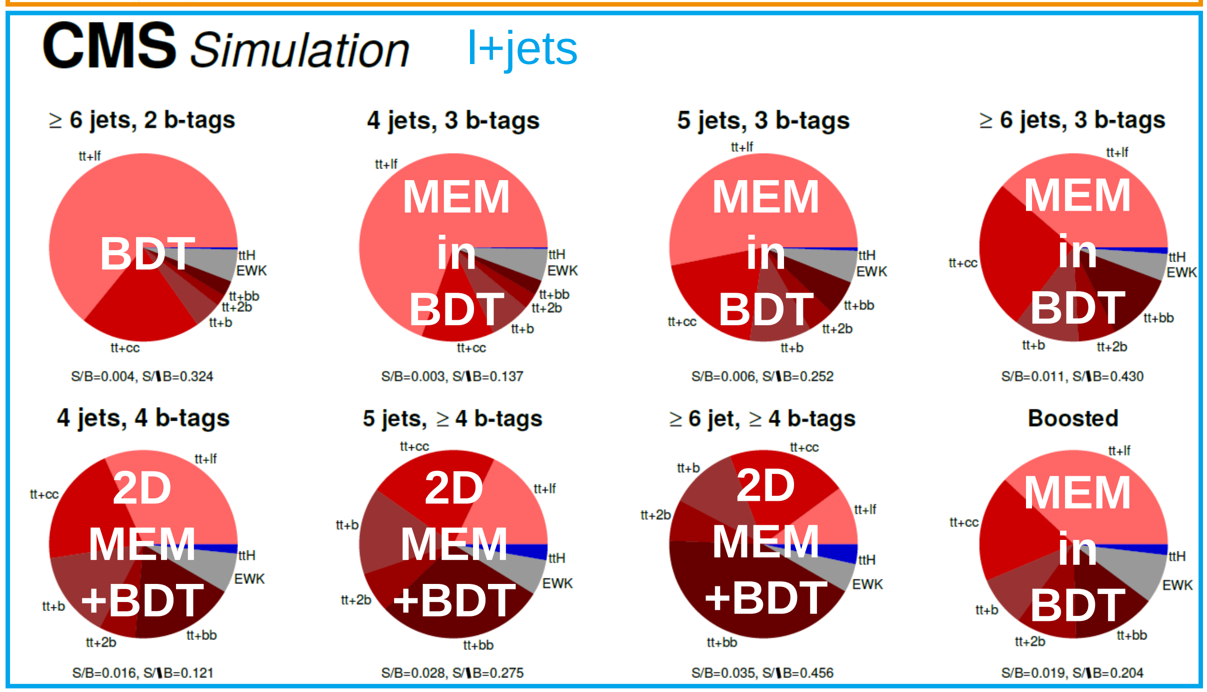
- In l+jets, inclusion of Matrix Element Method (MEM)
  - Use tt+bb as background hypothesis, permute over jet-quark associations
  - MEM discriminant as input variable in 3 b-tag categories and boosted category



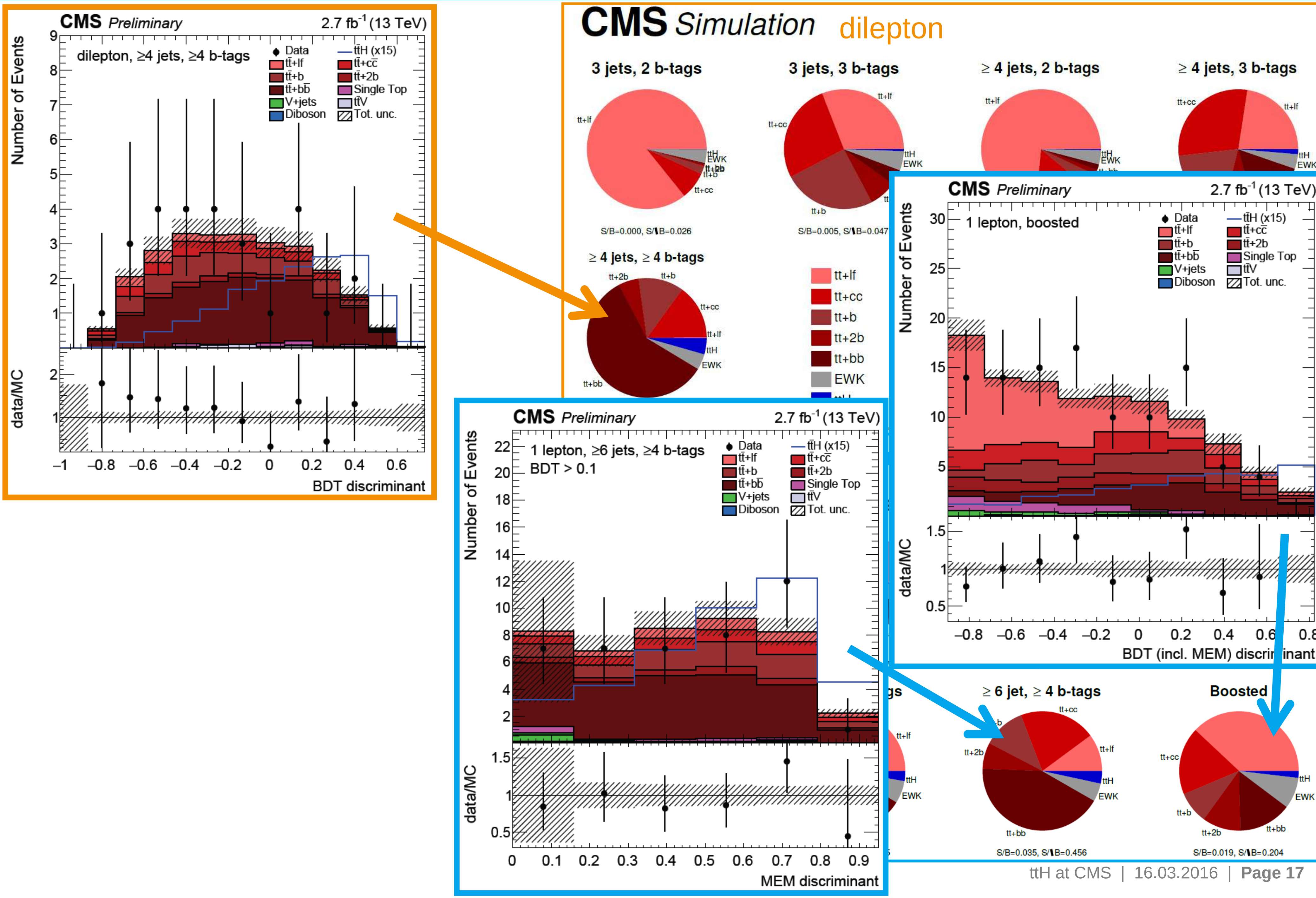
ttH(bb) – Signal Separation

- In l+jets, inclusion of Matrix Element Method (MEM)
  - Use tt+bb as background hypothesis, permute over jet-quark associations
  - MEM discriminant as input variable in 3 b-tag categories and boosted category
  - 2D BDT-MEM analysis in ≥4 b-tag categories





ttH(bb) – Signal Separation



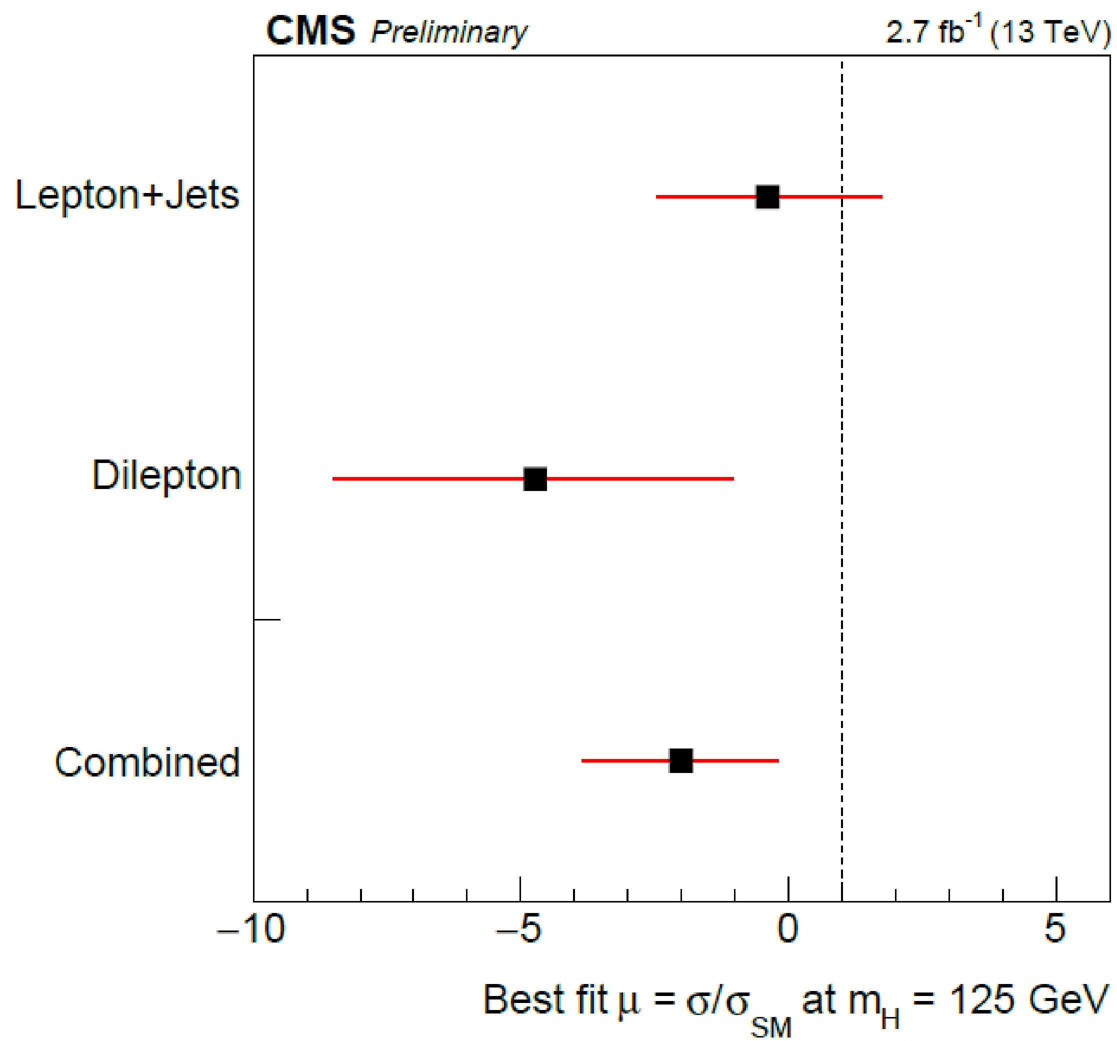
ttH(bb) – Results

- Combined fit of all categories

- Systematics dominated

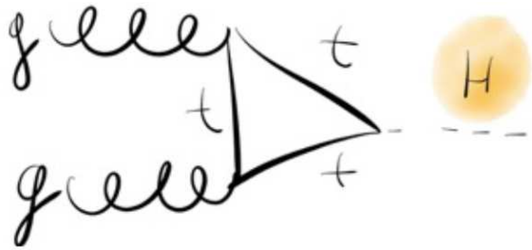
$$\hat{\mu}_{obs} = -2.0^{+1.8}_{-1.8}$$

1.7σ below SM expectation



Outlook

- Much more data to come
  - Expect ≈30 fb^{-1} in 2016
- ttH observation and Yukawa coupling measurement amongst priorities for Run 2 at LHC
  - Is ttH like in SM, reveals sings of new physics ?
- Key to find "hidden" loop contribution



- ttH of importance throughout whole LHC era

Summary

- Top-Higgs coupling only accessible via associated ttH production
  - Important to understand loop contributions
- First 13 TeV measurements performed
  - In γγ, multilepton and bb decay modes
  - Similar sensitivity as Run 1 analysis
  - Overall in agreement with SM
- Foundation with improved analysis techniques for 13 TeV
  - Many more results to come with incoming data

ttH(γγ) $\hat{\mu}_{obs} = 3.8^{+4.5}_{-3.6}$

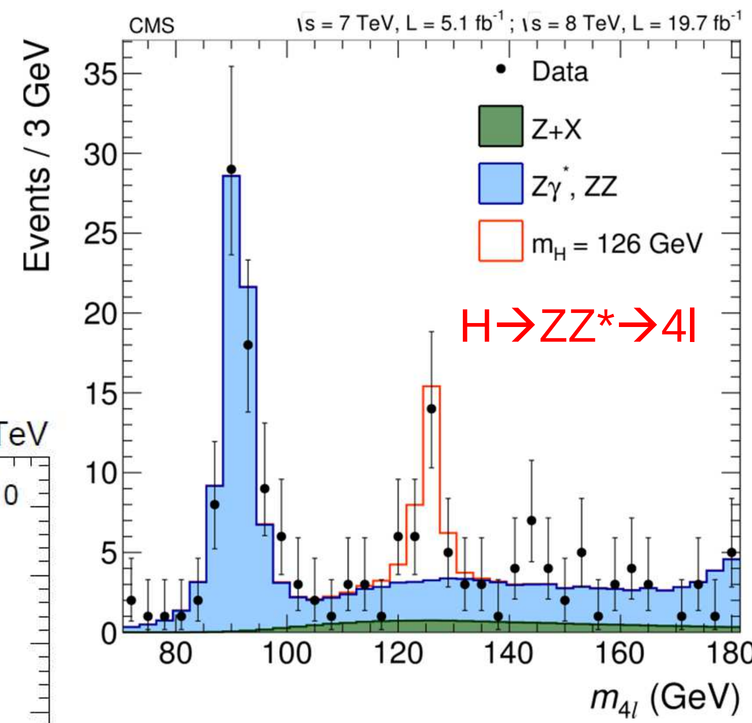
ttH(multilepton) $\hat{\mu}_{obs} = 0.6^{+1.4}_{-1.1}$

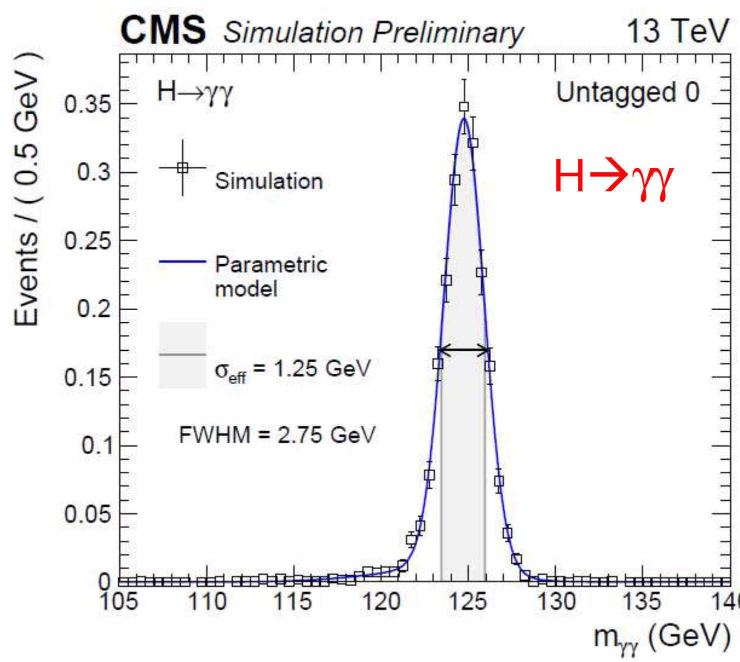
ttH(bb) $\hat{\mu}_{obs} = -2.0^{+1.8}_{-1.8}$

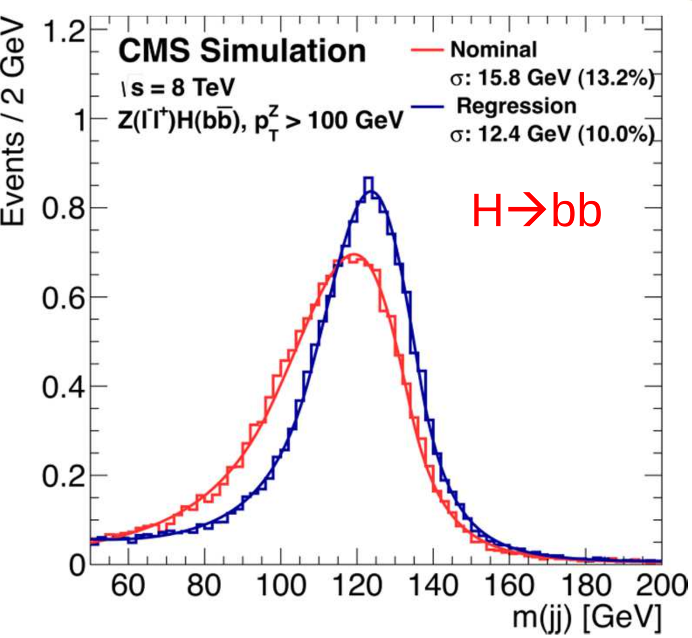
Backup

Mass Resolution of Higgs Decays

- Higgs mass can be reconstructed in certain decay channels
  - Excellent resolution for H→ZZ*→4l, and H→γγ
  - Poor mass resolution of H→bb







Cross section of ttbb and ratio to ttjj – Run 1

- Inclusive cross section (ratios) measured
  - 7 TeV (jet p_T>20 GeV), dilepton

$$\frac{\sigma_{\text{ttbb}}}{\sigma_{\text{ttjj}}} = (3.6 \pm 1.1\,(\text{stat}) \pm 0.9\,(\text{syst}))\,\%$$

PAS-TOP-12-024

  - 8 TeV (jet p_T>40 GeV), dilepton and l+jets

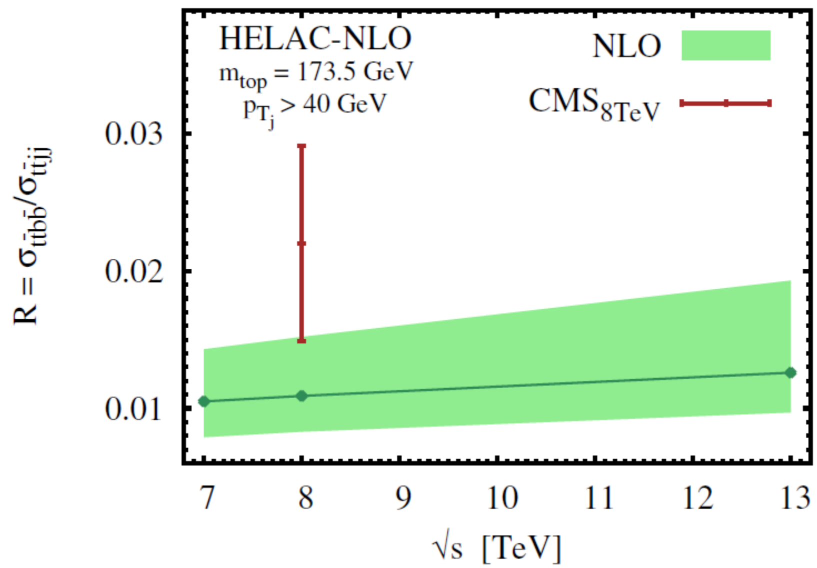
$$\frac{\sigma_{\text{ttbb}}}{\sigma_{\text{ttjj}}} = (2.2 \pm 0.3\,(\text{stat}) \pm 0.5\,(\text{syst}))\,\%$$

Phys. Lett. B 746 (2015) 132-153

$$\frac{\sigma_{\text{ttbb}}}{\sigma_{\text{ttjj}}} = (1.2 \pm 0.4\,(\text{stat}) \pm 0.03\,(\text{syst}))\,\%$$

PAS-TOP-13-016

- And calculated



jet p_T > 40 GeV

ttbb/ttjj (NLO)	Cross-section ratio
7 TeV	1.05%
8 TeV	1.09%
13 TeV	1.26%

ttjj (NLO)	Cross section
7 TeV	13.6 pb
8 TeV	21.0 pb
13 TeV	85.5 pb

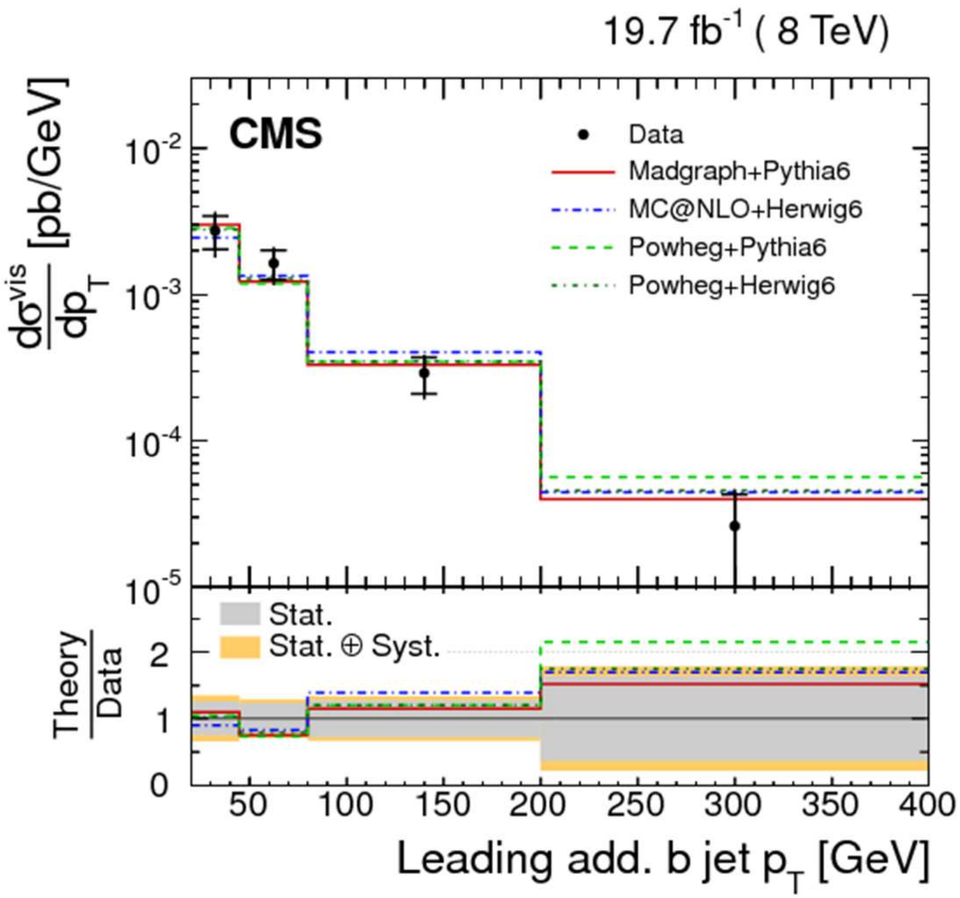
ttbb (NLO)	Cross section
7 TeV	142 fb
8 TeV	229 fb
13 TeV	1078 fb

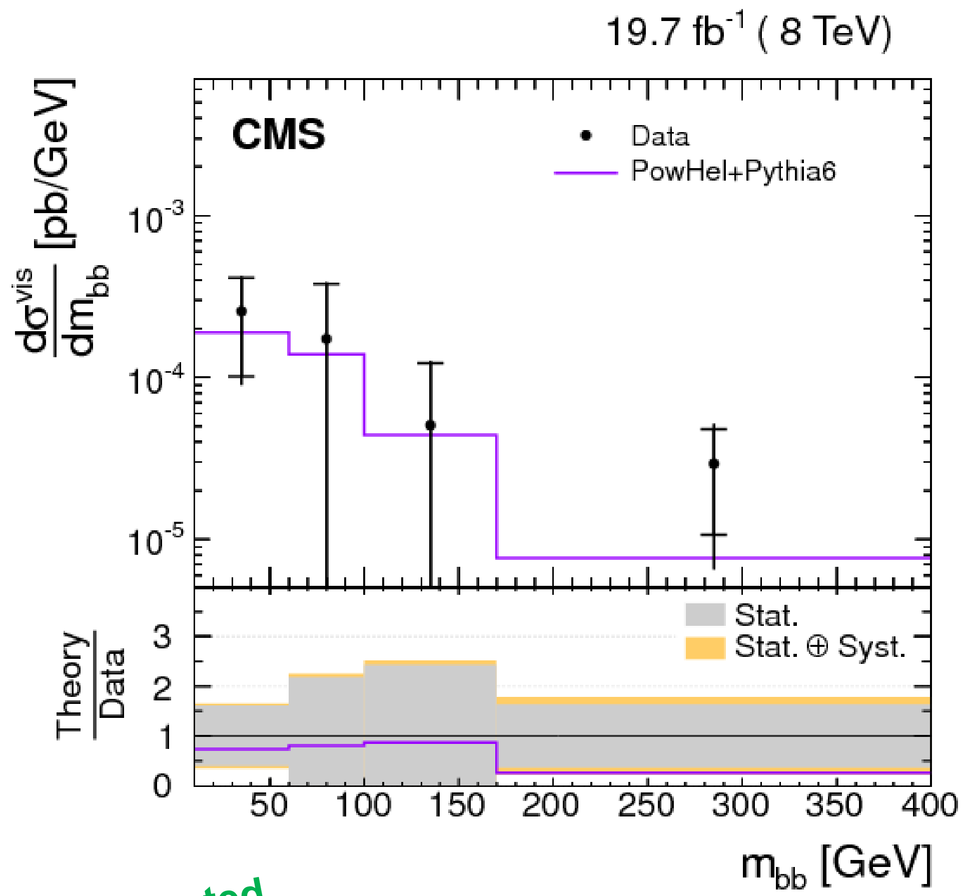
JHEP 07 (2014) 135

Differential Cross Sections of ttbb – Run 1

> Differential cross sections of properties of additional b jets

- Shape comparisons to different simulations
- Comparison to full next-to-leading order (NLO) calculation

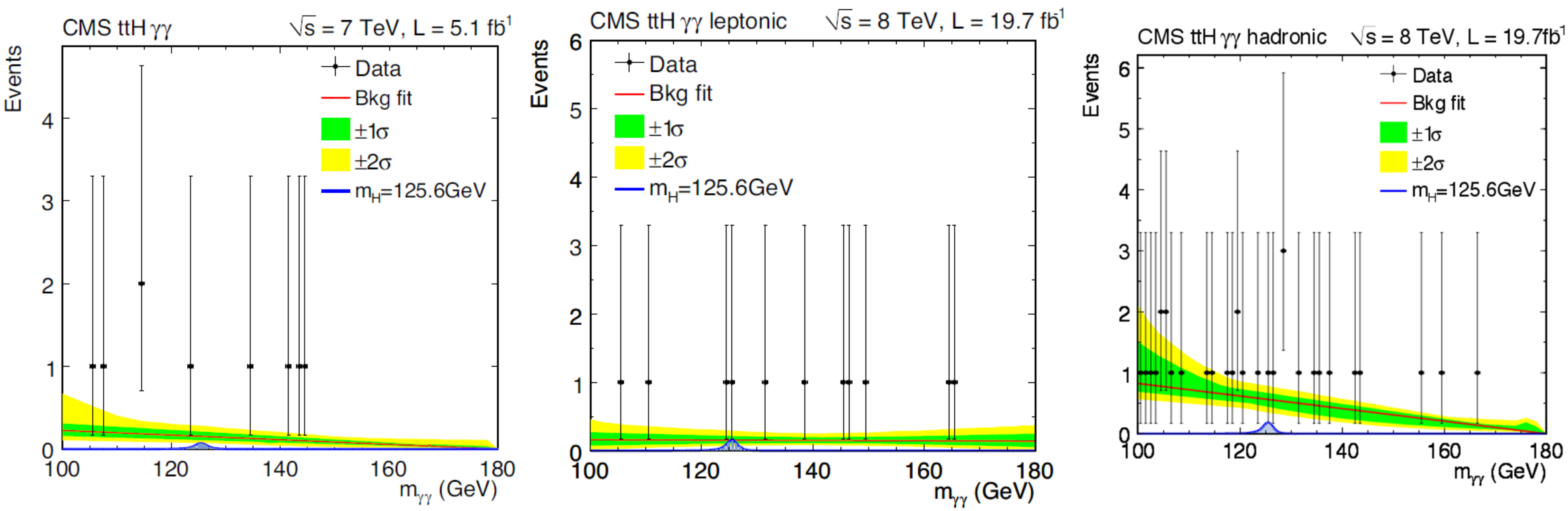




Statistics dominated

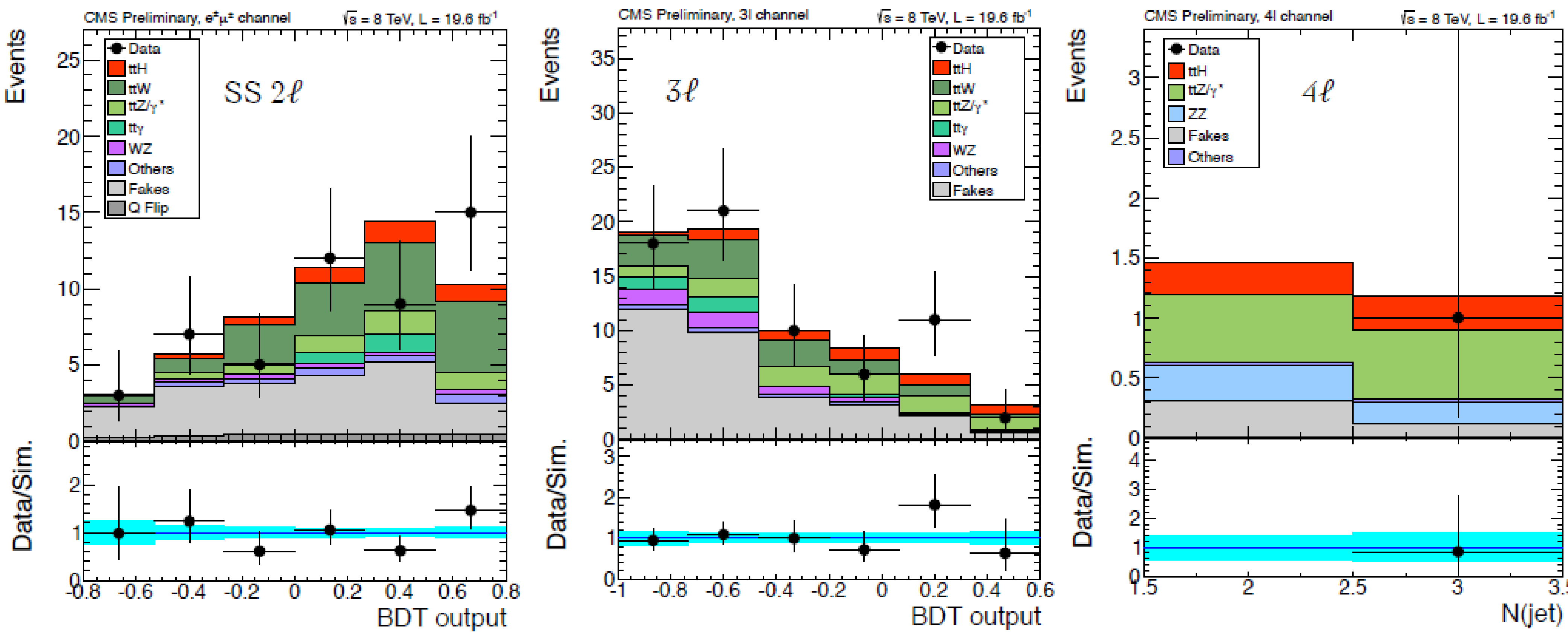
ttH(γγ) – Run 1

- Analysis of 7 TeV (1 inclusive channel), and 8 TeV (leptonic and hadronic)



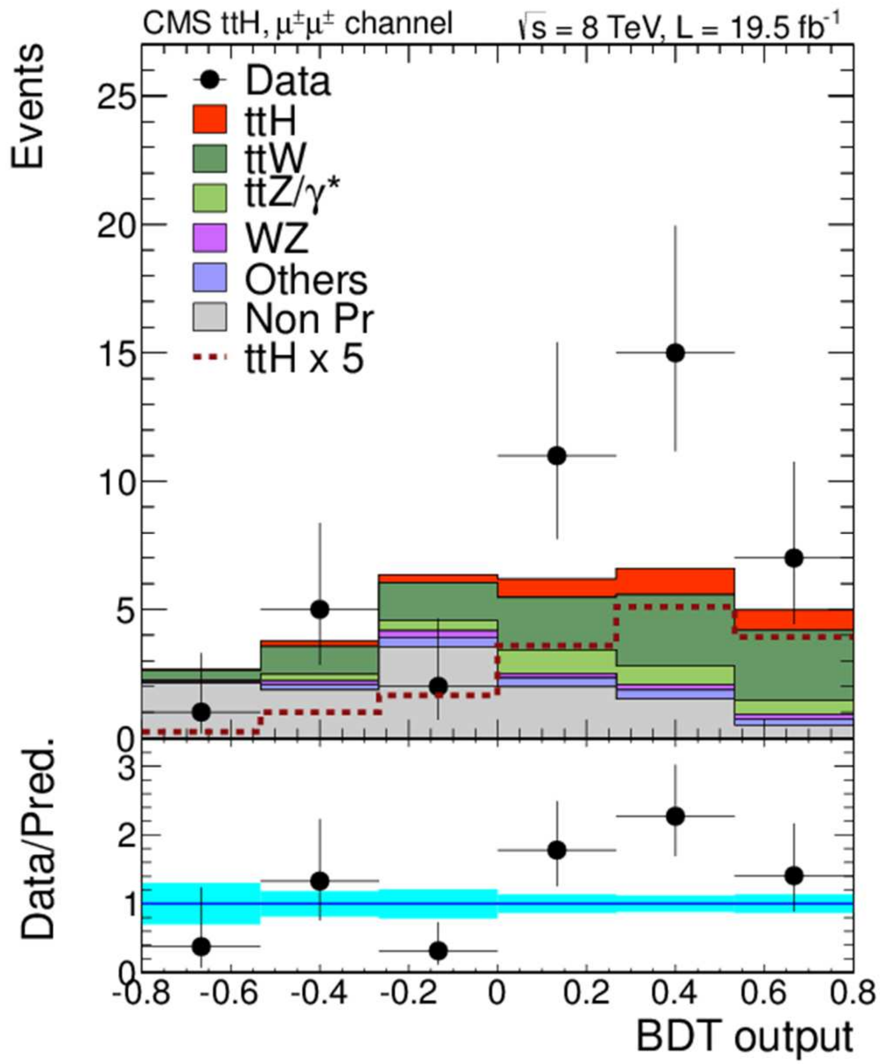
ttH(multileptons) – Run 1

- Categorise by 2, 3, 4 leptons
  - Sub-categories for signal-like and background-like selections
- BDT in 2 and 3 lepton categories, jet multiplicity in 4 lepton category



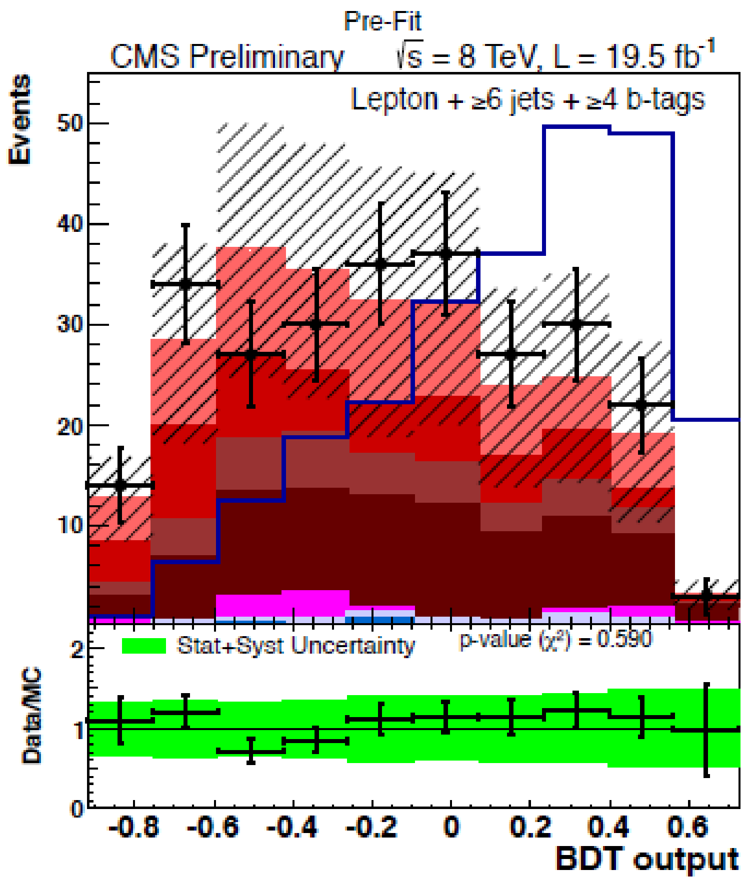
ttH(multileptons) – Run 1

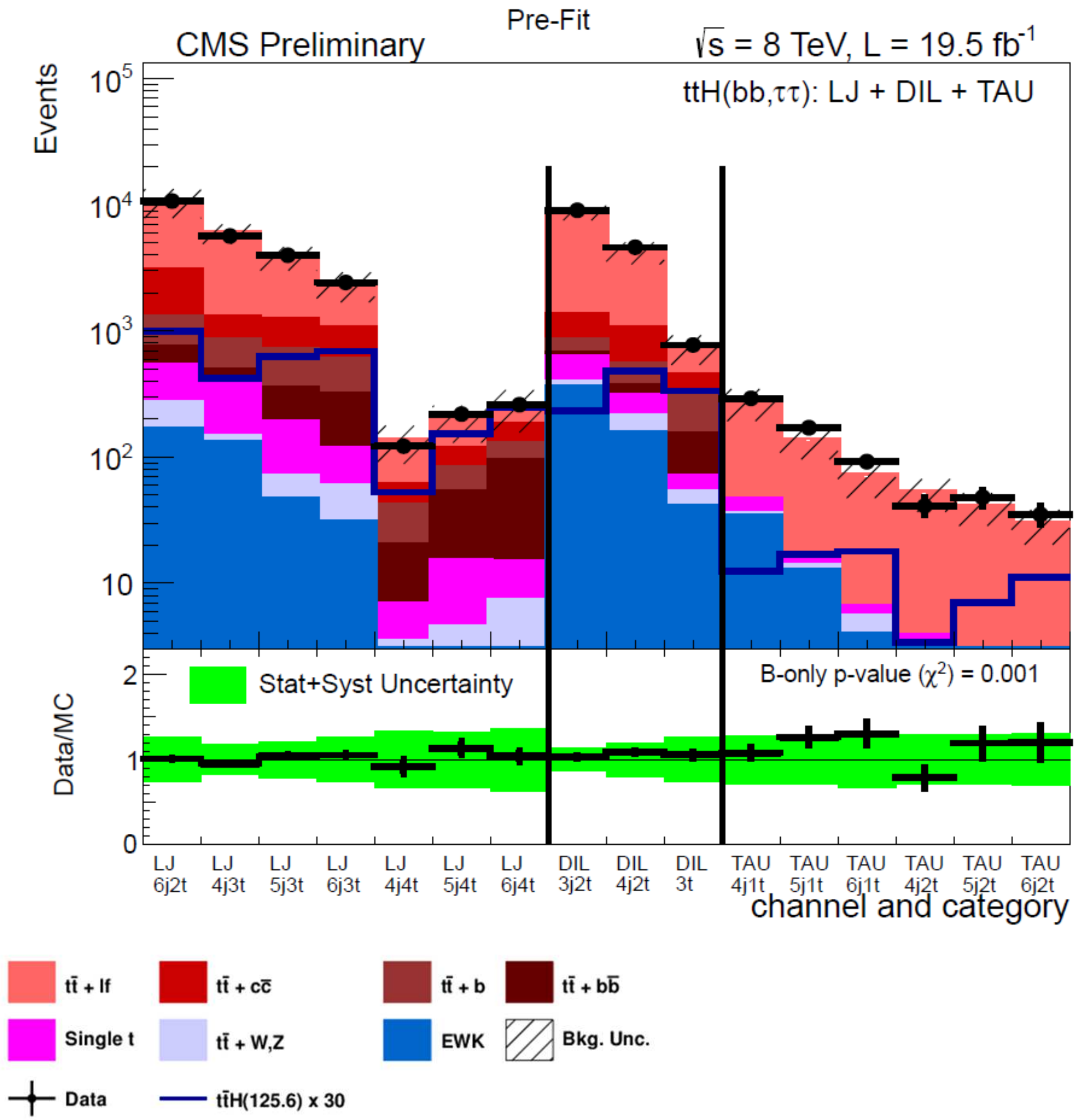
- Mild signal excess in same-sign dimuon channel



ttH(bb) and ttH($\tau_{had}\tau_{had}$) – Run 1

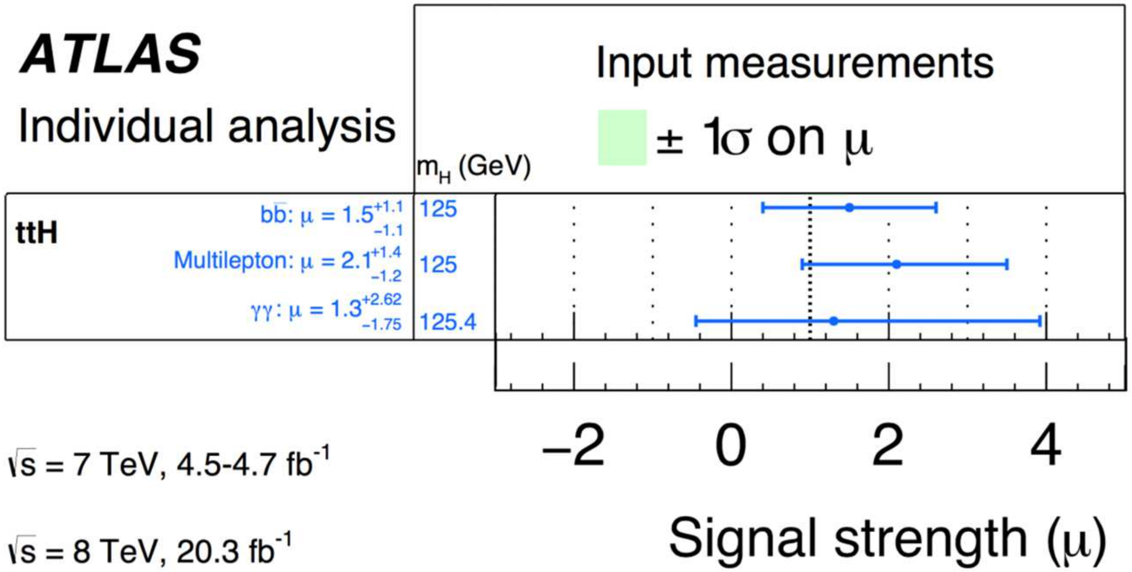
- Analysis channels dilepton, l+jets, hadronic τ's
- Categorise by (# jets, # b-tags)
- BDT analysis, optimised variables in each category





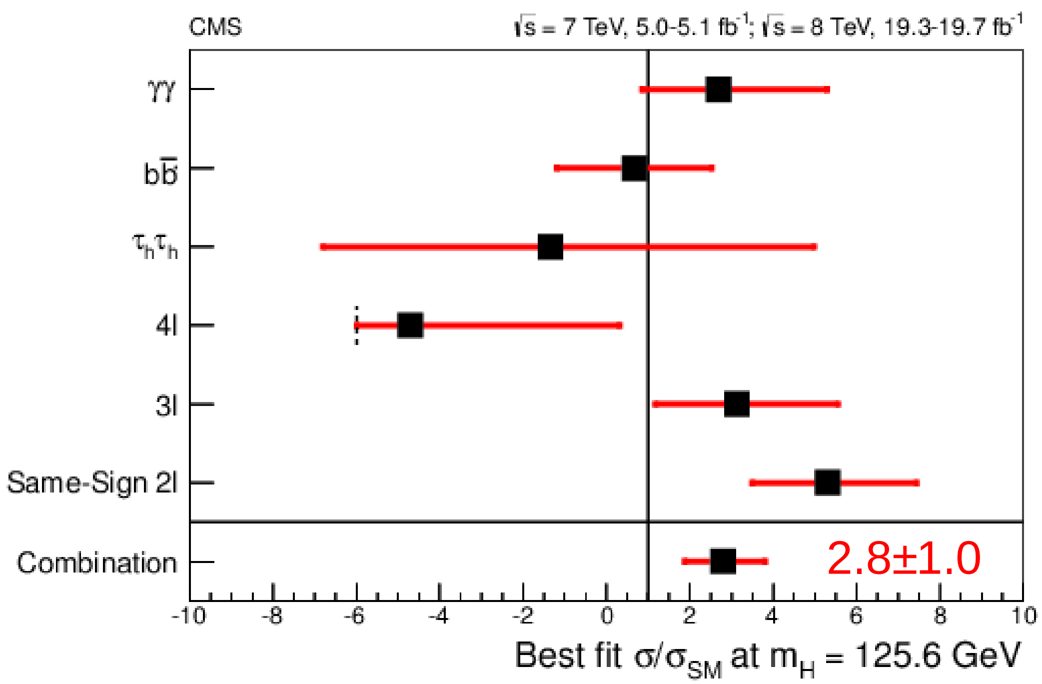
ttH – Results from Run 1

- Combine all orthogonal channels for best fit of SM ttH cross section



Observed (expected) significance
2.5σ (1.5σ)

EPJC (2016) 76:6



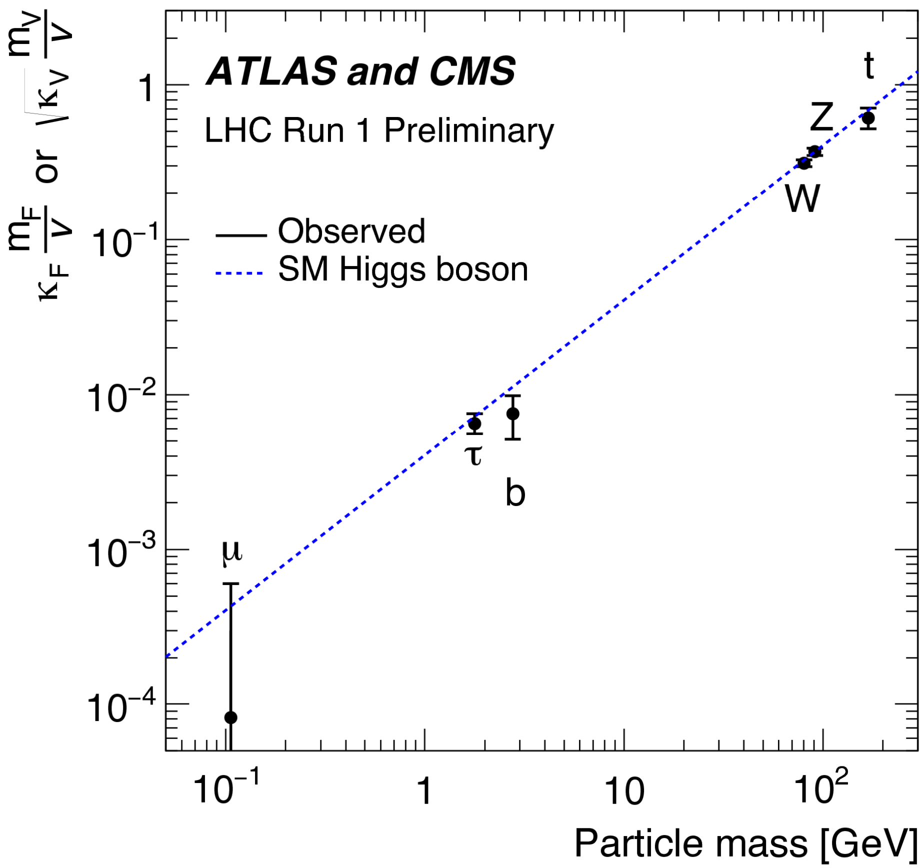
Observed (expected) significance
3.4σ (1.2σ)

equivalent to 2σ upward fluctuation compared to SM ttH

JHEP 09 (2014) 087

ATLAS+CMS Combination – Run 1

> Coupling strengths



ttH(bb) with Matrix Element Method – Run 1

- Analysis channels dilepton, l+jets
  - Categorise by tt+hf, tt+lf using likelihood from b-tag discriminator values → Low/high purity categories

$$\mathcal{F}(\xi) = \frac{f(\xi|t\bar{t}+\mathrm{hf})}{f(\xi|t\bar{t}+\mathrm{hf}) + f(\xi|t\bar{t}+\mathrm{lf})}$$

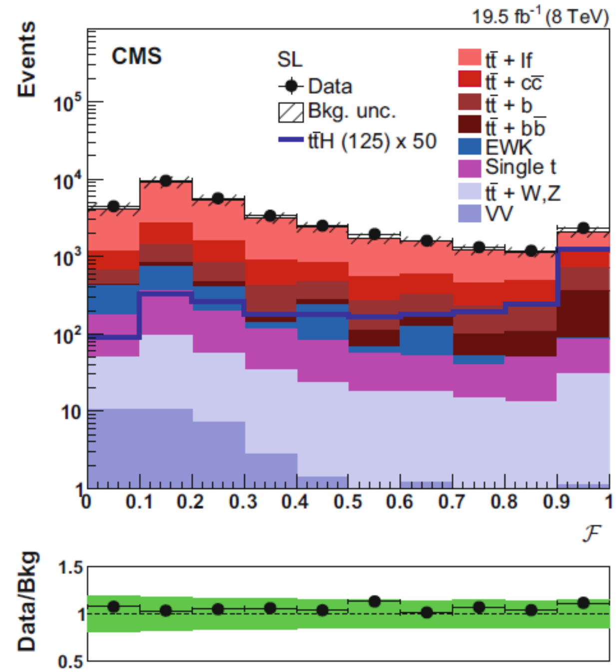
  Sub-categories by number of jets (in l+jets)

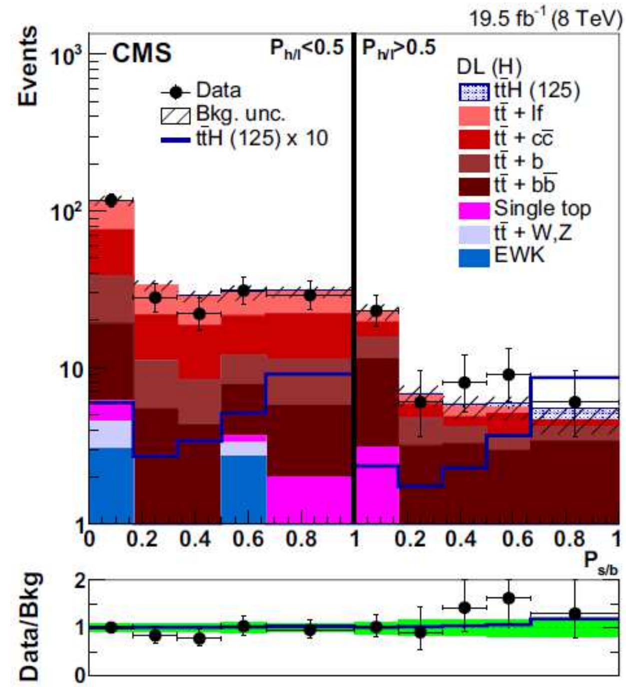
- MEM analysis with tt+bb as background hypothesis
  - Certain hypotheses for given number of jets
- Combination of 2 discriminants
  - Probability of ttH-like topology

$$P_{\mathrm{s/b}} = \frac{w(\mathbf{y}|t\bar{t}\mathrm{H})}{w(\mathbf{y}|t\bar{t}\mathrm{H}) + k_{\mathrm{s/b}}\, w(\mathbf{y}|t\bar{t}+b\bar{b})}$$

  - Probability of high b-jet multiplicity

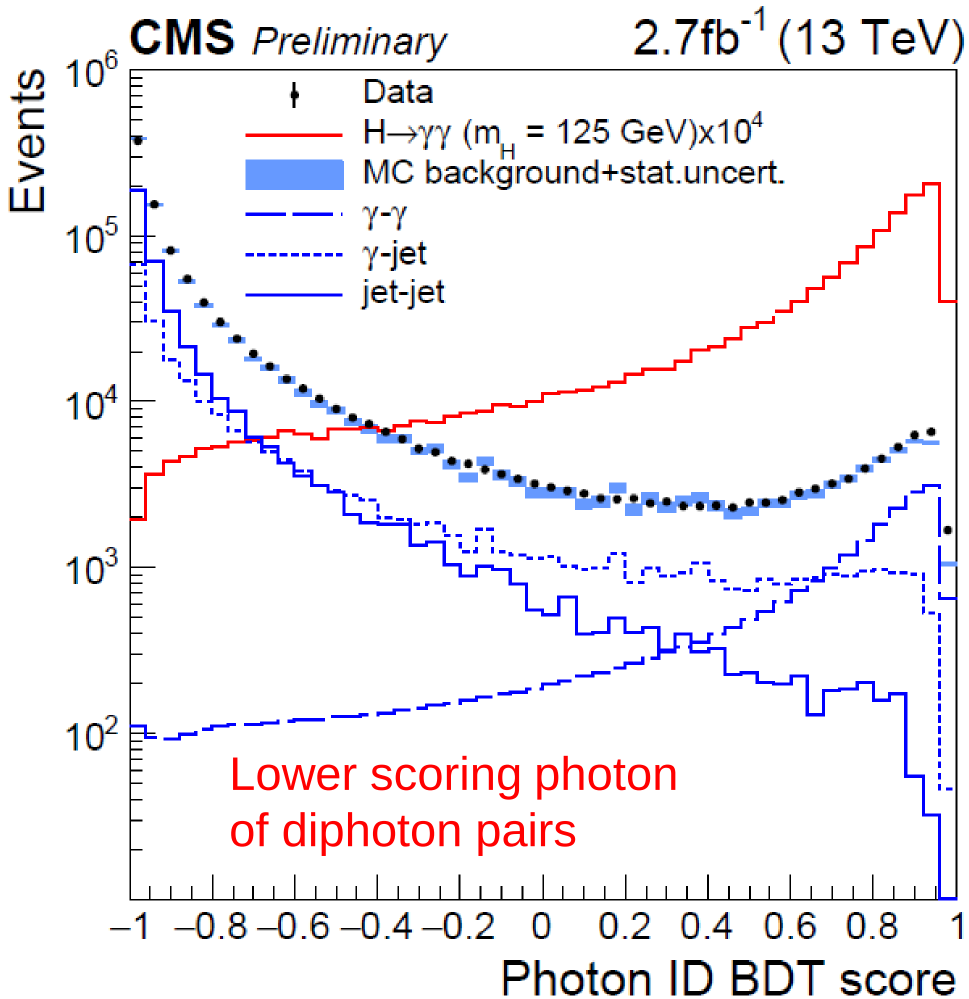
$$P_{\mathrm{h/l}} = \frac{f(\xi|t\bar{t}+\mathrm{hf})}{f(\xi|t\bar{t}+\mathrm{hf}) + k_{\mathrm{h/l}}\, f(\xi|t\bar{t}+\mathrm{lf})}$$





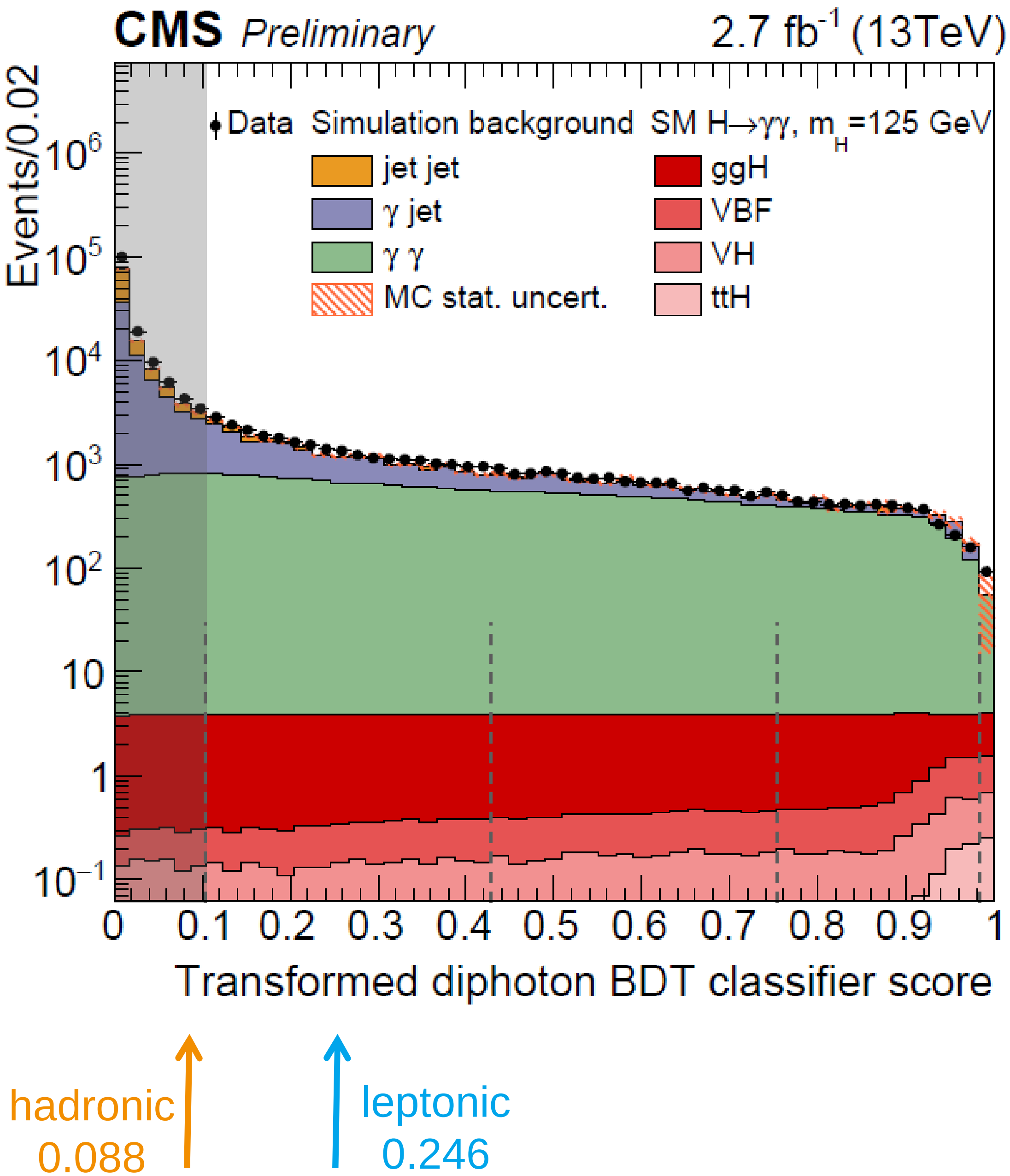
ttH(γγ) – Challenges

- Require excellent diphoton mass resolution, suppression of fake photons and backgrounds
  - Good photon reconstruction and energy calibration
  - Vertex association
  - Photon ID via BDT
  - Diphoton classifier via BDT



ttH(γγ) – Diphoton BDT

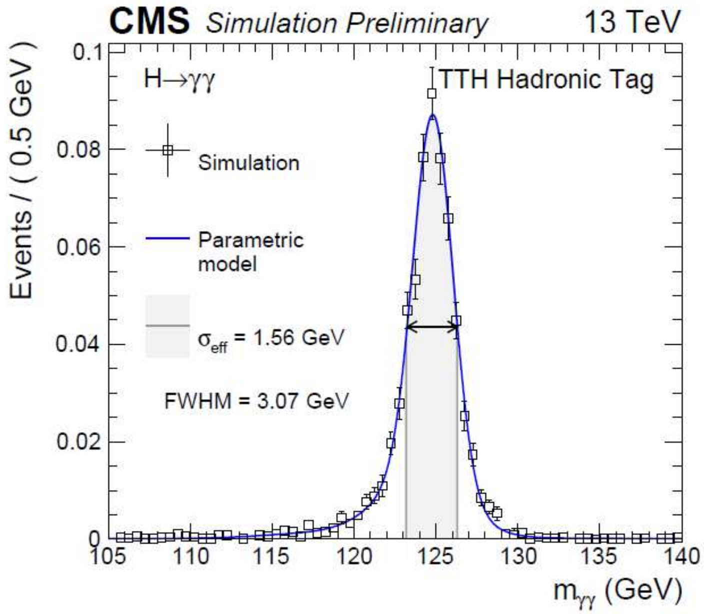
- Classify for
  - Signal-like kinematic characteristics
  - Good diphoton mass resolution events
  - Photon-like values from photon identification BDT
- Should be mass independent



ttH(γγ) – Signal and Background Model

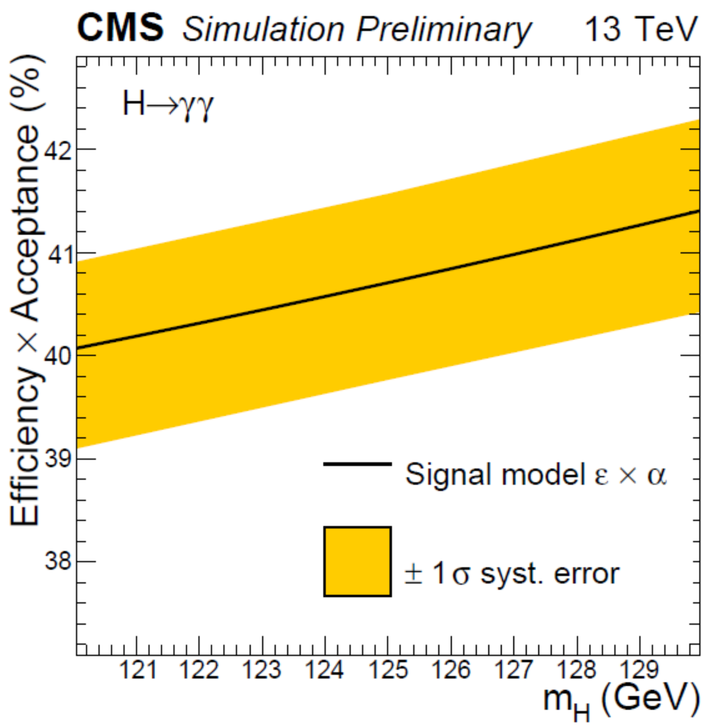
> Signal

- Simulated mass points 120, 125, 130 GeV
- Fit distribution of Higgs mass with parametric model, including systematic variations
- Sum of up to 4 Gaussians
- Continuous interpolation for any mass point
- Normalisation from linear interpolation of efficiency x acceptance



> Background

- Consider large set of candidate function families
- Treat choice of function as discrete parameter in likelihood fit
- Exclude low and high order functions
- Add penalty to account for number of floating parameters

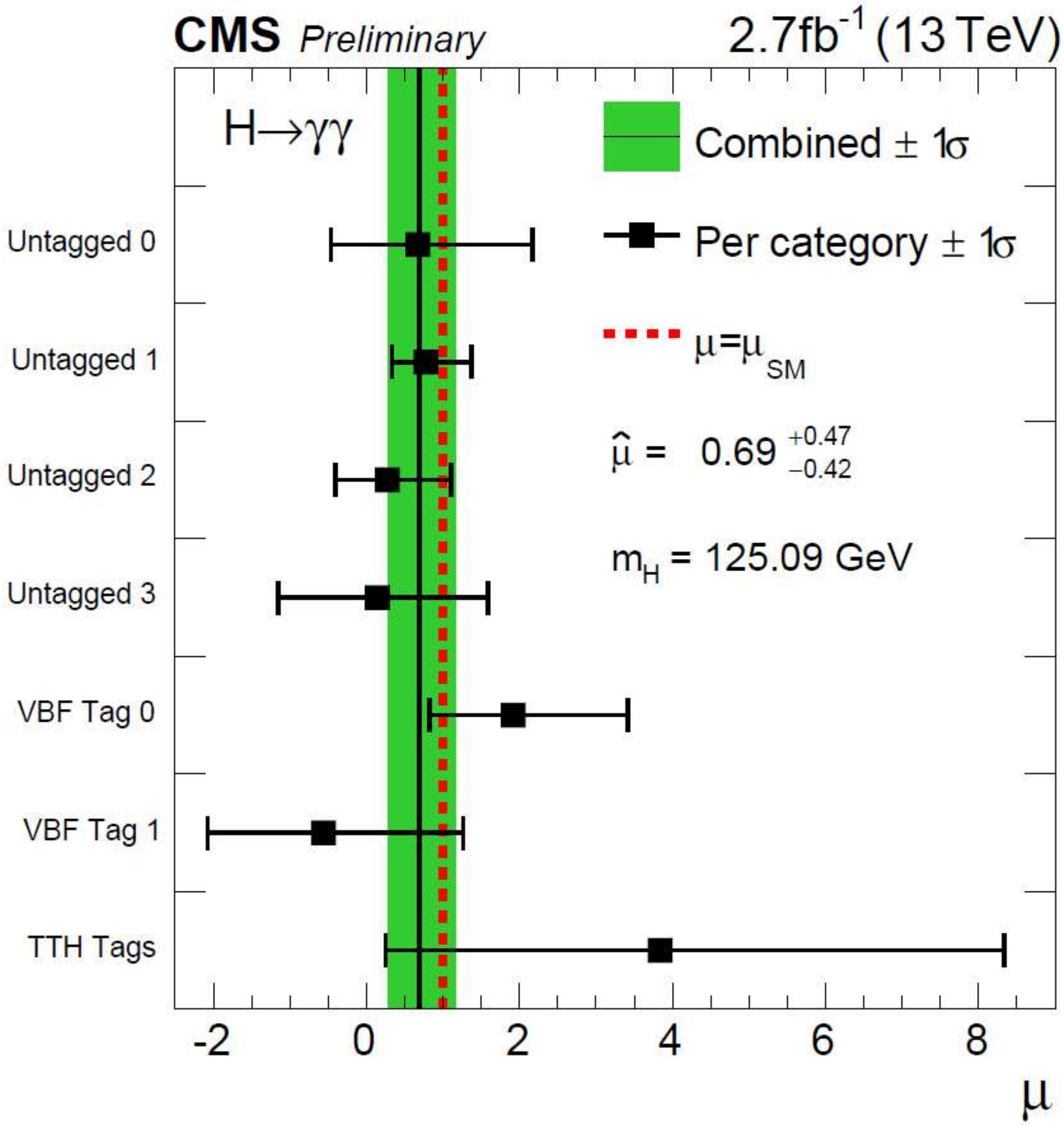


ttH(γγ) – Event Yields

Event Categories	SM 125 GeV Higgs boson expected signal yield								Bkg (GeV^{-1})
	Total	ggH	VBF	WH	ZH	$t\bar{t}$H	σ_{eff} (GeV)	σ_{HM} (GeV)	
Untagged 0	2.08	**76.19** %	10.06 %	7.45 %	3.98 %	2.32 %	1.25	1.17	0.93
Untagged 1	30.44	**86.24** %	7.13 %	3.73 %	2.12 %	0.79 %	1.41	1.22	61.19
Untagged 2	43.36	**91.16** %	4.80 %	2.39 %	1.29 %	0.36 %	1.86	1.50	165.52
Untagged 3	42.18	**92.18** %	4.21 %	2.05 %	1.16 %	0.40 %	2.63	2.20	350.94
VBF Tag 0	3.00	35.28 %	**63.48** %	0.68 %	0.19 %	0.36 %	1.61	1.24	1.57
VBF Tag 1	4.08	53.14 %	**43.62** %	1.69 %	0.85 %	0.69 %	1.77	1.35	6.85
TTH Hadronic Tag	0.64	8.76 %	0.41 %	1.66 %	2.10 %	**87.06** %	1.56	1.31	0.90
TTH Leptonic Tag	0.23	0.14 %	0.09 %	2.91 %	1.31 %	**95.55** %	1.73	1.56	0.03
Total	126.00	86.92 %	7.87 %	2.62 %	1.45 %	1.14 %	1.94	1.49	587.92

H→γγ Combination

- Combination of all orthogonal analysis channels



ttH(multilepton) – Lepton Fake Rate

- Lepton MVA
  - ID, kinematics, isolation, impact parameter, lepton-jet relations

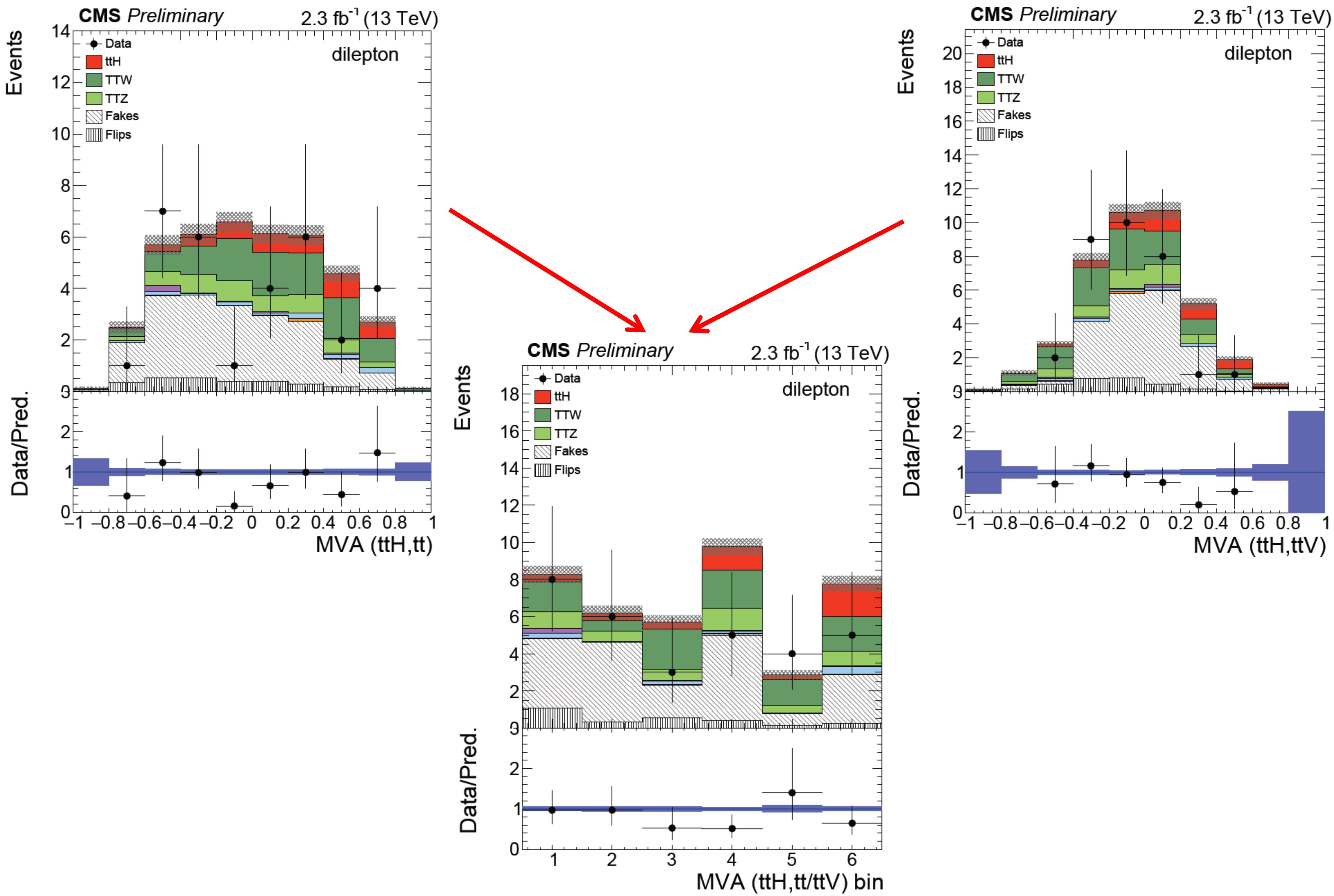
- Background fake leptons (jet mis-identification, heavy flavour decays)
  - Control region inverting MVA ID requirement
  - Apply transfer factor: probability for fake lepton to pass ID
  - Fake rate measured (high-pt): QCD events triggered by single lepton paths
  - Fake rate measured (low-pt): inclusive QCD events (μ), Z+l events (e)

- Charge mis-assignment of electrons from m_{ee} in SS and OS lepton pairs

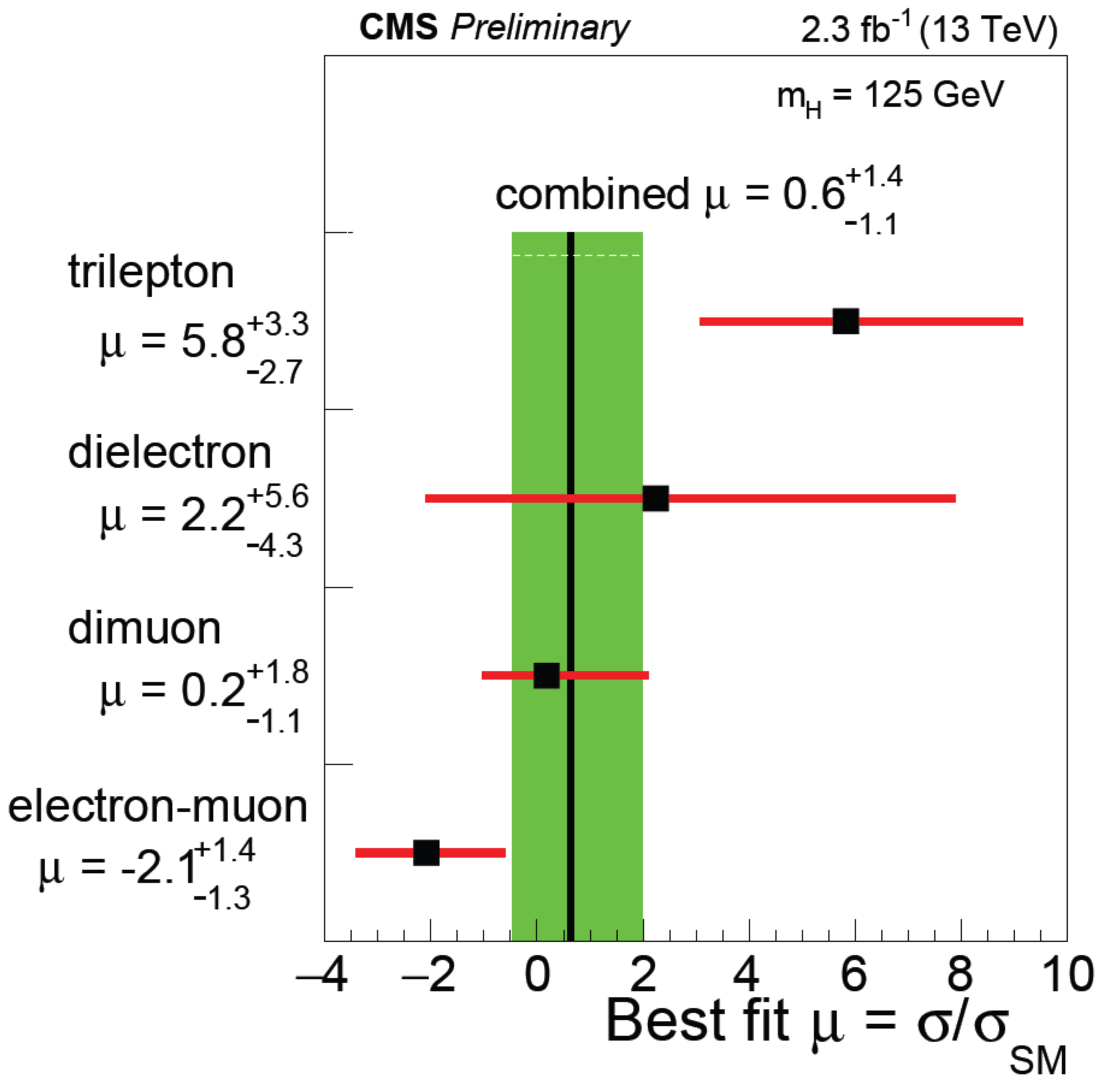
ttH(multilepton) – Event Yields

	$\mu\mu$	ee	$e\mu$	3ℓ
$t\bar{t}H$	1.53 ± 0.08	0.69 ± 0.05	2.27 ± 0.10	2.12 ± 0.09
$t\bar{t}W$	3.22 ± 0.16	1.47 ± 0.11	4.95 ± 0.19	2.56 ± 0.14
$t\bar{t}Z/\gamma^*$	0.82 ± 0.03	1.14 ± 0.14	2.42 ± 0.17	3.75 ± 0.18
WZ	0.09 ± 0.05	0.06 ± 0.06	0.25 ± 0.11	0.33 ± 0.11
tttt	0.19 ± 0.03	0.11 ± 0.02	0.28 ± 0.03	0.22 ± 0.03
tZq	0.10 ± 0.06	0.00 ± 0.00	0.12 ± 0.13	0.44 ± 0.17
rare SM bkg.	0.06 ± 0.03	0.04 ± 0.04	0.13 ± 0.06	0.16 ± 0.59
non-prompt (data)	3.99 ± 0.38	3.58 ± 0.38	10.10 ± 0.65	8.08 ± 0.67
charge mis-ID (data)		1.11 ± 0.05	1.65 ± 0.05	
signal	1.53 ± 0.08	0.69 ± 0.05	2.27 ± 0.10	2.12 ± 0.09
all backgrounds	8.47 ± 0.42	7.52 ± 0.44	19.90 ± 0.73	15.55 ± 0.95
data	9	11	11	28

ttH(multilepton) – Signal Extraction in 2 SS Leptons

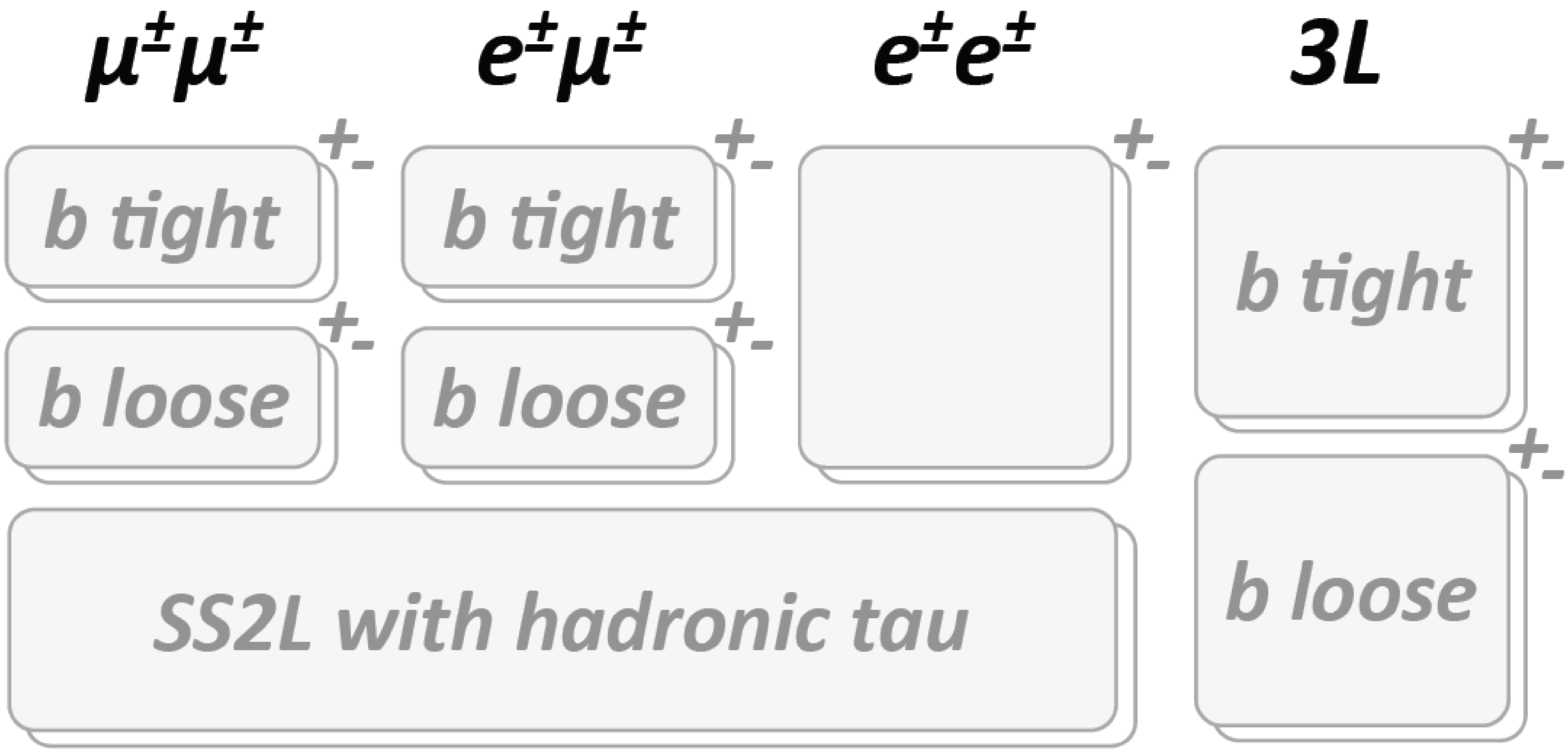


ttH(multilepton) – Results split by Flavour



ttH(multilepton) Categories

- 16 sub-categories – increase sensitivity due to different S+B composition
  - Lepton flavour: different background compositions, and fake contributions (charge flips only in electrons)
  - Lepton charge: Charge asymmetry of several backgrounds
  - Presence of 2 b-tags: Non-tt backgrounds
  - Presence of hadronic τ: ttH($\tau\tau$) with low backgrounds



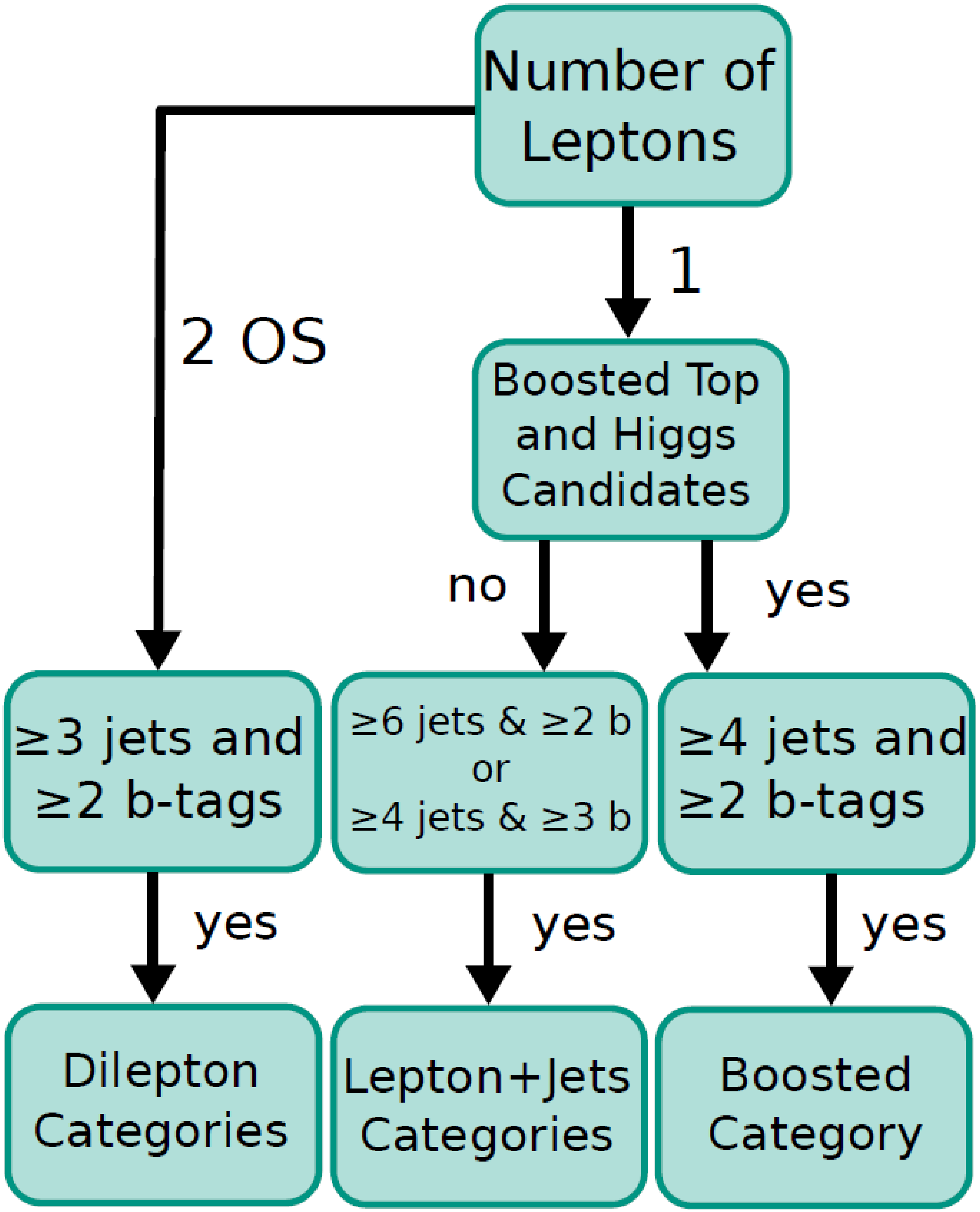
Definition of tt+xx Processes in ttH(bb)

- Split inclusive tt+jets based on heavy-flavour content of additional jets
  - Presence of ghost b/c hadron clustered to generator jet
  - Additional jets defined by p_T>20 GeV, |η|<2.4
- Processes: ttbb, ttb, tt2b, ttcc, tt+lf
  - ttbb and ttb in principle same process, well separated jets
    → Can be treated perturbatively
  - tt2b theoretically and experimentally different, collinear gluon splitting
    → Mainly from parton shower,
    needs (arbitrary) cut-off, matter of tuning

  - ttcc inclusive for all processes with at least one additional c jet
    Similar issues as for b jets, but less relevant background
  - tt+lf: events without additional heavy-flavour jet

ttH(bb) Categories

- 13 categories
  - 5 dilepton
  - 7 l+jets
  - 1 boosted (in l+jets)



ttH(bb) – Event Yields (l+jets)

Process	≥ 6 jets, 2 b-tags	4 jets, 3 b-tags	5 jets, 3 b-tags	≥ 6 jets, 3 b-tags
$t\bar{t}$+lf	5359.3 ± 1226.3	2026.1 ± 651.4	1000.2 ± 352.9	589.5 ± 199.7
$t\bar{t} + c\bar{c}$	1722.2 ± 849.5	363.2 ± 190.9	368.1 ± 191.3	396.6 ± 209.5
$t\bar{t}$+b	393.7 ± 188.2	203.1 ± 92.5	199.6 ± 90.8	170.8 ± 81.4
$t\bar{t}$+2b	165.2 ± 81.2	78.9 ± 38.0	87.2 ± 40.7	97.3 ± 46.8
$t\bar{t} + b\bar{b}$	226.4 ± 113.2	75.8 ± 35.3	114.1 ± 52.3	183.7 ± 86.7
Single Top	283.0 ± 49.0	115.3 ± 30.8	76.2 ± 19.5	47.5 ± 12.7
V+jets	130.5 ± 35.2	38.6 ± 17.8	22.8 ± 10.4	13.6 ± 6.4
$t\bar{t}$+V	43.5 ± 8.2	4.3 ± 1.2	6.4 ± 1.8	10.0 ± 2.7
Diboson	2.8 ± 1.3	2.1 ± 1.3	0.9 ± 0.5	0.2 ± 0.3
Total bkg	8326.7 ± 1788.6	2907.4 ± 836.5	1875.5 ± 534.7	1509.1 ± 423.7
$t\bar{t}$H	29.6 ± 2.1	7.4 ± 1.0	10.9 ± 1.2	16.7 ± 2.1
Data	7185	2793	1914	1386
S/B	0.0036	0.0026	0.0059	0.011
Data/B	0.9 ± 0.2	1.0 ± 0.3	1.0 ± 0.3	0.9 ± 0.3

Process	4 jets, ≥ 4 b-tags	5 jets, ≥ 4 b-tags	≥ 6 jets, ≥ 4 b-tags	boosted
$t\bar{t}$+lf	17.8 ± 10.8	17.7 ± 10.9	17.6 ± 11.3	45.1 ± 9.4
$t\bar{t} + c\bar{c}$	11.6 ± 8.2	22.1 ± 15.4	35.9 ± 24.9	21.8 ± 12.0
$t\bar{t}$+b	8.4 ± 4.4	14.8 ± 7.7	20.0 ± 10.9	10.3 ± 5.5
$t\bar{t}$+2b	3.5 ± 1.9	6.9 ± 3.7	12.3 ± 6.9	12.3 ± 6.6
$t\bar{t} + b\bar{b}$	10.1 ± 4.9	28.8 ± 13.9	73.4 ± 36.6	17.0 ± 8.4
Single Top	2.5 ± 1.1	4.3 ± 1.4	5.5 ± 2.0	7.0 ± 1.7
V+jets	1.0 ± 0.8	0.9 ± 0.8	1.4 ± 0.7	2.5 ± 0.8
$t\bar{t}$+V	0.3 ± 0.1	0.7 ± 0.3	1.6 ± 0.6	0.9 ± 0.3
Diboson	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0	0.1 ± 0.1
Total bkg	55.2 ± 23.0	96.5 ± 37.6	167.6 ± 65.7	117.0 ± 24.9
$t\bar{t}$H	0.9 ± 0.2	2.7 ± 0.6	5.9 ± 1.4	2.2 ± 0.3
Data	75	104	150	104
S/B	0.017	0.028	0.035	0.019
Data/B	1.4 ± 0.5	1.1 ± 0.4	0.9 ± 0.4	0.9 ± 0.2

ttH(bb) – Event Yields (dilepton)

	3 jets, 2 b-tags	3 jets, 3 b-tags	$\geq$ 4 jets, 2 b-tags	$\geq$ 4 jets, 3 b-tags	$\geq$ 4 jets, $\geq$ 4 b-tags
$t\bar{t}$+lf	2558.6 ± 542.7	26.6 ± 10.5	2271.6 ± 505.0	60.3 ± 25.6	0.9 ± 0.8
$t\bar{t} + c\bar{c}$	220.9 ± 103.4	22.7 ± 13.6	478.4 ± 234.4	78.4 ± 45.4	3.4 ± 2.9
$t\bar{t}$+b	65.4 ± 28.5	21.4 ± 10.2	126.2 ± 57.7	52.2 ± 25.1	2.7 ± 1.6
$t\bar{t}$+2b	16.9 ± 7.6	6.6 ± 3.1	42.9 ± 20.2	22.3 ± 10.7	1.2 ± 0.7
$t\bar{t} + b\bar{b}$	8.6 ± 4.2	3.6 ± 1.8	48.9 ± 23.7	39.8 ± 18.8	13.4 ± 7.1
Single Top	93.2 ± 16.7	3.0 ± 1.0	87.6 ± 15.8	7.3 ± 2.5	0.4 ± 0.4
V+jets	14.5 ± 11.0	1.3 ± 0.8	16.0 ± 7.4	0.0 ± 0.0	0.0 ± 0.0
$t\bar{t}$+V	3.6 ± 0.9	0.3 ± 0.2	16.4 ± 3.2	3.2 ± 0.9	0.5 ± 0.2
Diboson	1.7 ± 0.9	0.0 ± 0.0	1.2 ± 1.0	0.1 ± 0.0	0.0 ± 0.0
Total bkg	2983.4 ± 590.4	85.6 ± 25.6	3089.2 ± 650.6	263.6 ± 79.9	22.5 ± 9.8
$t\bar{t}$H	1.4 ± 0.2	0.4 ± 0.1	8.1 ± 1.1	3.6 ± 0.6	1.0 ± 0.3
Data	3123	115	2943	319	27
S/B	0.00047	0.0051	0.0026	0.014	0.046
Data/B	1.0 ± 0.2	1.3 ± 0.4	1.0 ± 0.2	1.2 ± 0.3	1.2 ± 0.5