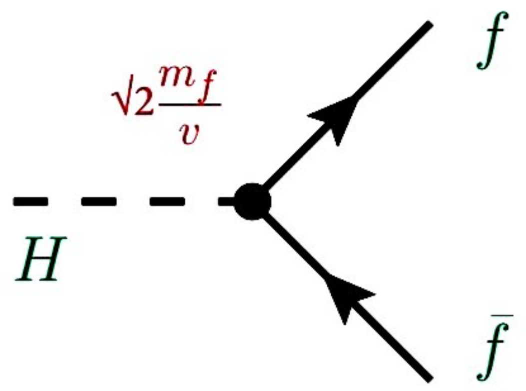
ttH measurements at LHC

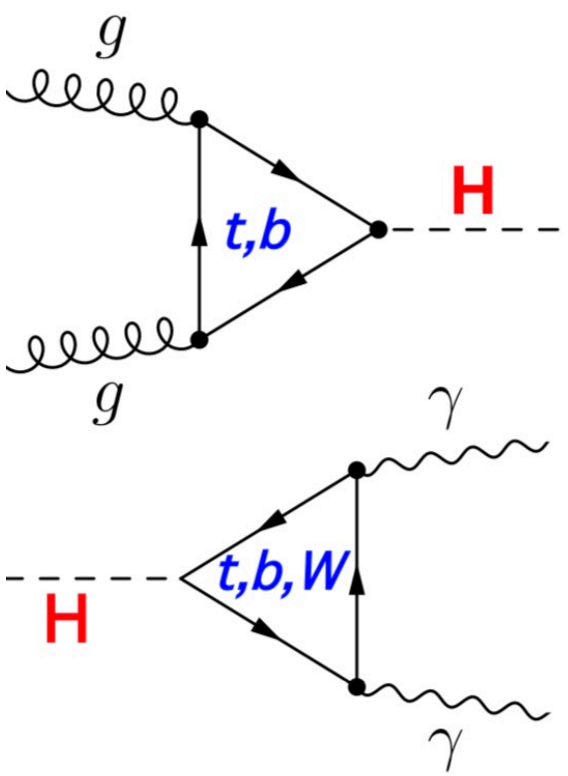
Rohin T Narayan on behalf of the ATLAS and CMS collaborations

Oct-25-2018, LCWS
Arlington

Motivation



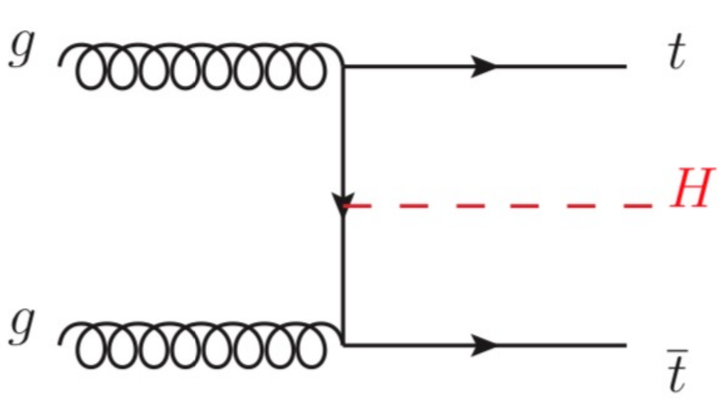
- Fermion masses are generated through Yukawa interaction
- Heaviest SM particle (top) expected to have largest Yukawa coupling (y_t) to the Higgs field.



- Indirect constraints on top Yukawa coupling possible through Higgs production through ggH and H->γγ decay processes.
- ATLAS+CMS Run-1 combination of ratio of measured coupling to SM expectation (κ_t)
  - Assuming no new particles in the loop

$$\kappa_t \equiv \frac{y_t}{y_t(SM)} = 0.87 \pm 0.15$$

JHEP08(2016)045



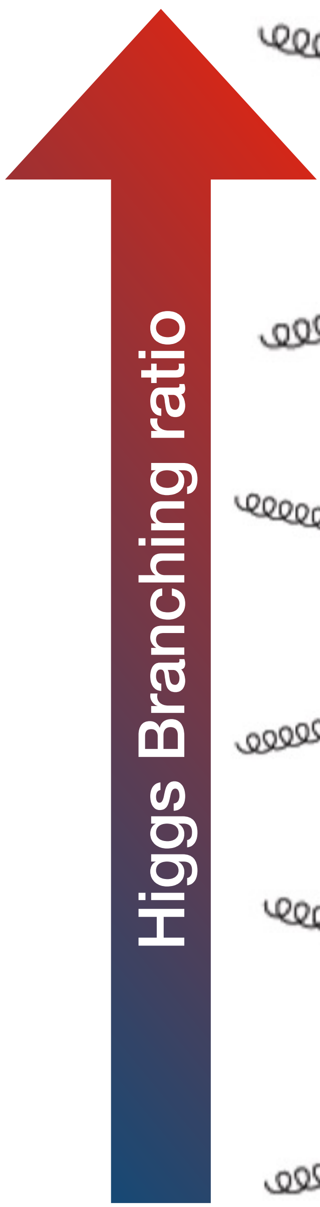
- Direct measurement of top-Yukawa coupling needed to disentangle any new physics effects.
- ttH/tH production cross-section measurement is the only direct way to measure y_t
  - tH cross-section too low

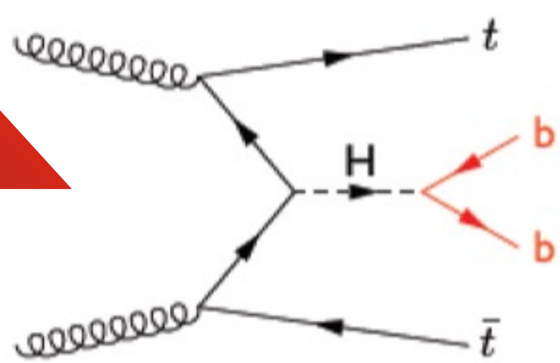
Experimental challenges

- Standard Model production cross section: ~507 fb: About 1% of total Higgs cross-section.
  - Many final states
  - Tiny signal and large backgrounds

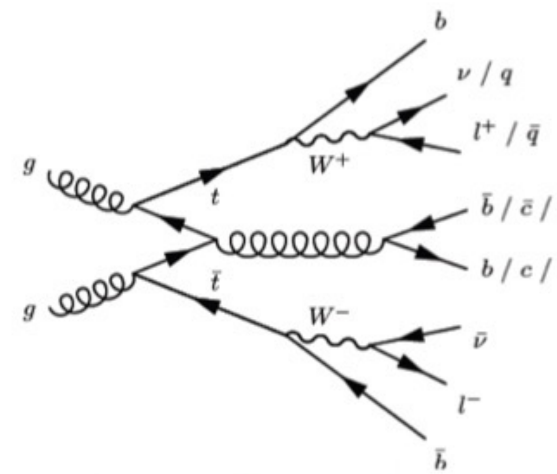
Signal

Background

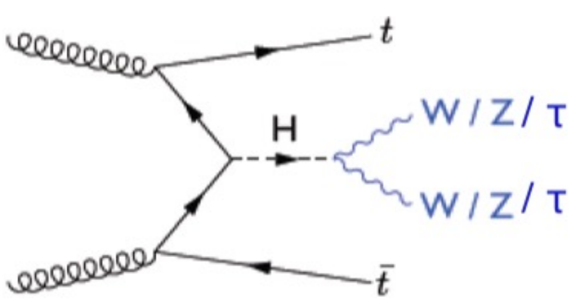




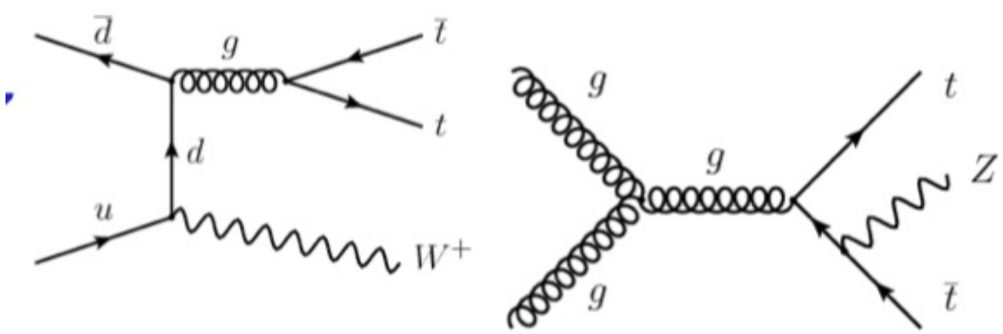
ttH(bb)



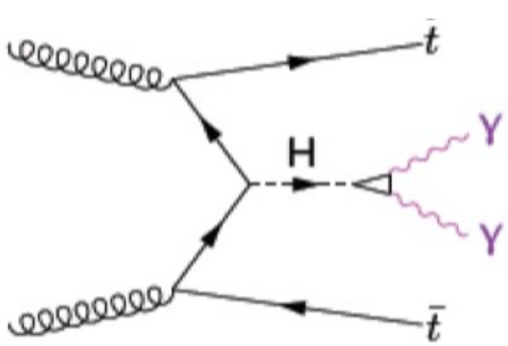
tt+(HF) jets (irreducible)



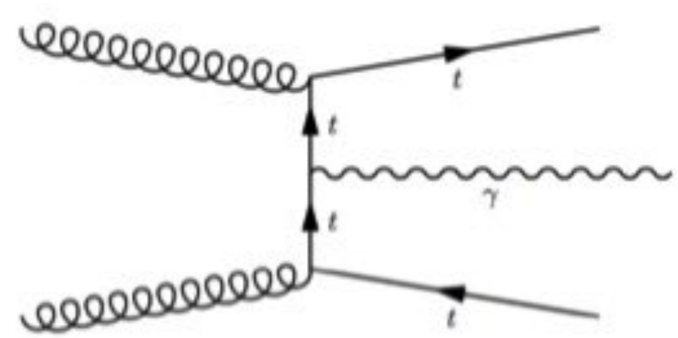
ttH-multileptons



ttV



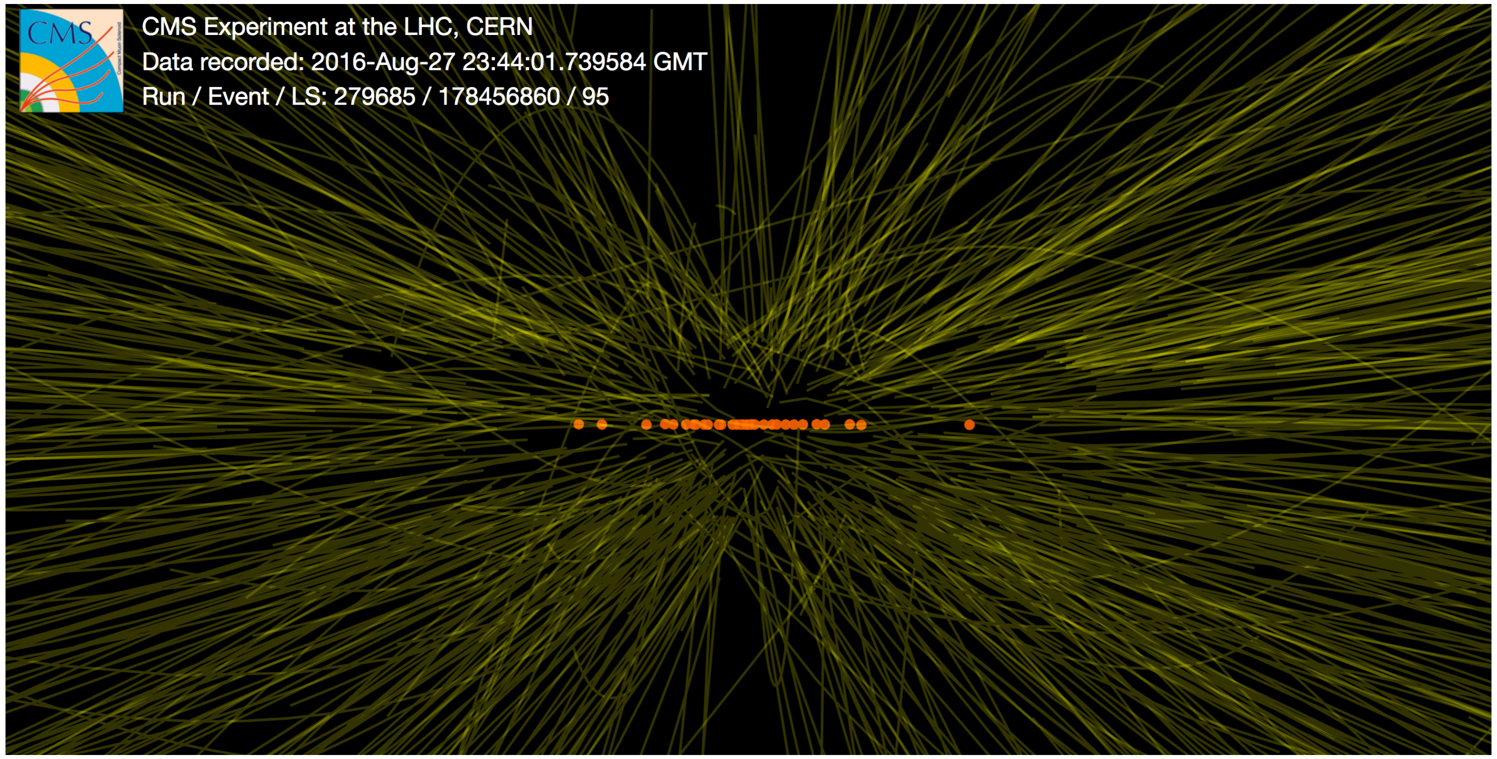
ttH(γγ)

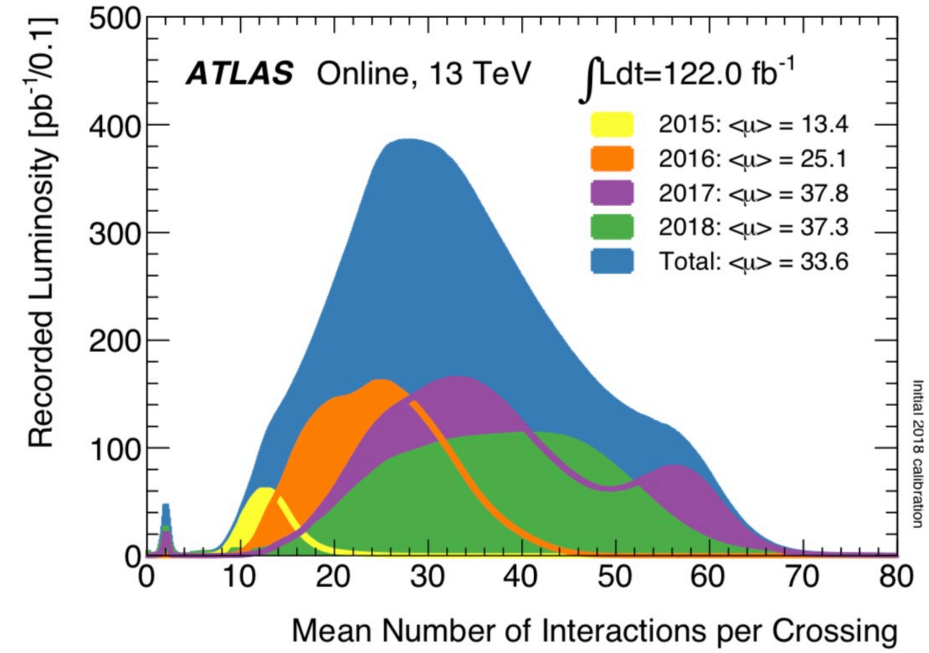


ttγ

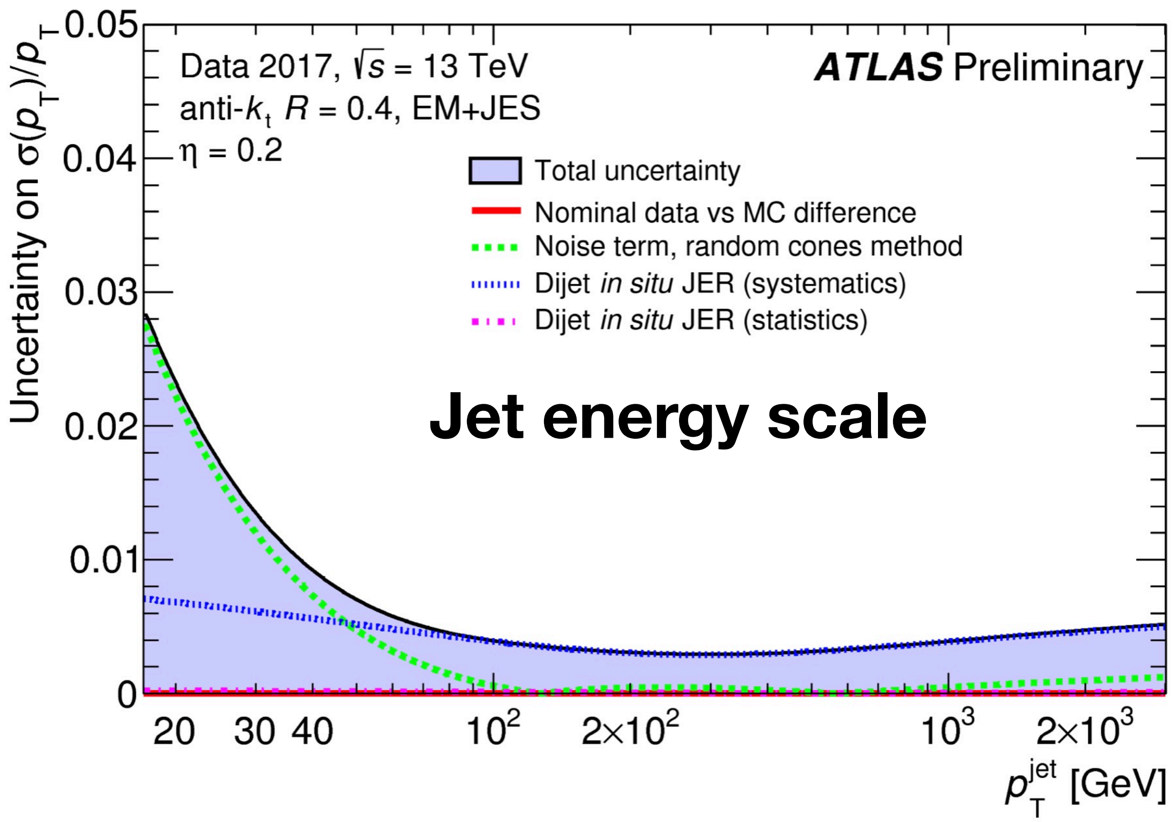
Experimental challenges

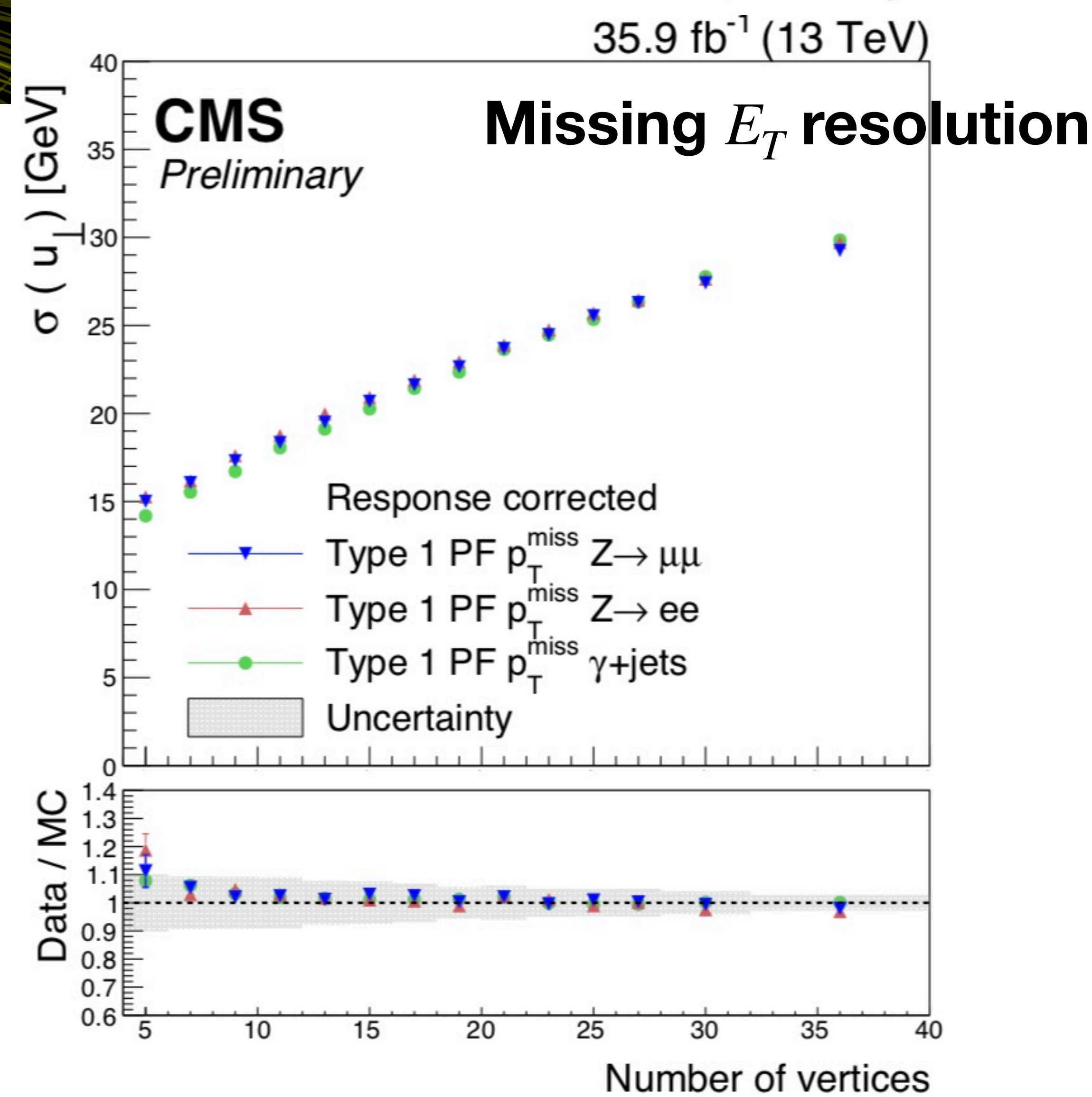
- Object identification and reconstruction among multiple proton collisions (pileup)



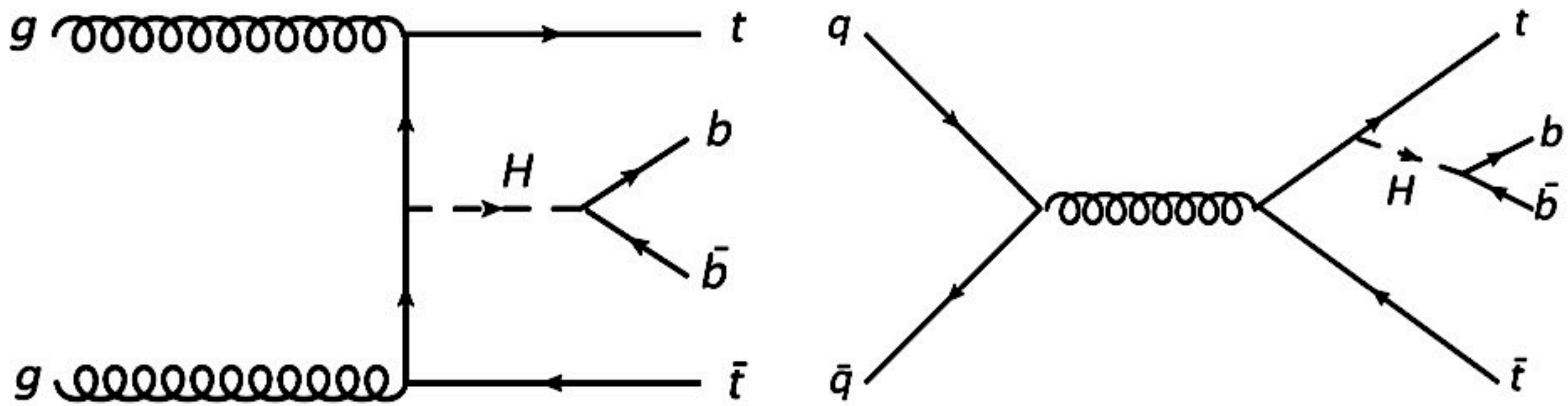


Excellent mitigation efforts





ttH(bb)



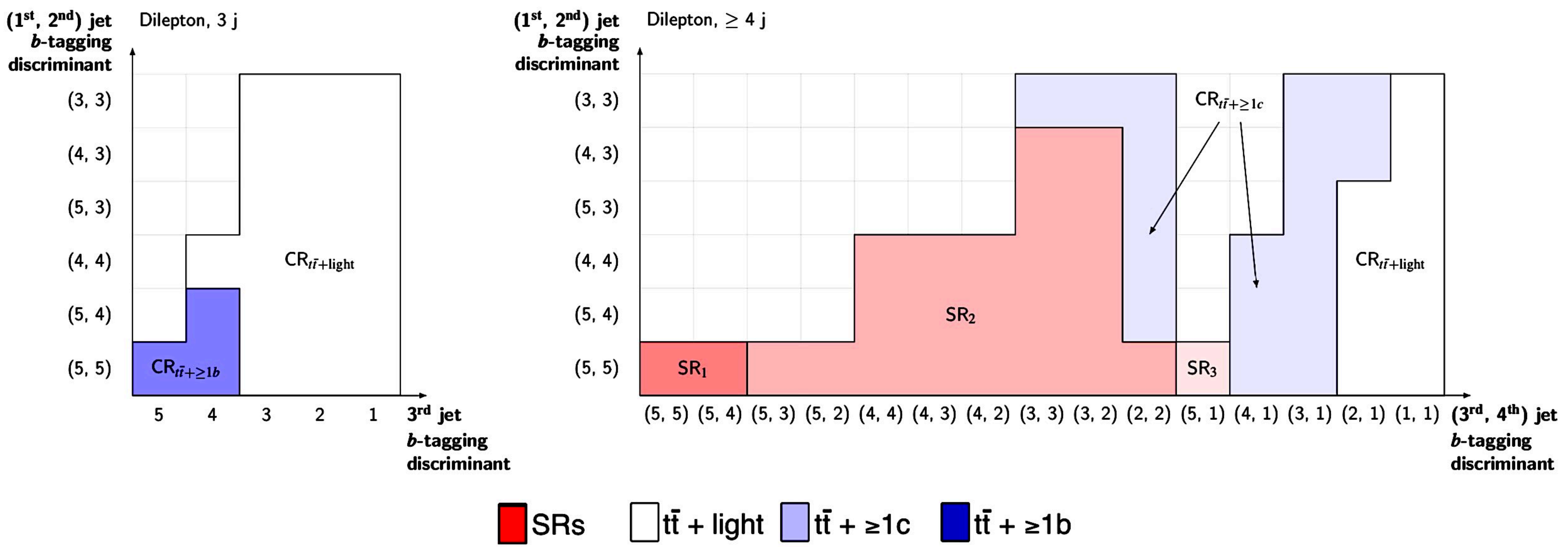
- Large $H \rightarrow b\bar{b}$ branching fraction.
- Challenges
  - Huge combinatorics in reconstructing $H \rightarrow b\bar{b}$ candidate
  - Large $tt + HF$ background + theory uncertainties
- Channels
  - Leptonic (dileptonic and single leptonic)
  - Hadronic

ttH(bb): dileptonic

- Exactly two opposite sign leptons
- Regions constructed for 3 and ≥4 jets.
    - SR's constructed with high b-tag quality
    - Three SR's: purity ranging from 1.8% to 5.4%
    - Highest signal purity in region with 3 very tight and 1 tight b-tag
- CR's for tt+b, tt+c and tt+light to constrain uncertainties, backgrounds,

	none	*loose*	*medium*	*tight*	*very-tight*
Efficiency	-	85%	77%	70%	60%
Discriminant value	1	2	3	4	5

b-tag working points



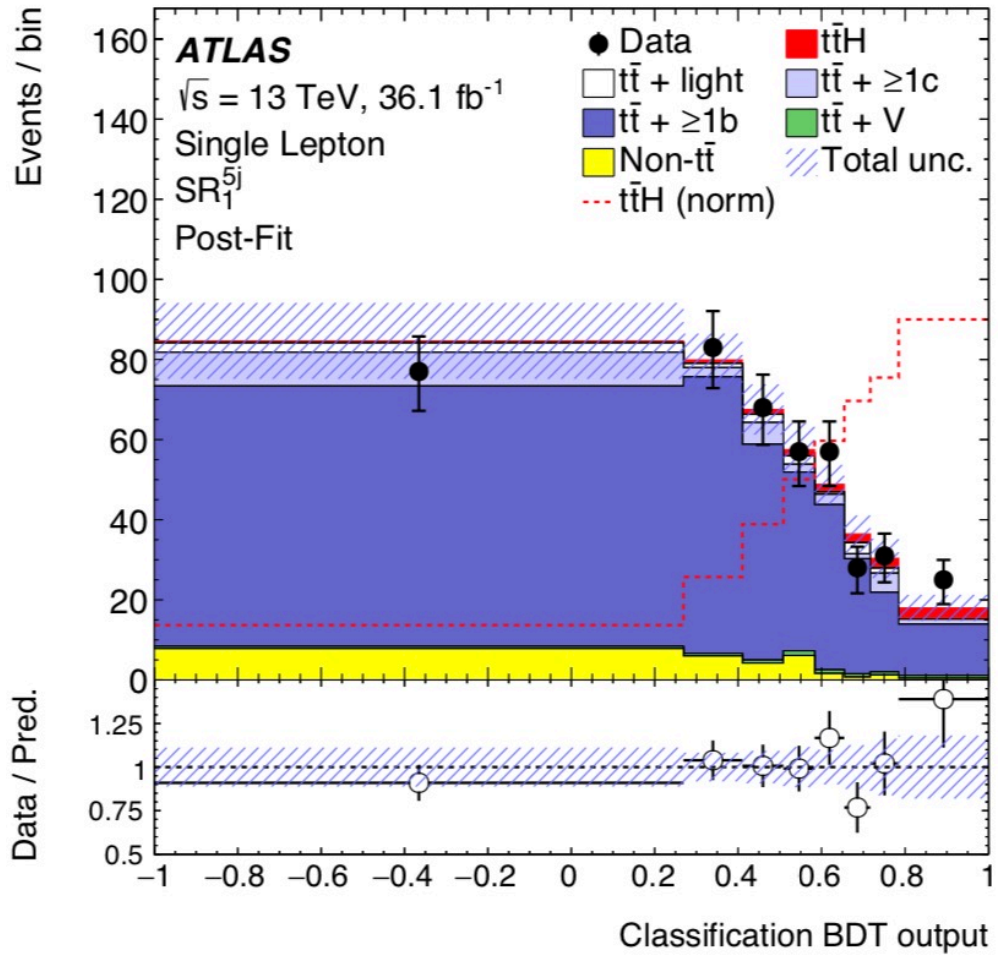
- Similar region splitting in single lepton category as well (**backup**)

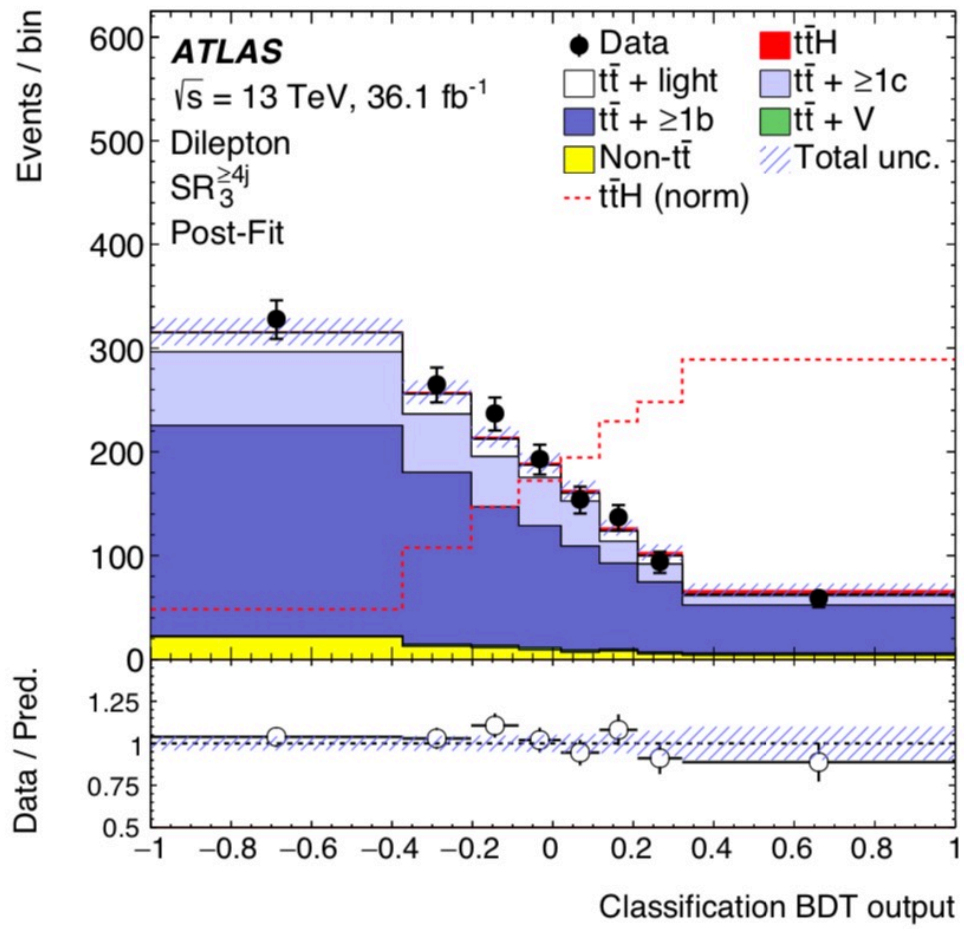
ttH(bb) multivariate analysis

Phys. Rev. D 97, 072016 (2018)

ATLAS EXPERIMENT

- In signal regions, MVA techniques are used to further separate signal and background
  - Reconstruction BDT: Attempts to reconstruct top quarks and H(bb)
  - Likelihood discriminant (LHD) : ttH signal vs tt+b background.
  - Matrix Element discriminant (MEM): ttH and tt+b PDF estimation using Matrix Element Method
- Final discriminant: Classification BDT
  - Reconstruction BDT, LHD, MEM, kinematic variables

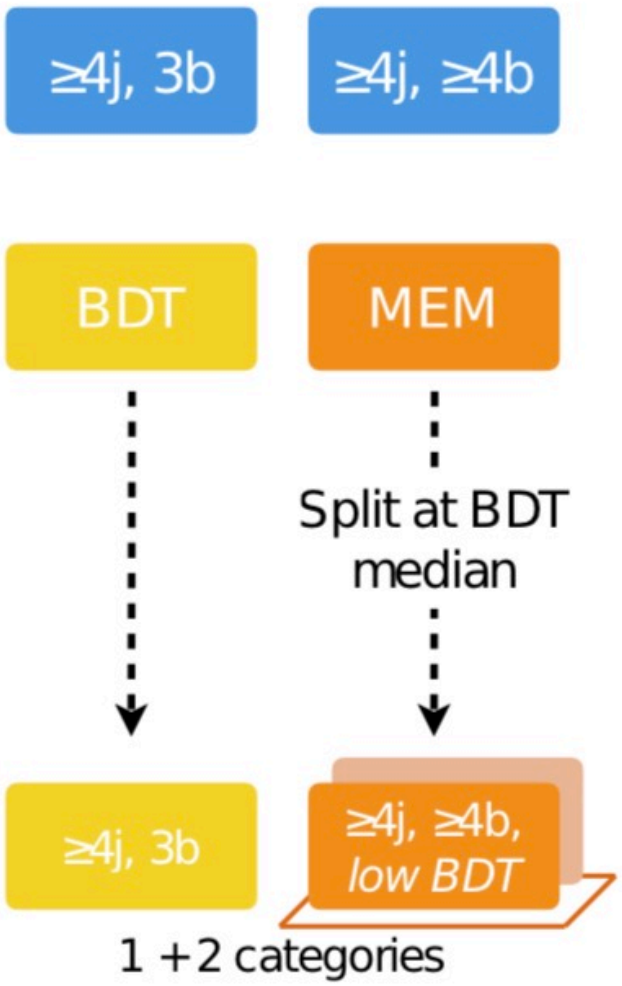


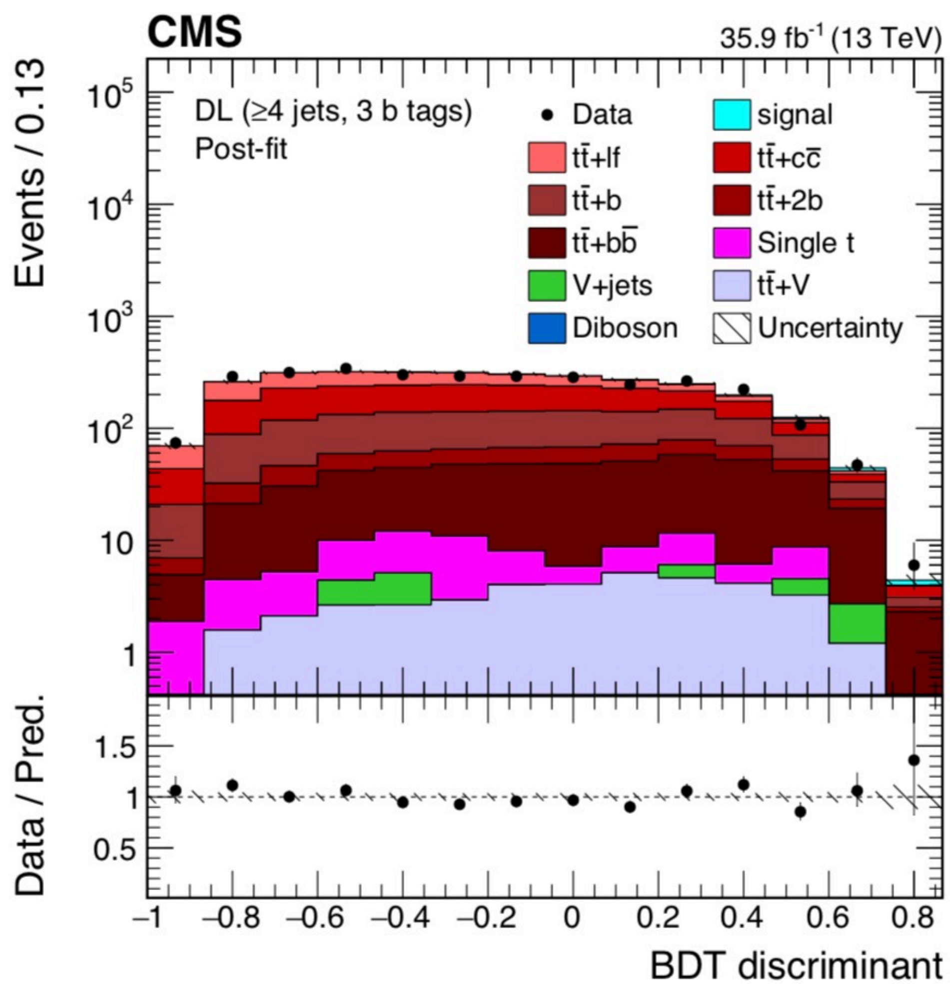


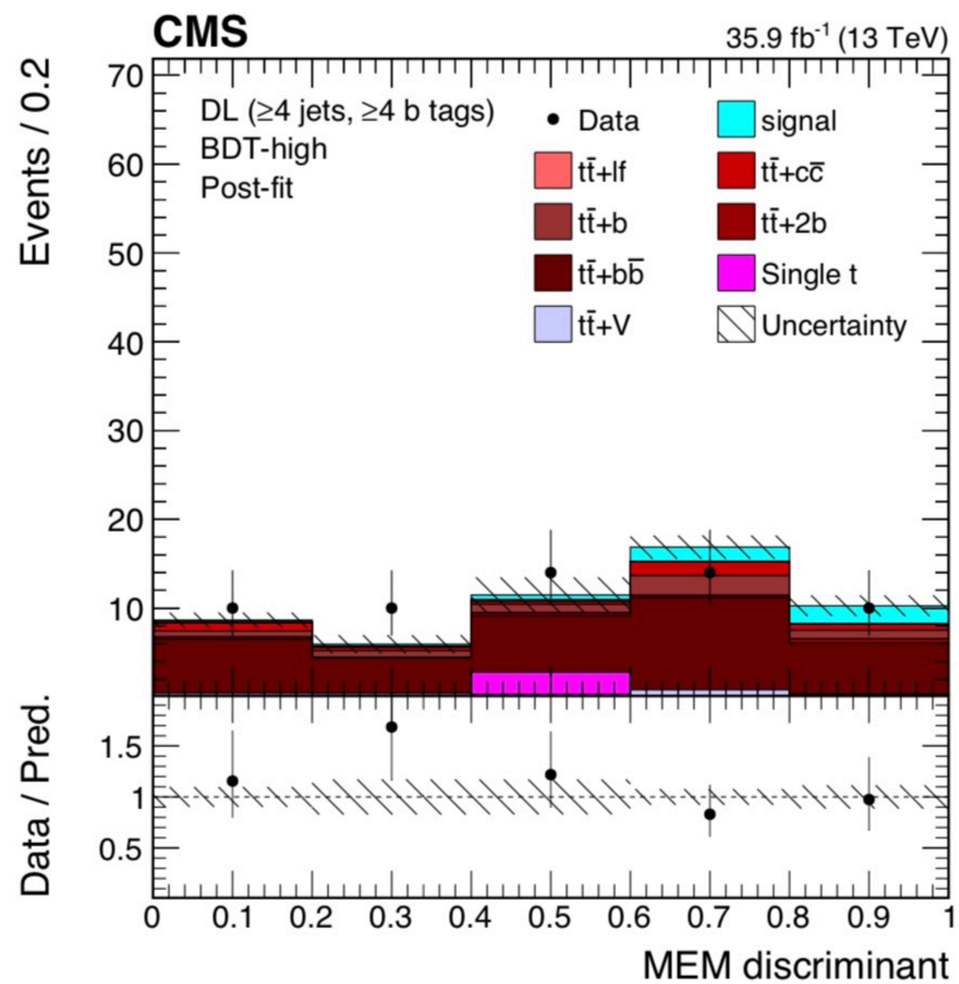
ttH(bb): Dilepton

arXiv:1804.03682

- Events categorized based on number of jets and b-tagged jets:







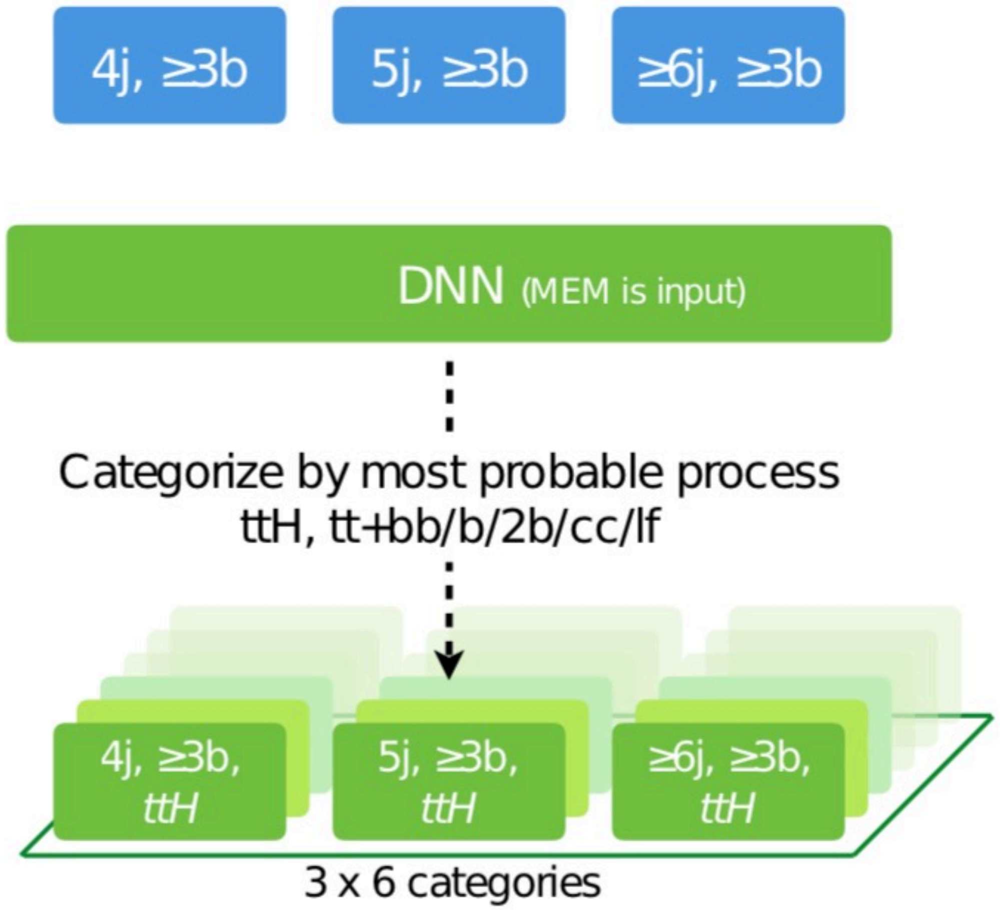
- ≥ 4j, 3b: A BDT separating signal and tt+jets background

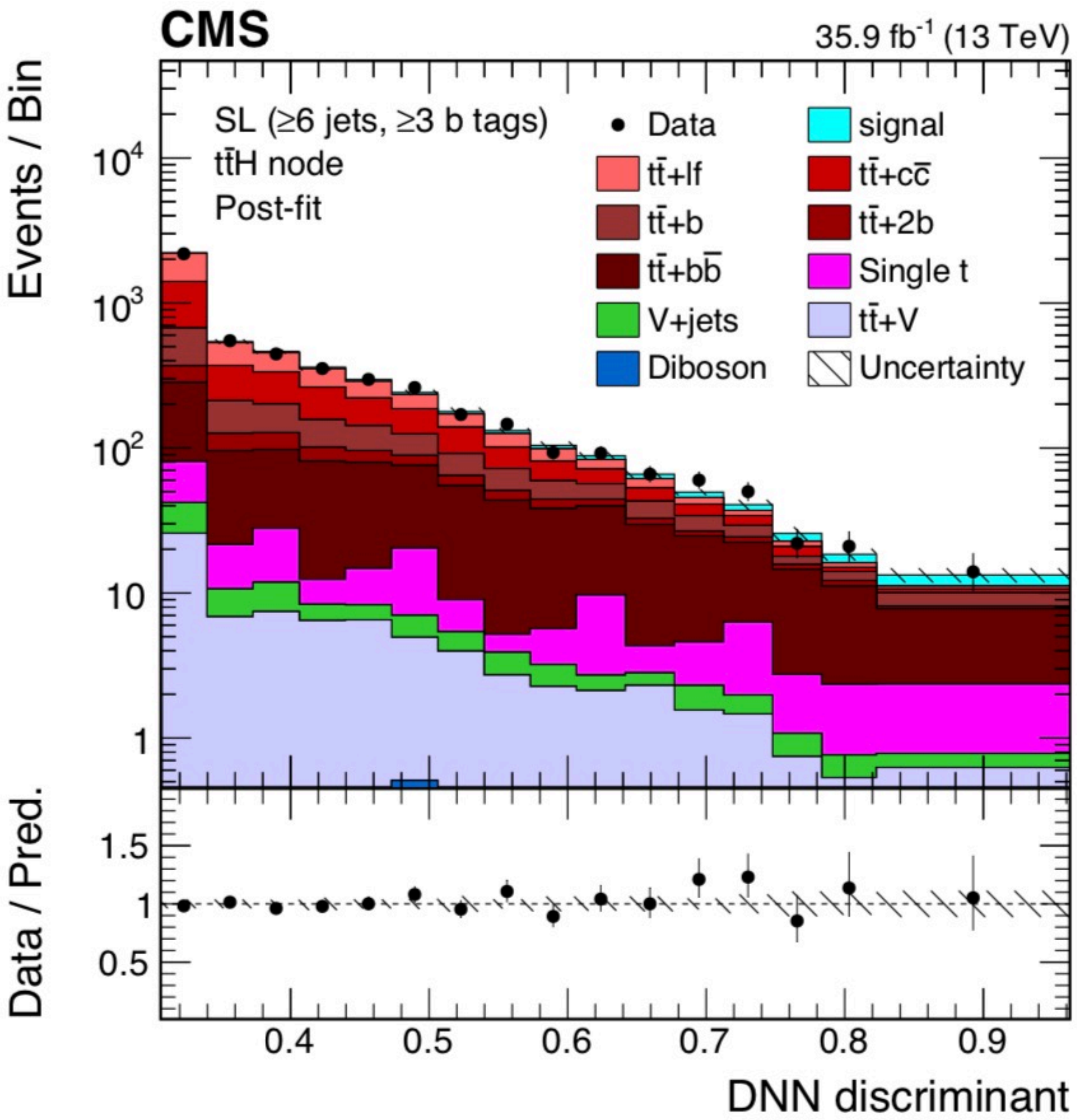
- ≥ 4j, ≥ 4b: Two categories based on BDT; Matrix Element Method (MEM) separating signal from tt+HF background

ttH(bb): SingleLepton

arXiv:1804.03682

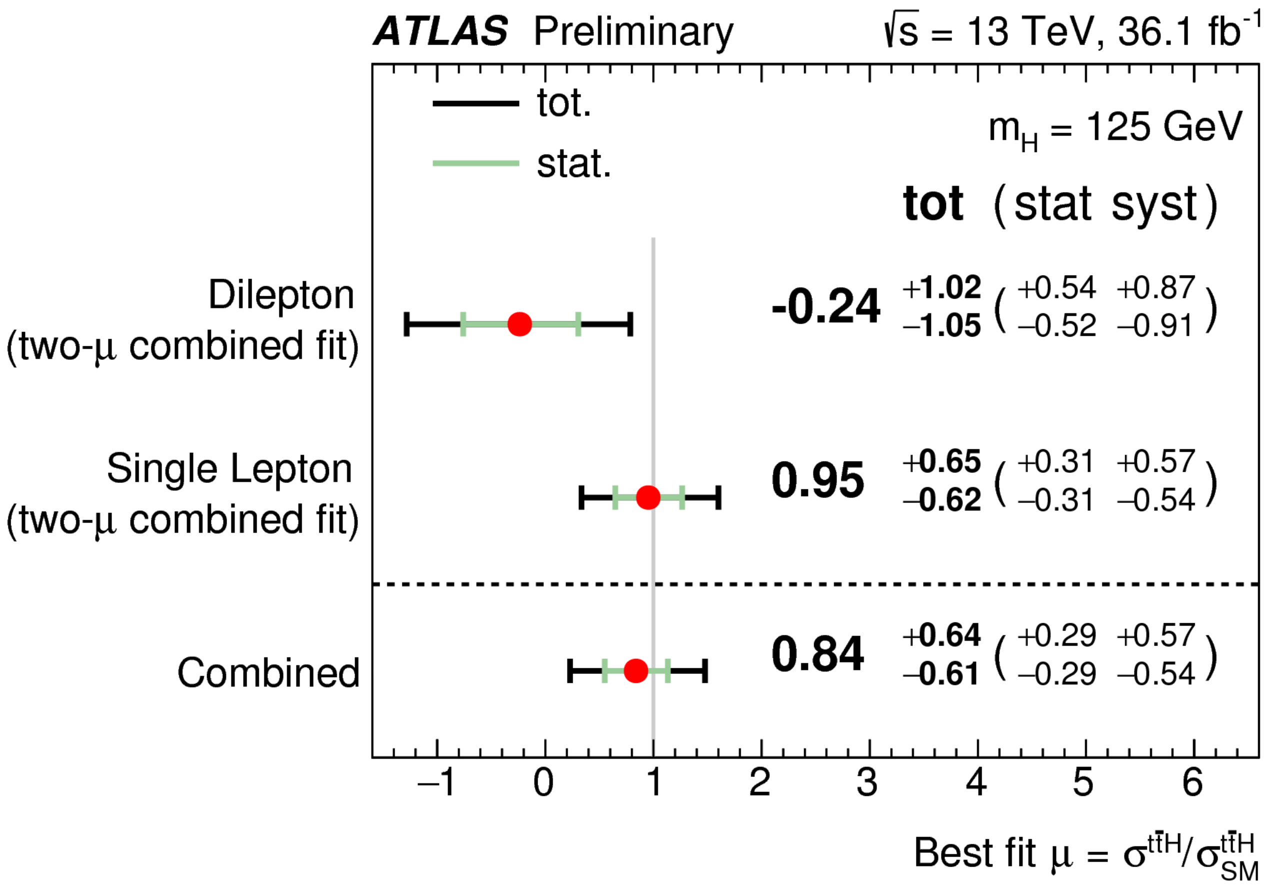
- Events categorized as number of jets and b-tagged jets





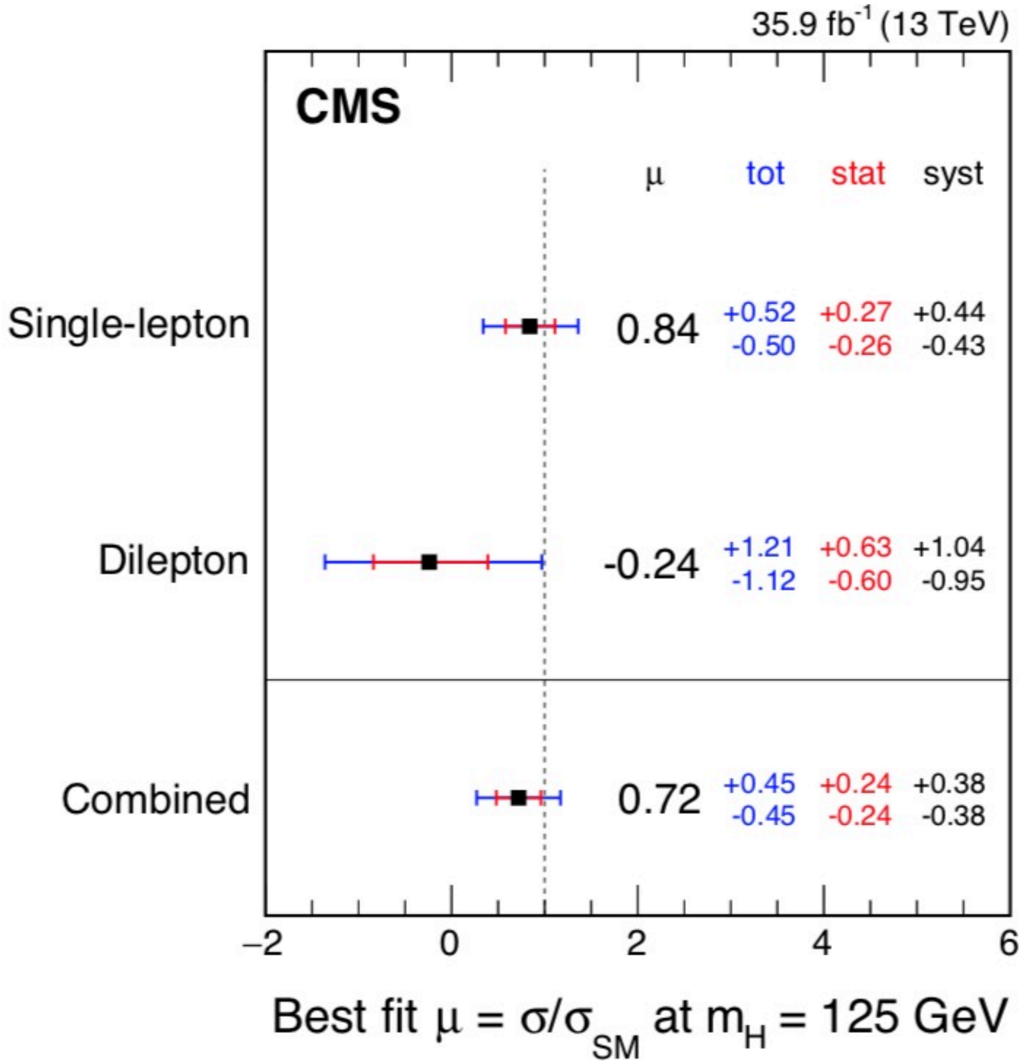
- Multi-classification Deep Neural Network per jet category:
  - 6 categories: 1 signal and 5 tt+HF categories
- Final discriminant is the DNN output

ttH(bb) Leptonic: Results



$$\mu = 0.84^{+0.64}_{-0.61}$$

significance: 1.4σ (expected 1.6σ)



$$\mu = 0.72^{+0.45}_{-0.45}$$

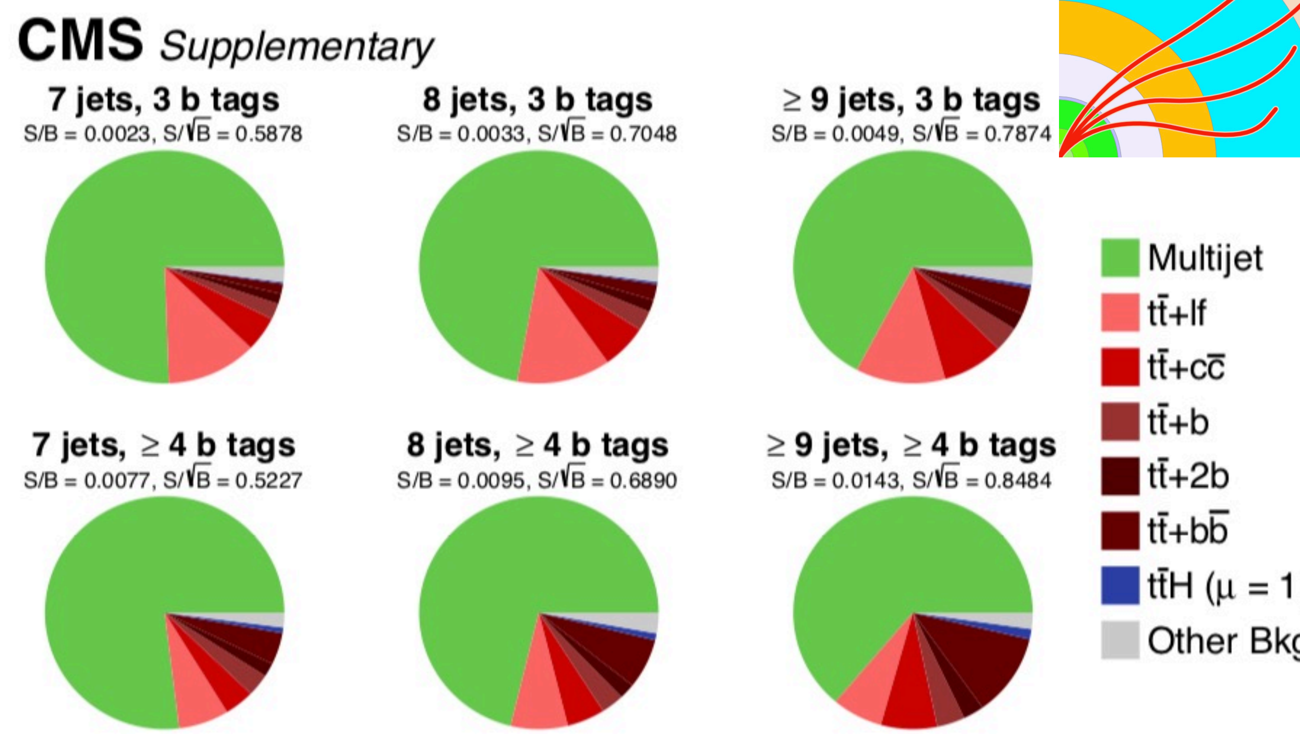
significance: 1.6σ (expected 2.2σ)

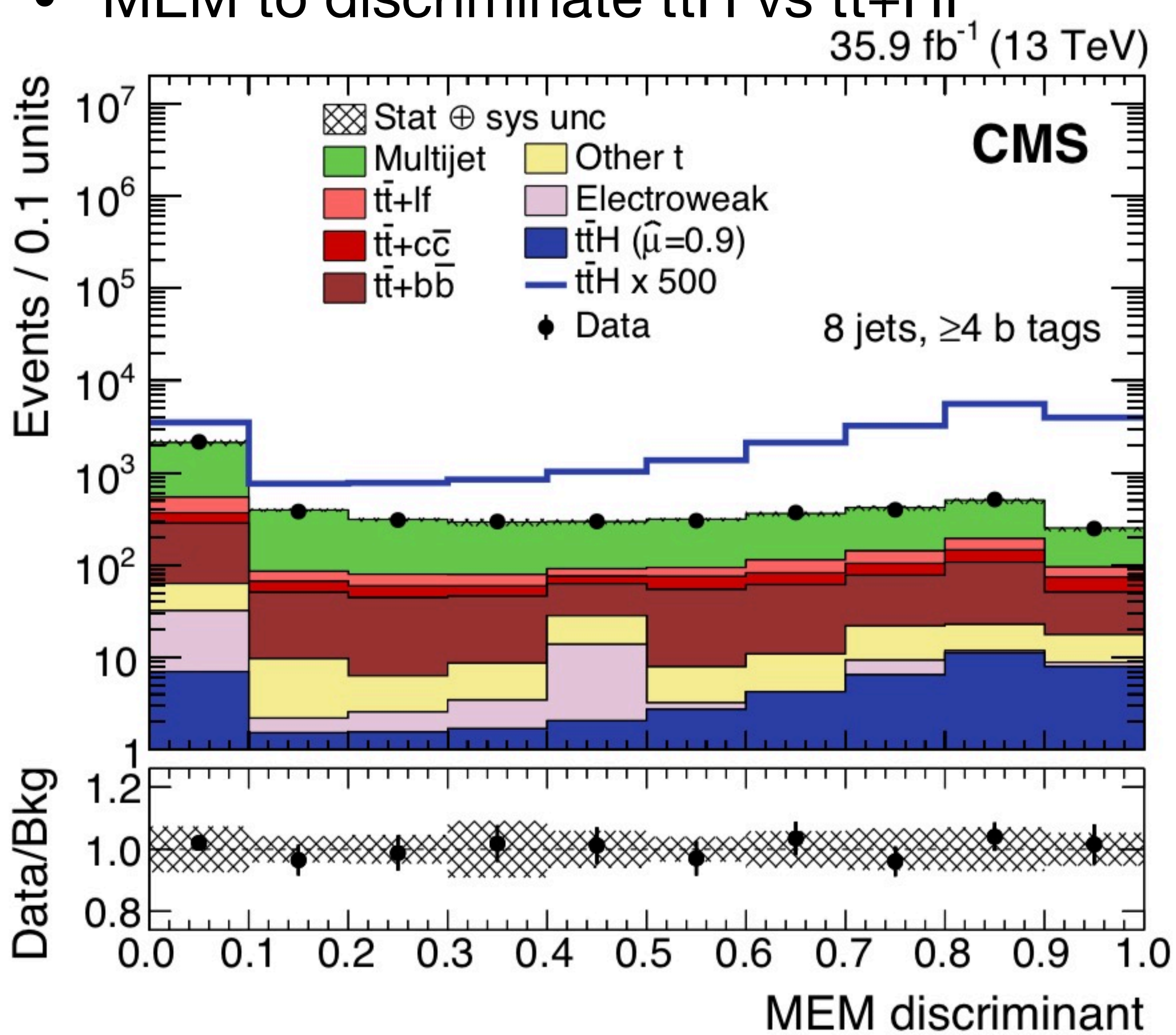
- Systematics:
    - Largest impact from tt+bb generator modeling
    - b-tagging
    - ATLAS considers additional tt+HF systematics (backup)

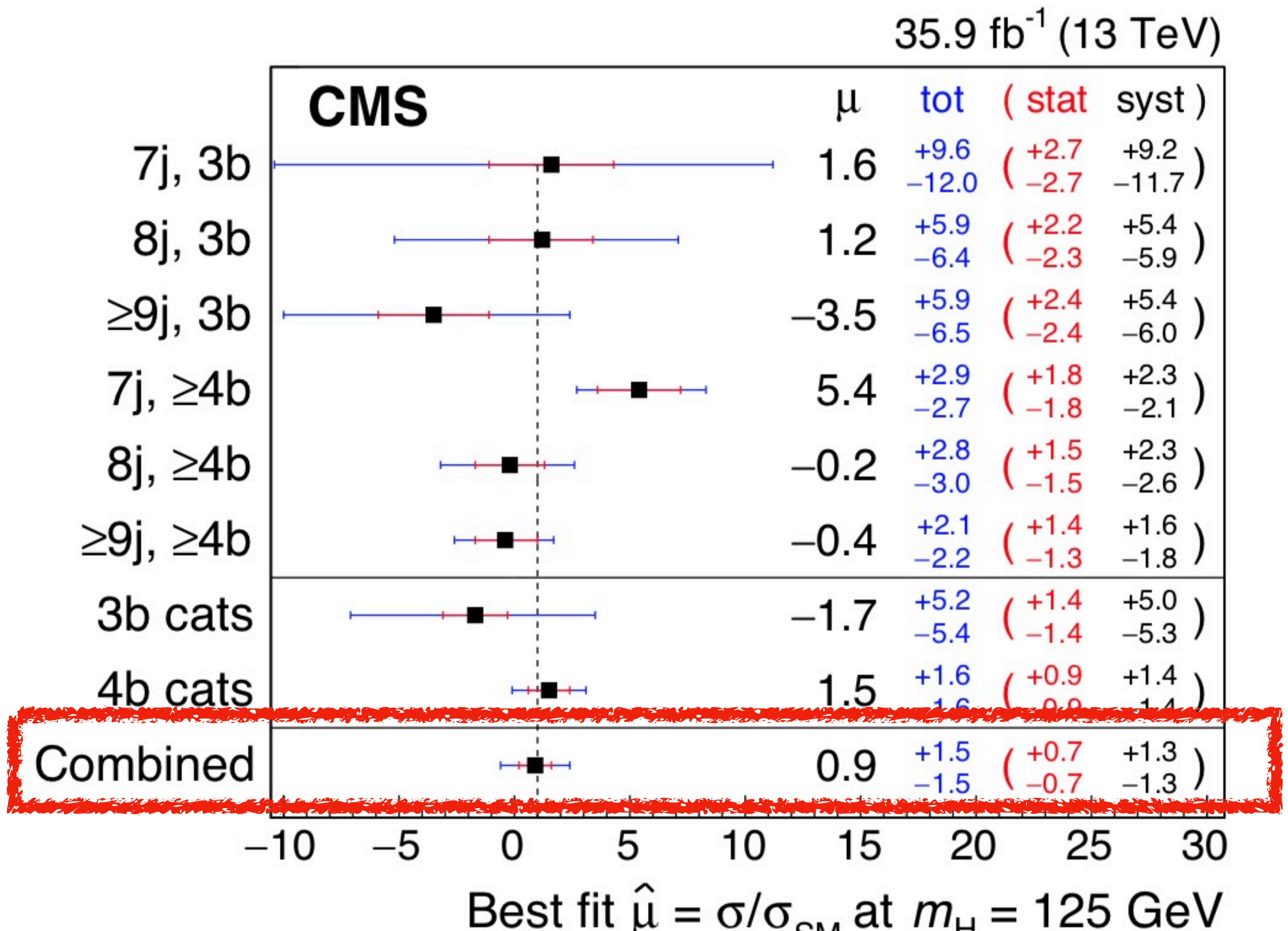
ttH(bb) hadronic

JHEP 06 (2018) 101

- Events categorized as number of jets and b-tagged jets
  - ≥ 7 jets, ≥ 3 b-tagged jets, HT > 500 GeV, no leptons
- Dominant background is QCD-multijet
  - Modeling
    - Shape from low b-tag multiplicity control region in data
    - Normalization from final fit to data
- MEM to discriminate ttH vs tt+HF

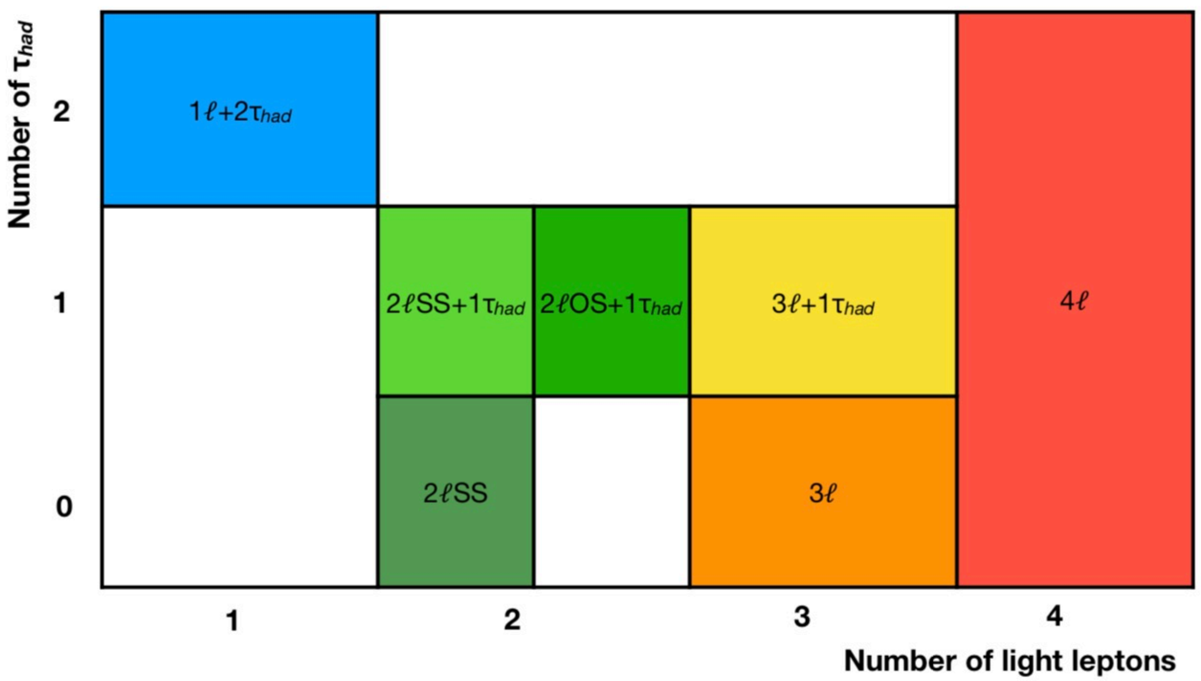


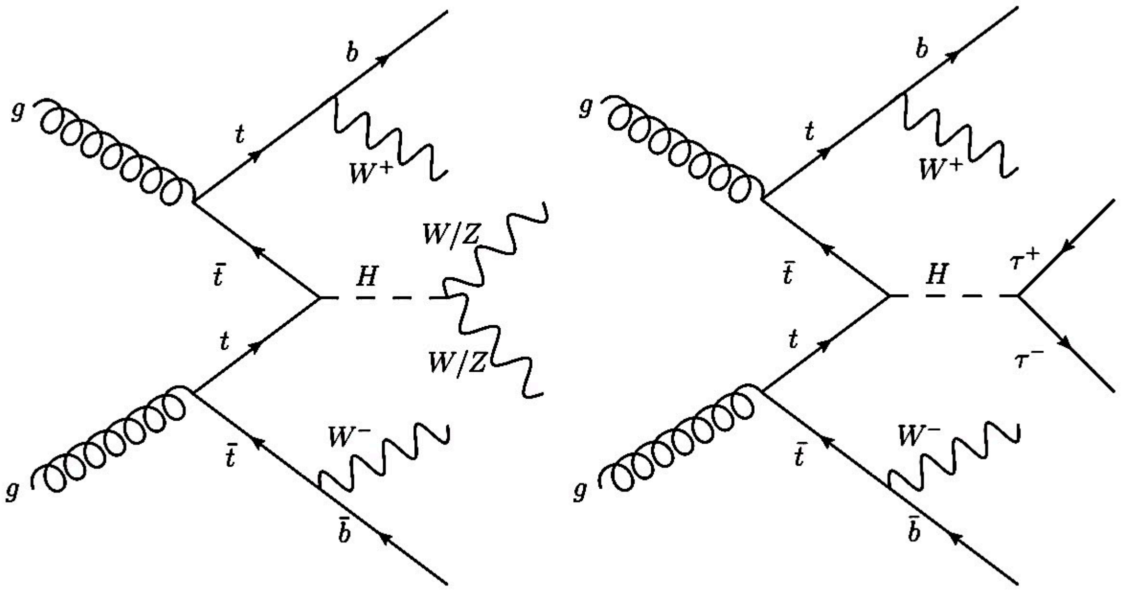




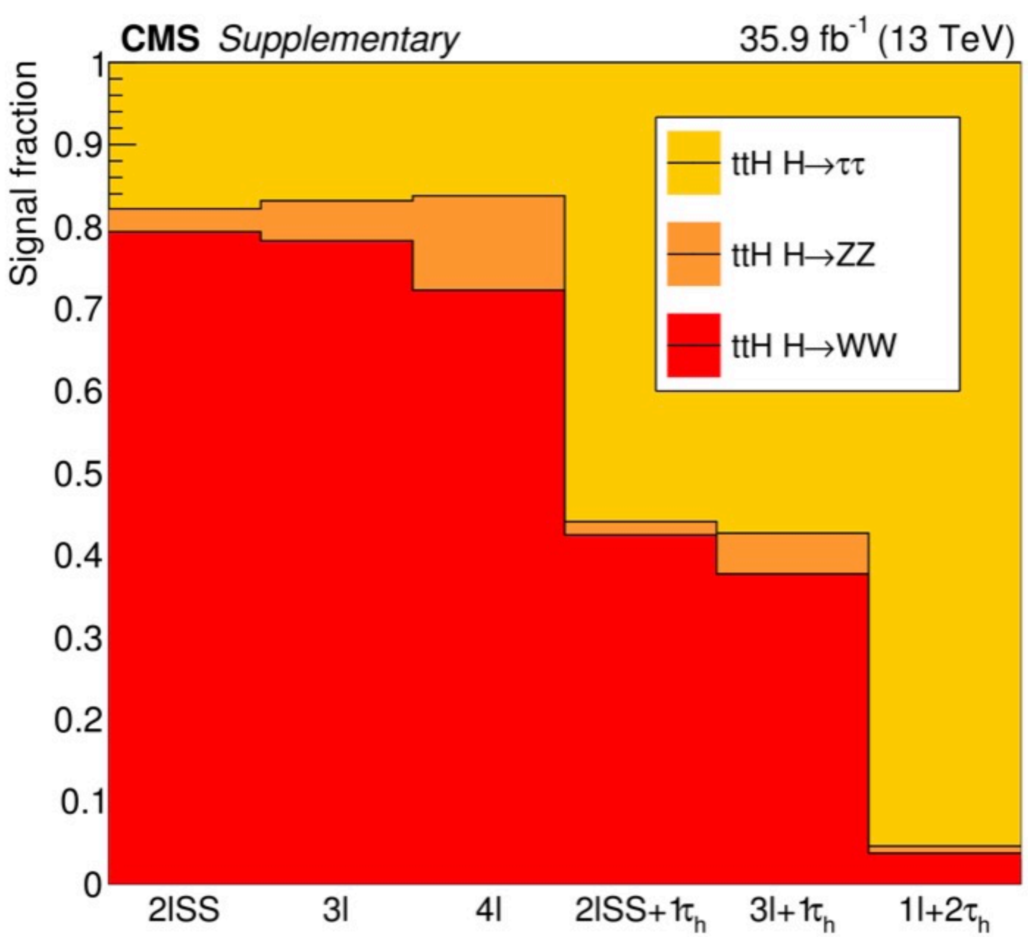
ttH Multileptons

- Targets $H \to ZZ^*, H \to WW^*, H \to \tau^+\tau^-$
- Events categorized based on number of light leptons and hadronic taus





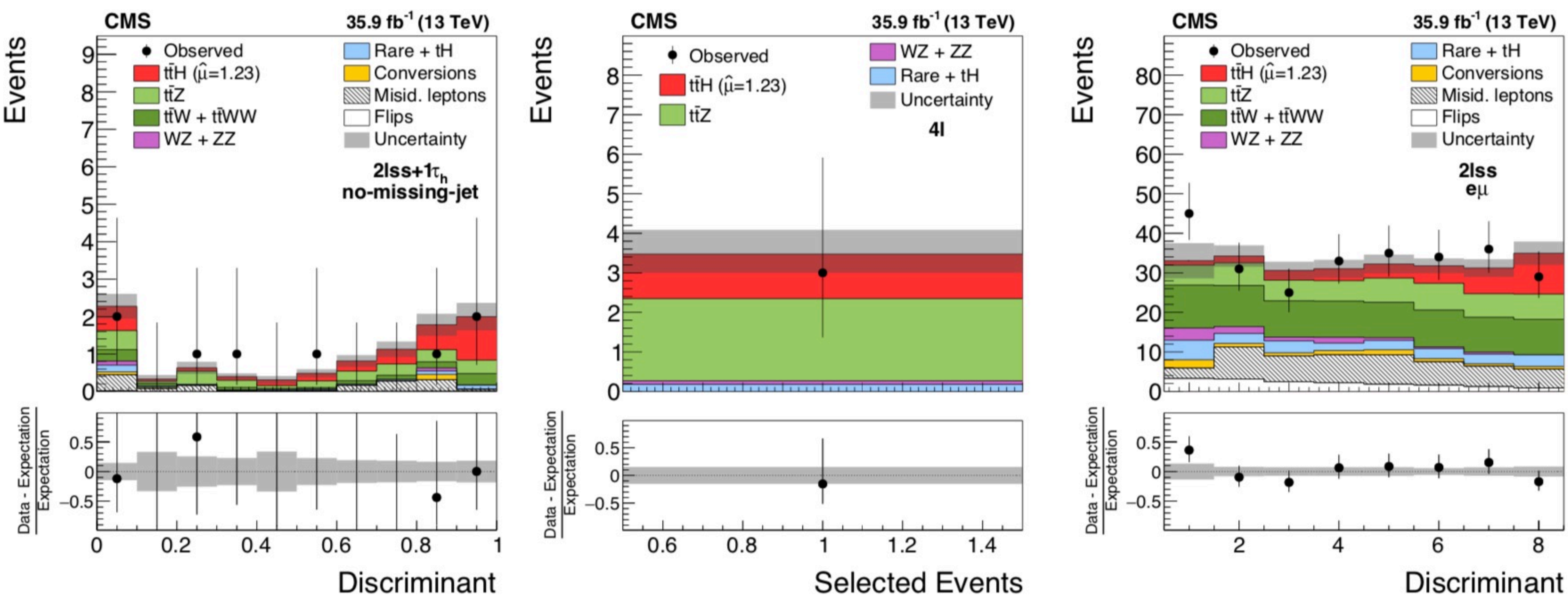
- Main backgrounds
  - Fake leptons (heavy flavor, conversions)
  - Irreducible background from ttV
- High lepton multiplicity reduces backgrounds



JHEP 08 (2018) 066

ttHML: CMS strategy

- Event categorization based on number of light leptons and hadronic taus.
- Final fit variables
  - MEM against ttZ (2 l same-sign + 1 τ_h)
  - Yields in 4-leptons
  - BDTs against tt+jets (1l+2 τ_h) and tt + jets, tt+ V (2 l same-sign, 3 l)



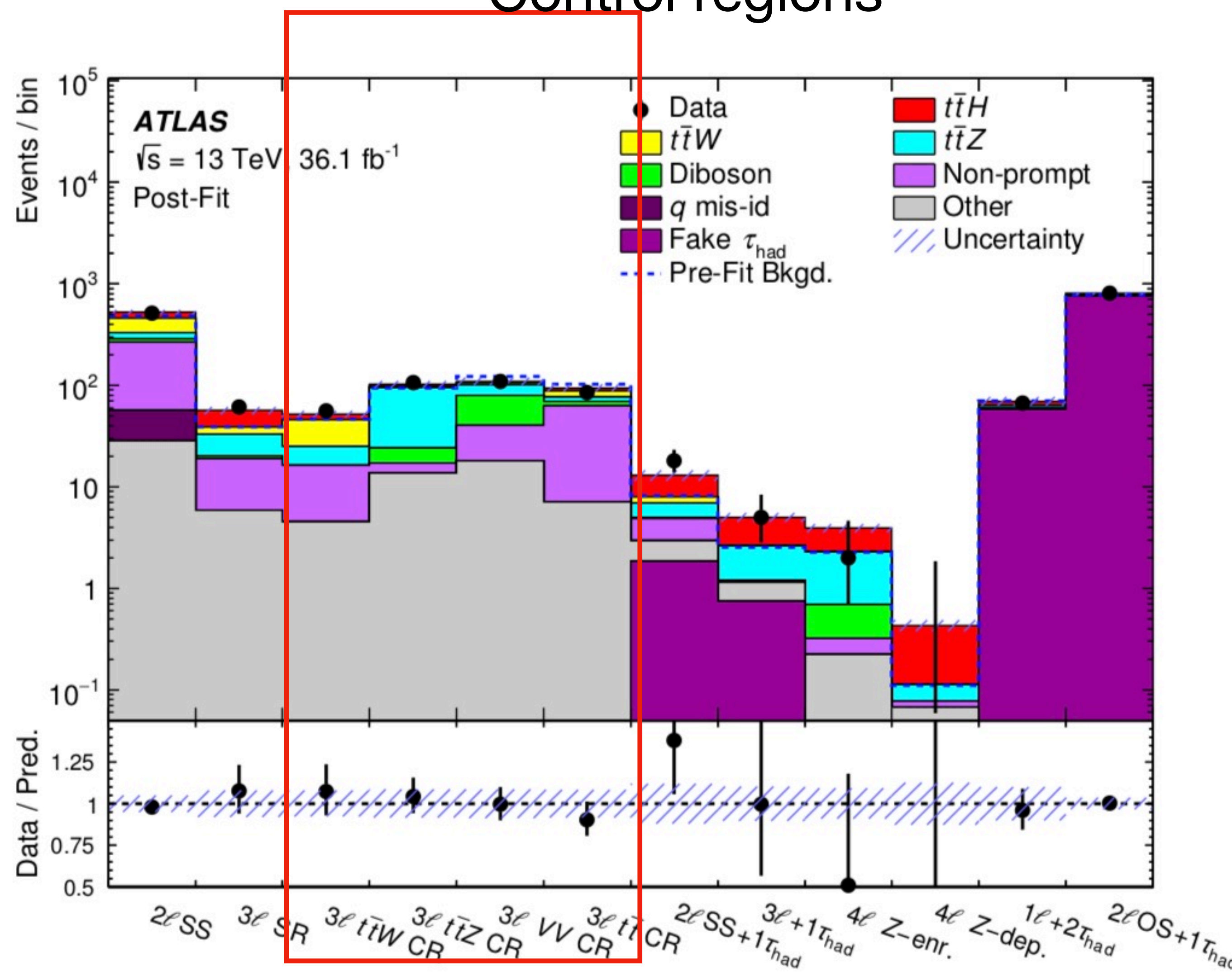
ttHML: ATLAS strategy

Phys. Rev. D 97, 072003

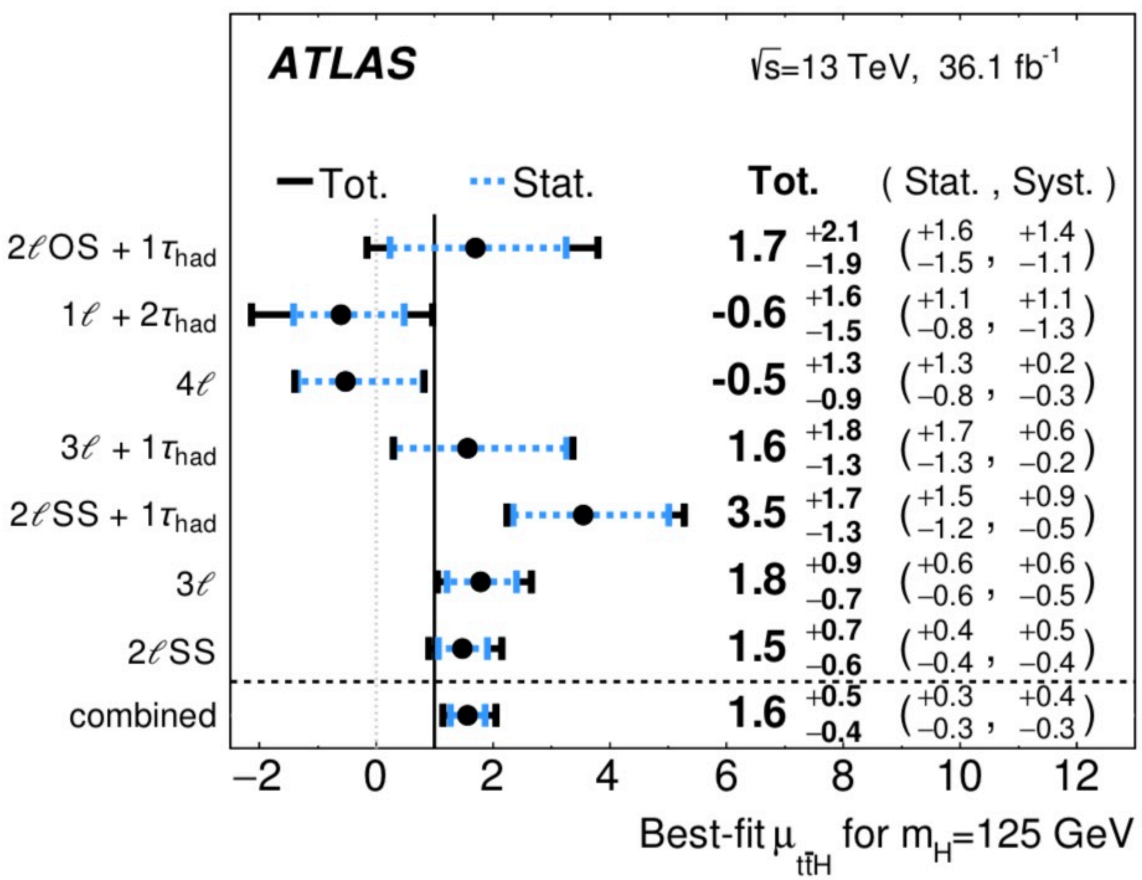
	$2\ell SS$	3ℓ	4ℓ	$1\ell+2\tau_{had}$	$2\ell SS+1\tau_{had}$	$2\ell OS+1\tau_{had}$	$3\ell+1\tau_{had}$
BDT trained against	Fakes and $t\bar{t}V$	$t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, VV	$t\bar{t}Z$ / -	$t\bar{t}$	all	$t\bar{t}$	-
Discriminant	2×1D BDT	5D BDT	Event count	BDT	BDT	BDT	Event count
Number of bins	6	5	1 / 1	2	2	10	1
Control regions	-	4	-	-	-	-	-

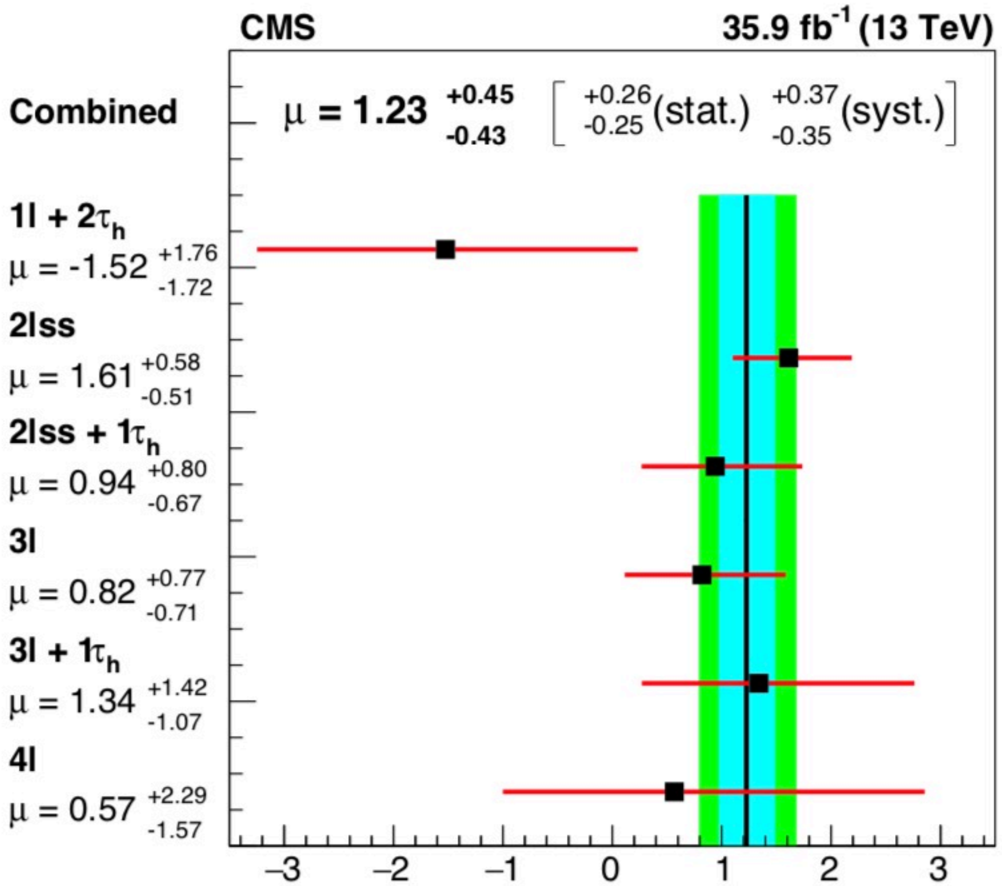
- Simultaneous fit in 12 regions (CR+SR)
  - Multiclass BDT in 3l
- Single bin used in 3l CR's as well as low stat SR's
  - 3l+1τ
  - 4l (z-enriched,z-depeleted)
- BDT shape information used in 5 SR's

Control regions



ttHML: Results





- Measured $\mu = 1.6^{+0.5}_{-0.4}$ with a significance with respect to background only hypothesis **4.1σ** **(expct. 2.8σ)**
- Measured $\mu = 1.23^{+0.45}_{-0.43}$ with a significance with respect to background only hypothesis 3.2 **σ** **(expct. 2.8σ)**

- Limited by fake lepton estimation, tau identification JES, ttH and ttV process modeling.
- Several channels are limited by statistics.

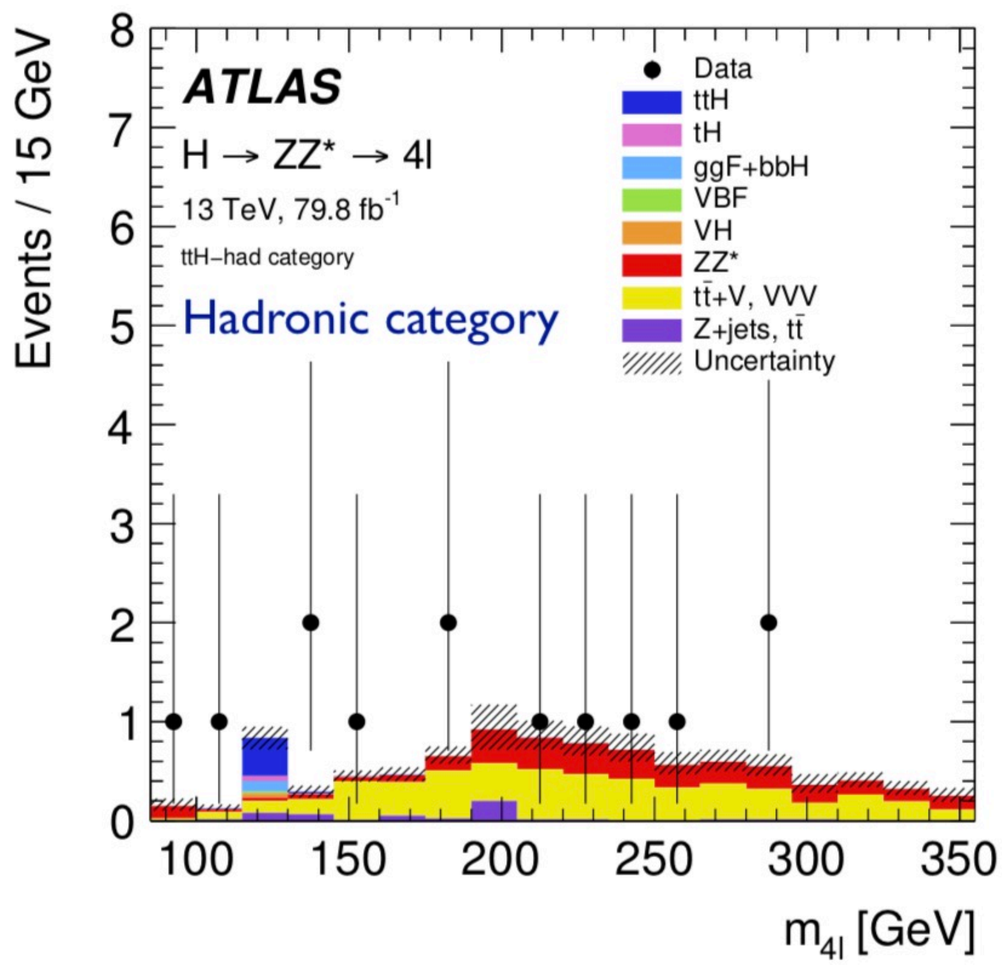
ttH(ZZ*)

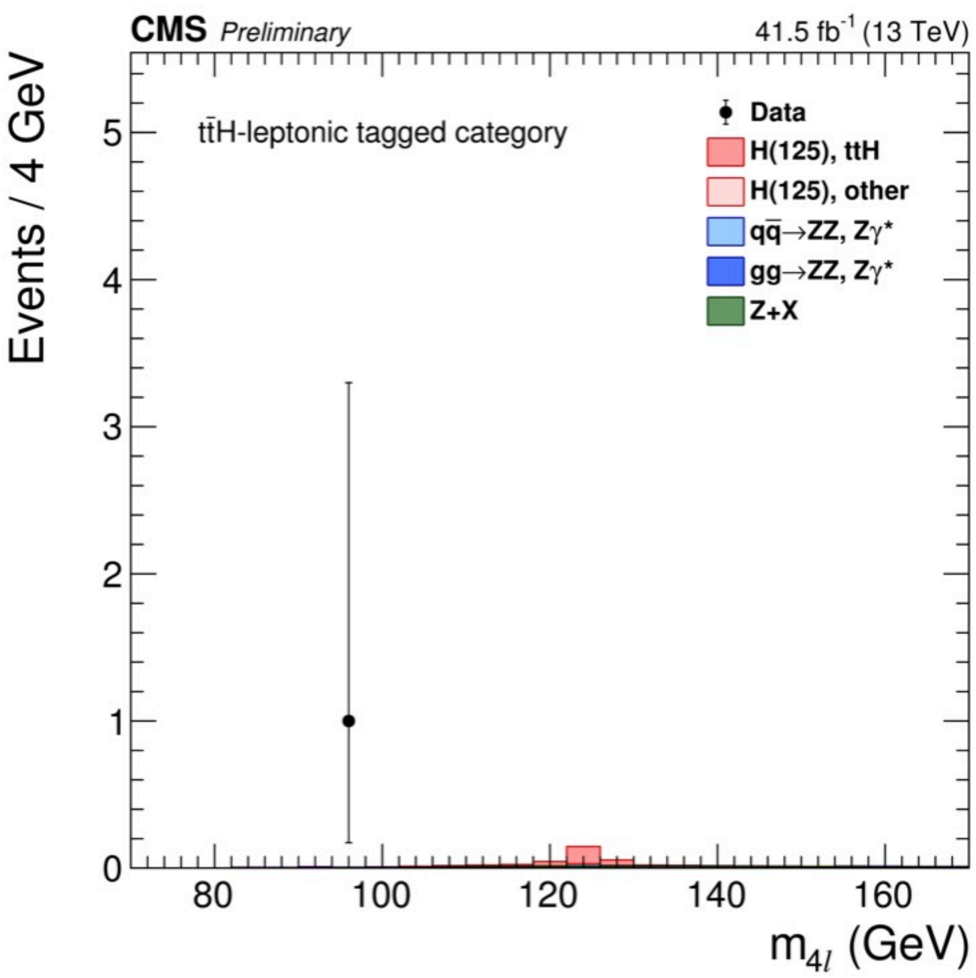
JHEP 03 (2018) 095

PAS-HIG-18-001

- Candidate event selection:
    - At least two opposite sign same flavor light lepton
- Extremely low rate (about 0.6 event expected with the statistics available); but clean final state with high S/B

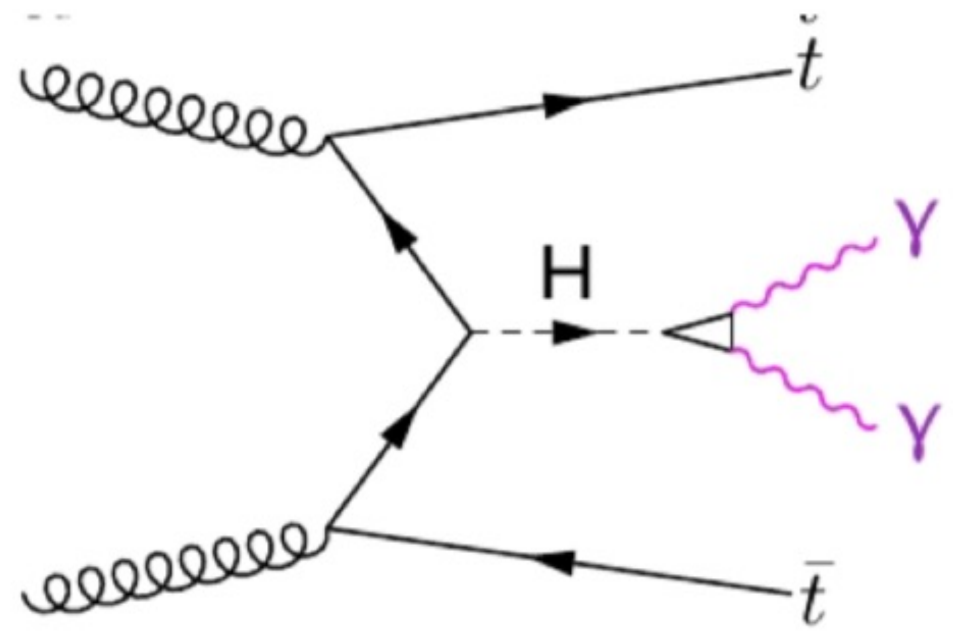
- Dedicated ttH channel, part of the global H(ZZ*) analysis





- No candidates in a window $115 < m_{4l} < 130$

ttH(γγ)

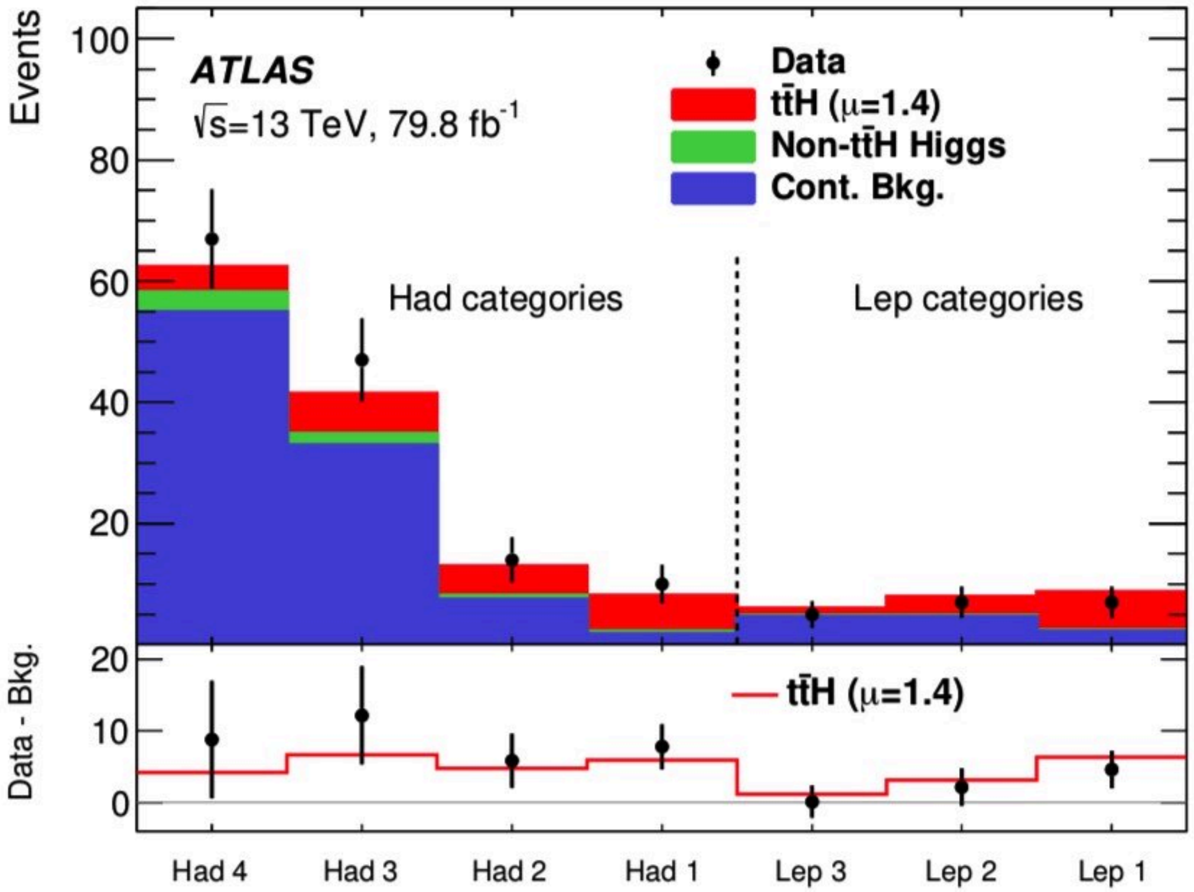


- Clear signature from the photons
- Higgs boson can be reconstructed from the two isolated photons.
- Dedicated ttH channel part of the H(γγ) analysis.

ttH(γγ) ATLAS strategy

Phys. Lett. B 784 (2018) 173

- Events classified into hadronic and leptonic channels based on decay topology of tt-bar system.
- BDT's used in hadronic and leptonic channels.
  - Separating ttH from main background ttγγ and γγ background processes.

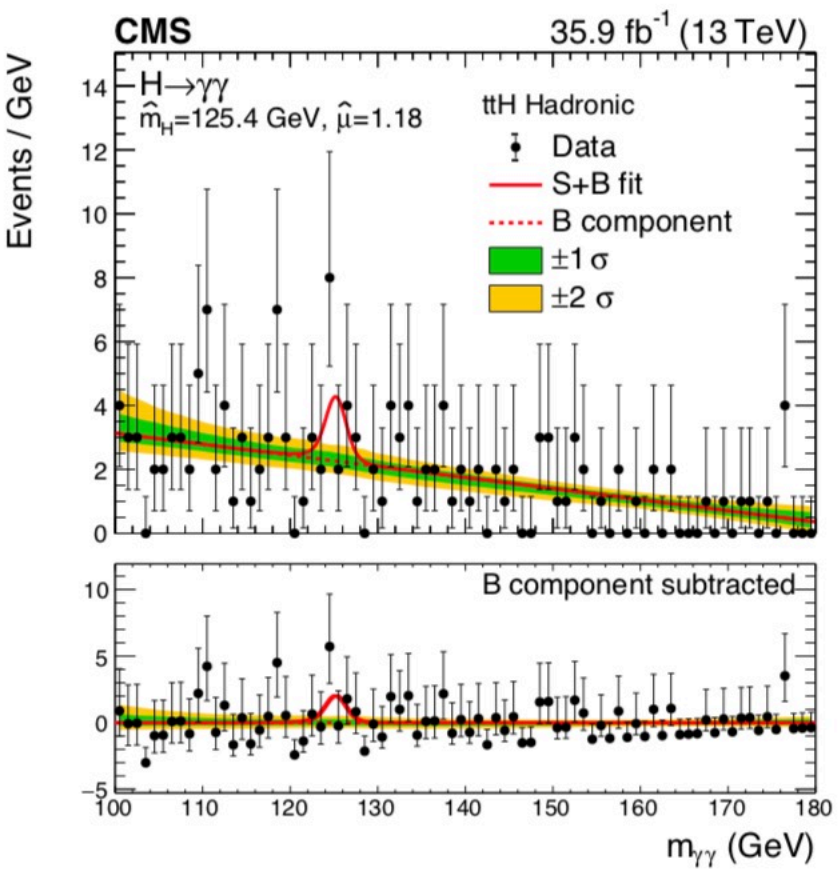


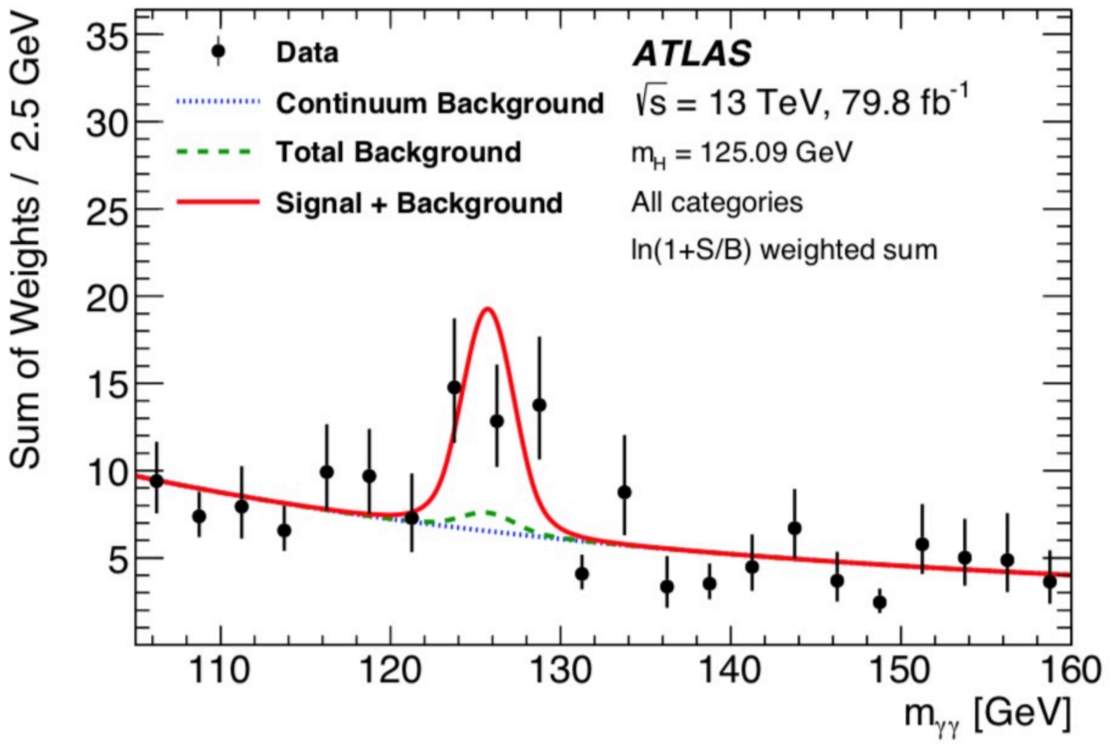
- Based on the BDT output 4 hadronic and 3 leptonic categories are defined for the fit.
  - S+B fit in each category between 105 < $m_{\gamma\gamma}$ < 160 GeV

ttH(γγ) Results

Phys. Lett. B 784 (2018) 173

arXiv:1804.02716





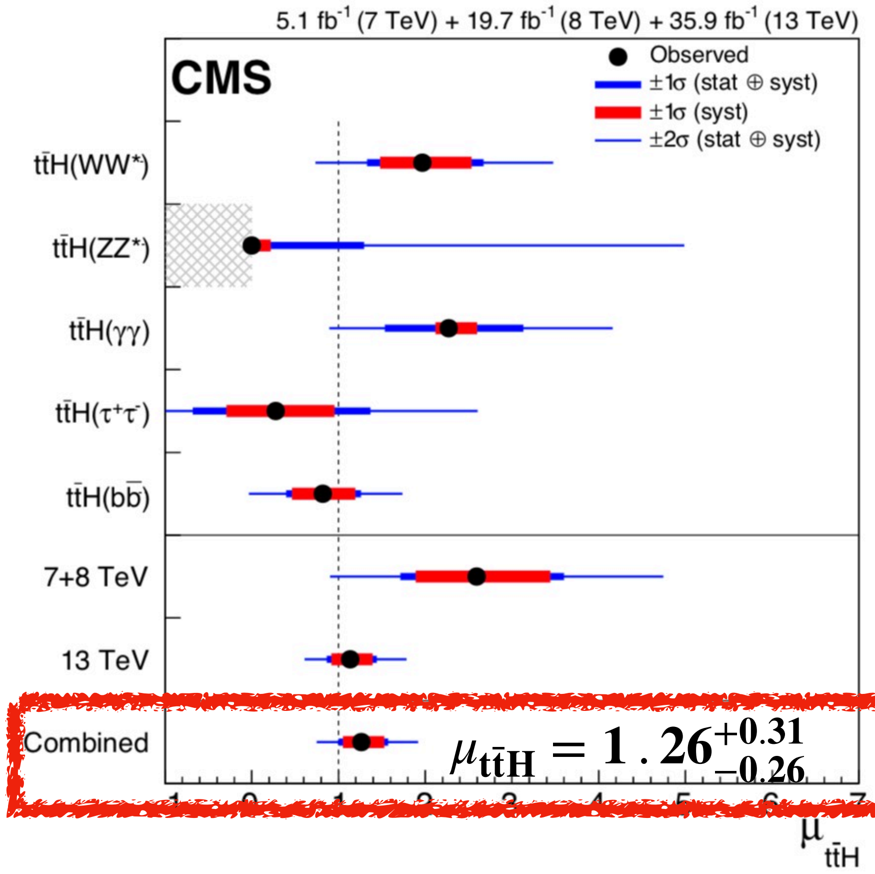
- Signal extracted from a fit to $m_{\gamma\gamma}$

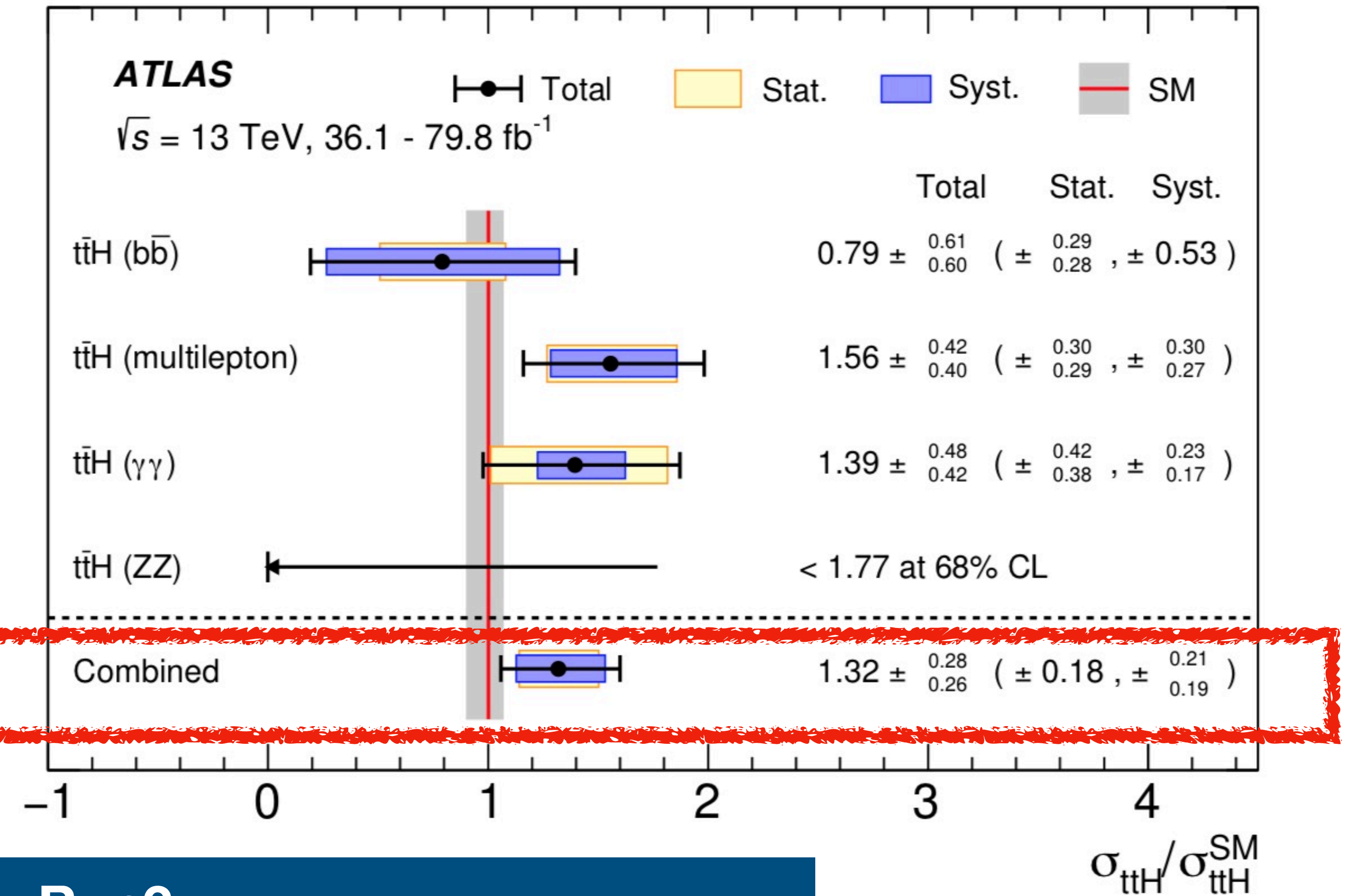
ATLAS: $\mu = 1.39^{+0.48}_{-0.42}$ observed (expected) significance: 4.1σ (3.7σ)

CMS: $\mu = 2.2^{+0.9}_{-0.8}$

ttH Combination

- Combine different decay modes assuming SM for the decay BR
- Combine also with run 1 data (lower sensitivity, 20 (5) fb^{-1} @ 8(7) TeV)





	Run1+Run2	Run2
ATLAS	6.3σ observed (5.1σ expected)	5.8σ observed (4.9σ expected)
CMS	5.2σ observed(4.2σ expected)	

ATLAS uses 2017 data as well for ttH(γγ), ttH(ZZ*) analysis

kappa

	ATLAS	CMS
Kt	$1.09^{+0.15}_{-0.14}$	$0.98^{+0.14}_{-0.14}$

ATLAS-CONF-2018-031 CMS-HIG-17-031

Conclusion and Outlook

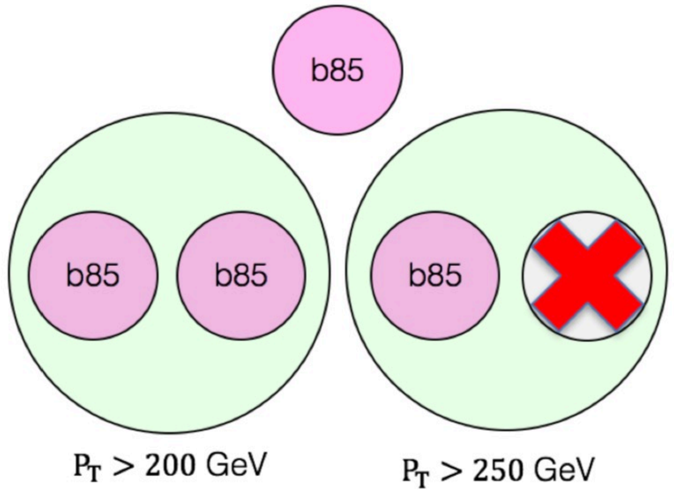
- Both CMS and ATLAS has observed ttH process at the LHC
- Leading measurement uncertainties
    - Process modeling: tt+HF and ttV
    - Light lepton fake estimates, jet systematics
- New data (about 140 fb^{-1} collected)
    - Statistically limited channels will gain sensitivity
    - Impacts statistically limited systematics: eg. fake estimates
    - Differential measurements
    - Simultaneous measurements, ratios
- HL-LHC
    - Rare decay channels become accessible eg: ttH(μμ), ttH(Zγ)

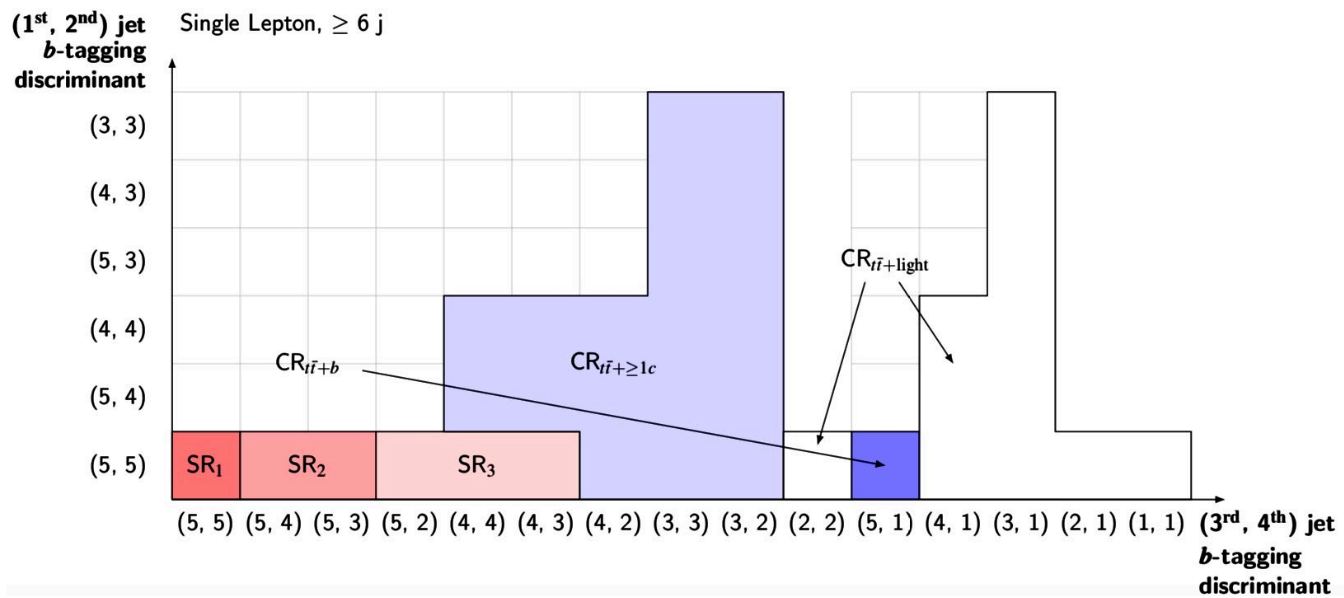
Backup

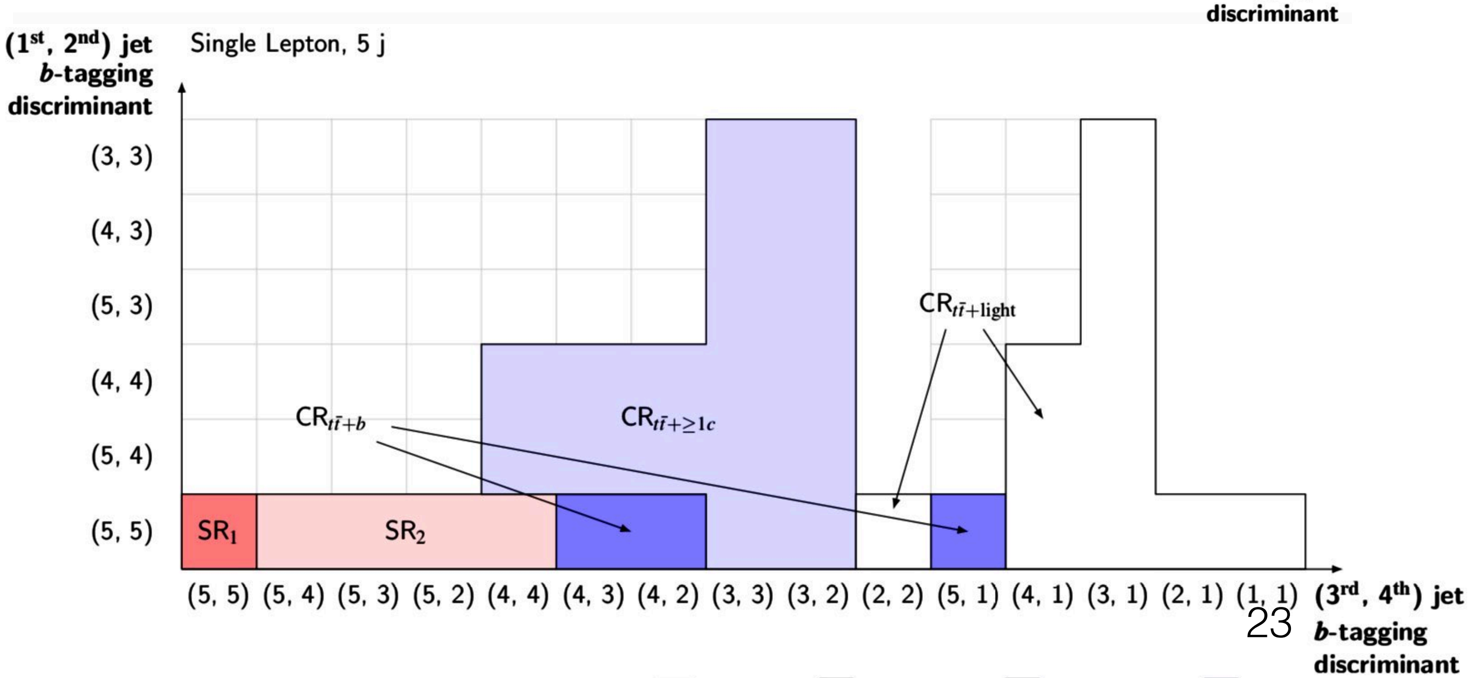
ttH(bb): Single lepton

Sub channels

- Exactly 1 lepton.
- Boosted Category:
  - Presence of two R=1.0 reclustered jets (for Higgs and top candidate)
  - Remove jets which have pT < 50GeV and pT>1500 GeV
- Resolved Category: Failing boosted selection
  - Require ≥5 jets and ≥2 very-tight or ≥3 medium b-tagged jets







- SR's defined for 5 and ≥6 jets
- Boosted channel is not categorized further.
- Highest purity in 4 very tight b-tag bins (≥6 jets)
- CR's for tt+ ≥1b, tt+ ≥1c and tt + light

CMS ttH(bb)Leptonic

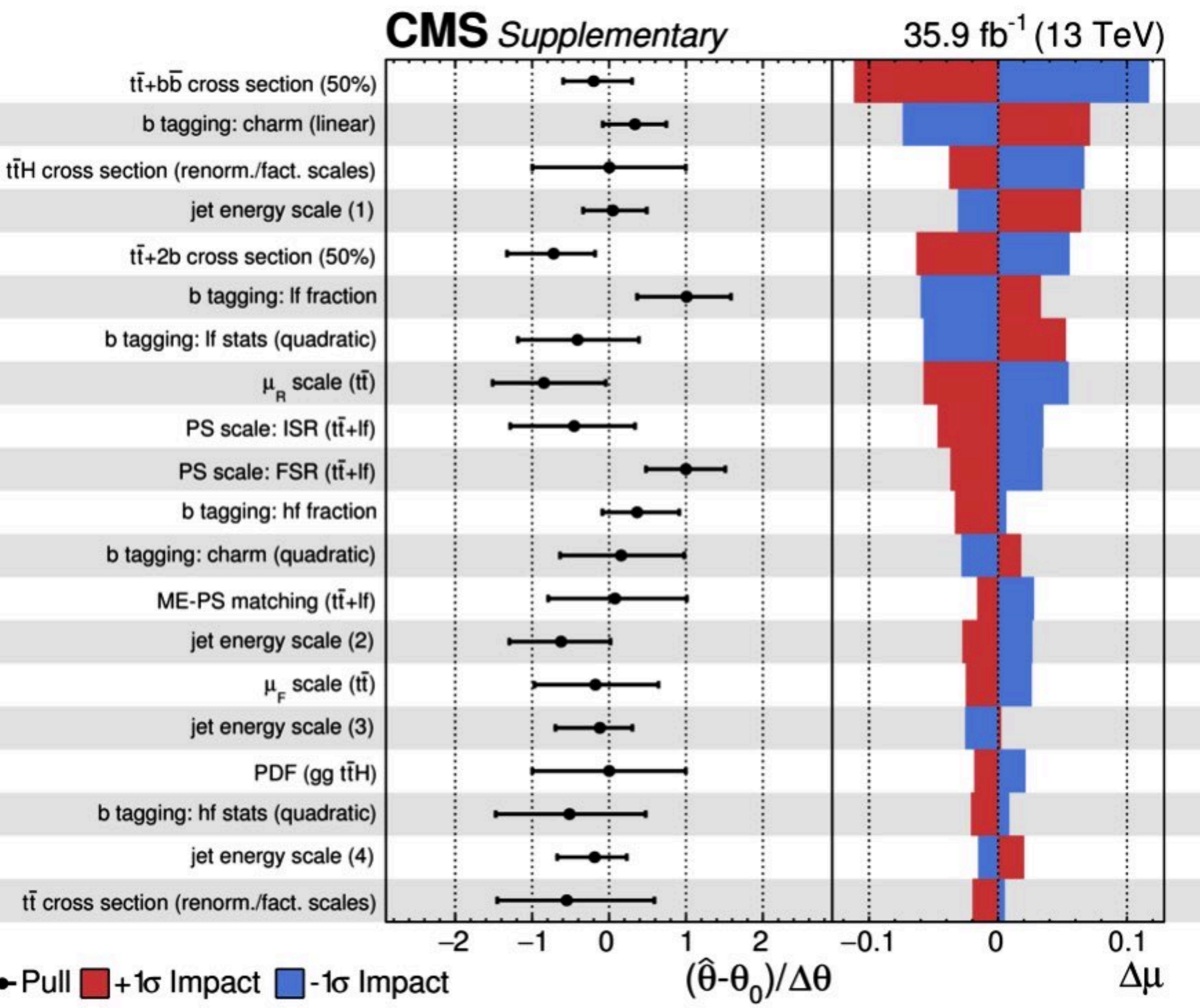
Uncertainties

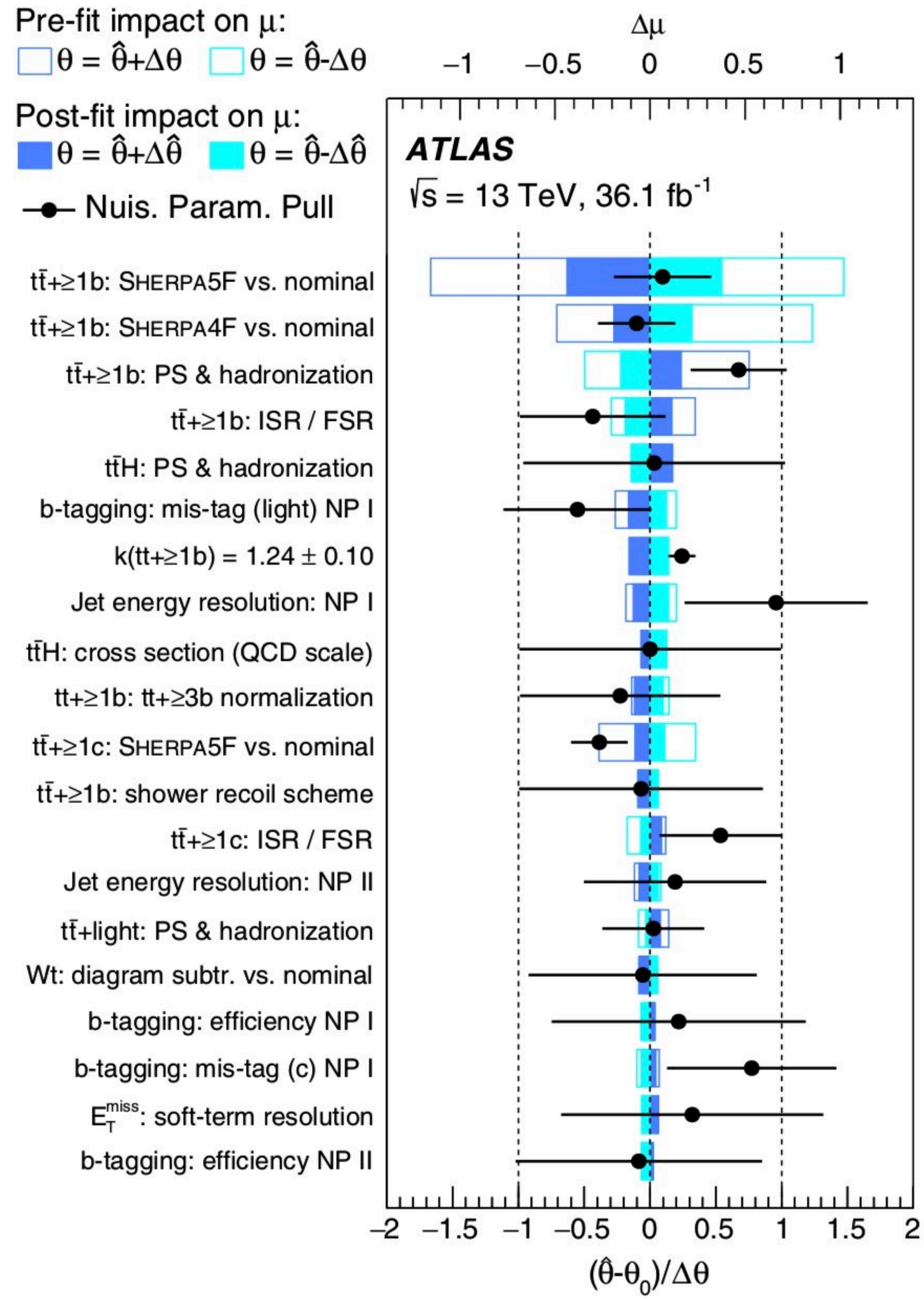
Uncertainty source	$\pm\Delta\mu$ (observed)	$\pm\Delta\mu$ (expected)
Total experimental	$+0.15/-0.16$	$+0.19/-0.17$
b tagging	$+0.11/-0.14$	$+0.12/-0.11$
jet energy scale and resolution	$+0.06/-0.07$	$+0.13/-0.11$
Total theory	$+0.28/-0.29$	$+0.32/-0.29$
$t\bar{t}$+hf cross section and parton shower	$+0.24/-0.28$	$+0.28/-0.28$
Size of the simulated samples	$+0.14/-0.15$	$+0.16/-0.16$
Total systematic	$+0.38/-0.38$	$+0.45/-0.42$
Statistical	$+0.24/-0.24$	$+0.27/-0.27$
Total	$+0.45/-0.45$	$+0.53/-0.49$

CMS Combinations

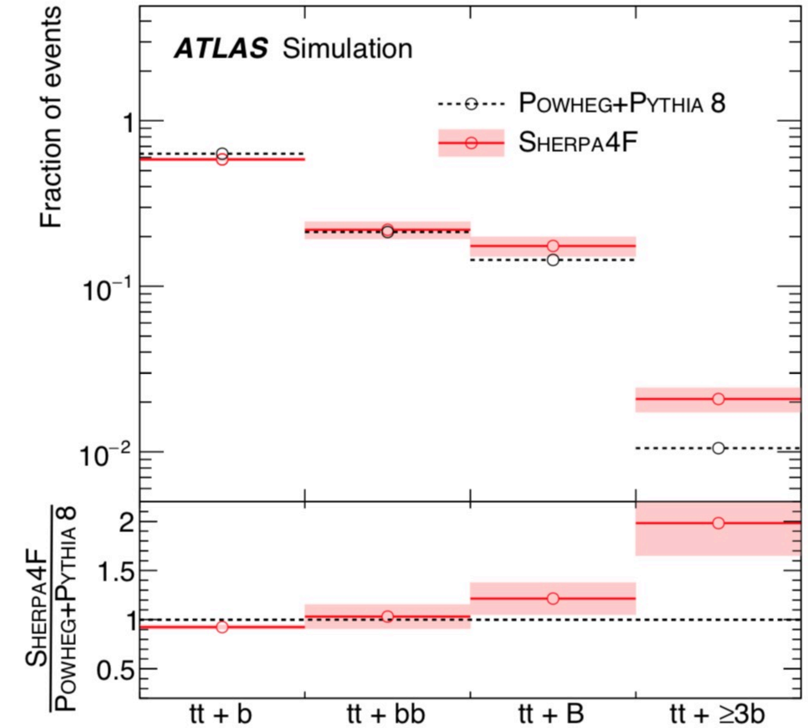
		Uncertainty			
Parameter	Best fit	Stat	Expt	Thbgd	Thsig
$\mu^{WW^*}_{t\bar{t}H}$	$1.97^{+0.71}_{-0.64}$ $\left(^{+0.57}_{-0.54}\right)$	$^{+0.42}_{-0.41}$ $\left(^{+0.39}_{-0.38}\right)$	$^{+0.46}_{-0.42}$ $\left(^{+0.36}_{-0.34}\right)$	$^{+0.21}_{-0.21}$ $\left(^{+0.17}_{-0.17}\right)$	$^{+0.25}_{-0.12}$ $\left(^{+0.12}_{-0.03}\right)$
$\mu^{ZZ^*}_{t\bar{t}H}$	$0.00^{+1.30}_{-0.00}$ $\left(^{+2.89}_{-0.99}\right)$	$^{+1.28}_{-0.00}$ $\left(^{+2.82}_{-0.99}\right)$	$^{+0.20}_{-0.00}$ $\left(^{+0.51}_{-0.00}\right)$	$^{+0.04}_{-0.00}$ $\left(^{+0.15}_{-0.00}\right)$	$^{+0.09}_{-0.00}$ $\left(^{+0.27}_{-0.00}\right)$
$\mu^{\gamma\gamma}_{t\bar{t}H}$	$2.27^{+0.86}_{-0.74}$ $\left(^{+0.73}_{-0.64}\right)$	$^{+0.80}_{-0.72}$ $\left(^{+0.71}_{-0.64}\right)$	$^{+0.15}_{-0.09}$ $\left(^{+0.09}_{-0.04}\right)$	$^{+0.02}_{-0.01}$ $\left(^{+0.01}_{-0.00}\right)$	$^{+0.29}_{-0.13}$ $\left(^{+0.13}_{-0.05}\right)$
$\mu^{\tau^+\tau^-}_{t\bar{t}H}$	$0.28^{+1.09}_{-0.96}$ $\left(^{+1.00}_{-0.89}\right)$	$^{+0.86}_{-0.77}$ $\left(^{+0.83}_{-0.76}\right)$	$^{+0.64}_{-0.53}$ $\left(^{+0.54}_{-0.47}\right)$	$^{+0.10}_{-0.09}$ $\left(^{+0.09}_{-0.08}\right)$	$^{+0.20}_{-0.19}$ $\left(^{+0.14}_{-0.01}\right)$
$\mu^{b\bar{b}}_{t\bar{t}H}$	$0.82^{+0.44}_{-0.42}$ $\left(^{+0.44}_{-0.42}\right)$	$^{+0.23}_{-0.23}$ $\left(^{+0.23}_{-0.22}\right)$	$^{+0.24}_{-0.23}$ $\left(^{+0.24}_{-0.23}\right)$	$^{+0.27}_{-0.27}$ $\left(^{+0.26}_{-0.27}\right)$	$^{+0.11}_{-0.03}$ $\left(^{+0.11}_{-0.04}\right)$
$\mu^{7+8\,TeV}_{t\bar{t}H}$	$2.59^{+1.01}_{-0.88}$ $\left(^{+0.87}_{-0.79}\right)$	$^{+0.54}_{-0.53}$ $\left(^{+0.51}_{-0.49}\right)$	$^{+0.53}_{-0.49}$ $\left(^{+0.48}_{-0.44}\right)$	$^{+0.55}_{-0.49}$ $\left(^{+0.50}_{-0.44}\right)$	$^{+0.37}_{-0.13}$ $\left(^{+0.14}_{-0.02}\right)$
$\mu^{13\,TeV}_{t\bar{t}H}$	$1.14^{+0.31}_{-0.27}$ $\left(^{+0.29}_{-0.26}\right)$	$^{+0.17}_{-0.16}$ $\left(^{+0.16}_{-0.16}\right)$	$^{+0.17}_{-0.17}$ $\left(^{+0.17}_{-0.16}\right)$	$^{+0.13}_{-0.12}$ $\left(^{+0.13}_{-0.12}\right)$	$^{+0.14}_{-0.06}$ $\left(^{+0.11}_{-0.05}\right)$
$\mu_{t\bar{t}H}$	$1.26^{+0.31}_{-0.26}$ $\left(^{+0.28}_{-0.25}\right)$	$^{+0.16}_{-0.16}$ $\left(^{+0.15}_{-0.15}\right)$	$^{+0.17}_{-0.15}$ $\left(^{+0.16}_{-0.15}\right)$	$^{+0.14}_{-0.13}$ $\left(^{+0.13}_{-0.12}\right)$	$^{+0.15}_{-0.07}$ $\left(^{+0.11}_{-0.05}\right)$

ATLAS + CMS ttH(bb) systematics





ATLAS tt+HF systematics

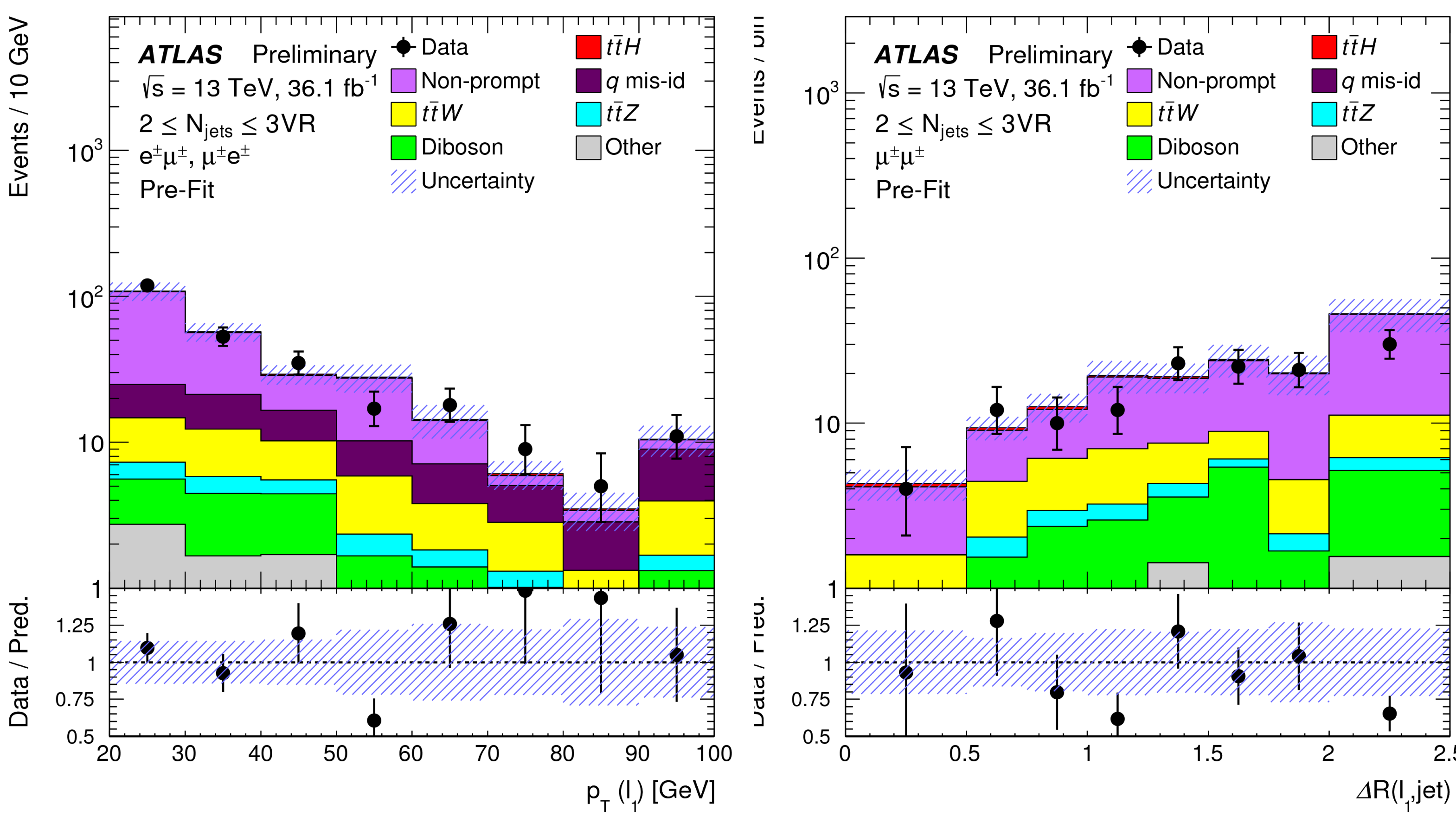


Systematic source	Description	$t\bar{t}$ categories
$t\bar{t}$ cross-section	Up or down by 6%	All, correlated
$k(t\bar{t} + \geq 1c)$	Free-floating $t\bar{t} + \geq 1c$ normalization	$t\bar{t} + \geq 1c$
$k(t\bar{t} + \geq 1b)$	Free-floating $t\bar{t} + \geq 1b$ normalization	$t\bar{t} + \geq 1b$
Sherpa5F vs. nominal	Related to the choice of NLO event generator	All, uncorrelated
PS & hadronization	Powheg+Herwig 7 vs. Powheg+Pythia 8	All, uncorrelated
ISR / FSR	Variations of μ_R, μ_F, h_{damp} and A14 Var3c parameters	All, uncorrelated
$t\bar{t} + \geq 1c$ ME vs. inclusive	MG5_aMC@NLO+Herwig++: ME prediction (3F) vs. incl. (5F)	$t\bar{t} + \geq 1c$
$t\bar{t} + \geq 1b$ Sherpa4F vs. nominal	Comparison of $t\bar{t} + b\bar{b}$ NLO (4F) vs. Powheg+Pythia 8 (5F)	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ renorm. scale	Up or down by a factor of two	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ resumm. scale	Vary μ_Q from $H_T/2$ to μ_{CMMPS}	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ global scales	Set μ_Q, μ_R, and μ_F to μ_{CMMPS}	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ shower recoil scheme	Alternative model scheme	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ PDF (MSTW)	MSTW vs. CT10	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ PDF (NNPDF)	NNPDF vs. CT10	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ UE	Alternative set of tuned parameters for the underlying event	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 1b$ MPI	Up or down by 50%	$t\bar{t} + \geq 1b$
$t\bar{t} + \geq 3b$ normalization	Up or down by 50%	$t\bar{t} + \geq 1b$

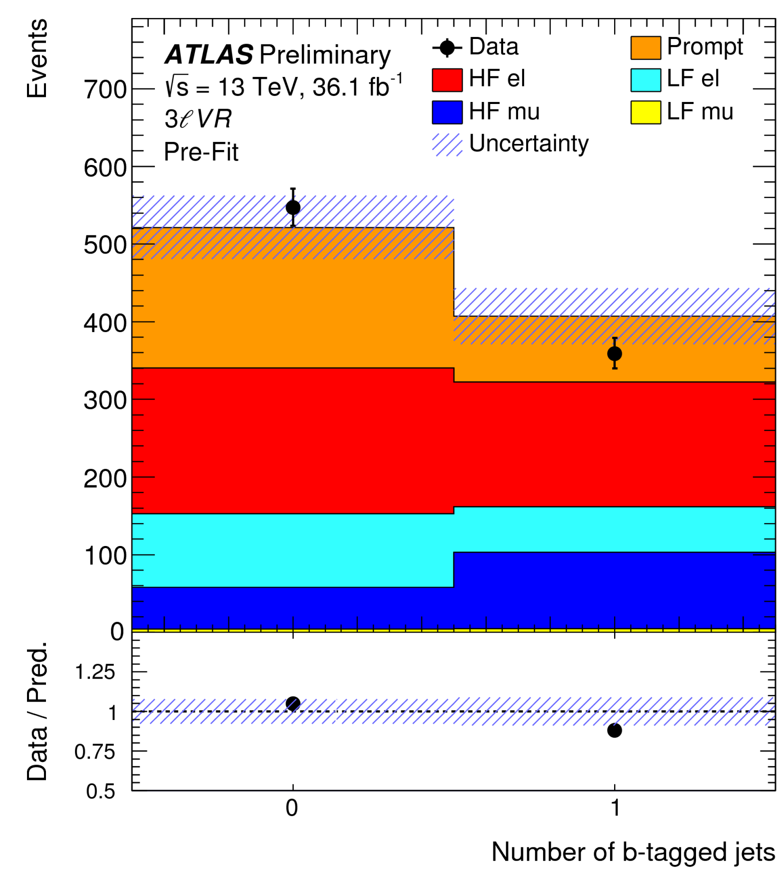
ttHML: light lepton fakes

	2ℓSS	3ℓ	4ℓ	$1\ell+2\tau_{had}$	$2\ell SS+1\tau_{had}$	$2\ell OS+1\tau_{had}$	$3\ell+1\tau_{had}$
Non-prompt lepton strategy	DD (MM)	DD (MM)	semi-DD (SF)	MC	DD (FF)	MC	MC
Fake tau strategy	–	–	–	DD (SS data)	semi-DD (SF)	DD (FF)	semi-DD (SF)

2lSS/3l



4l

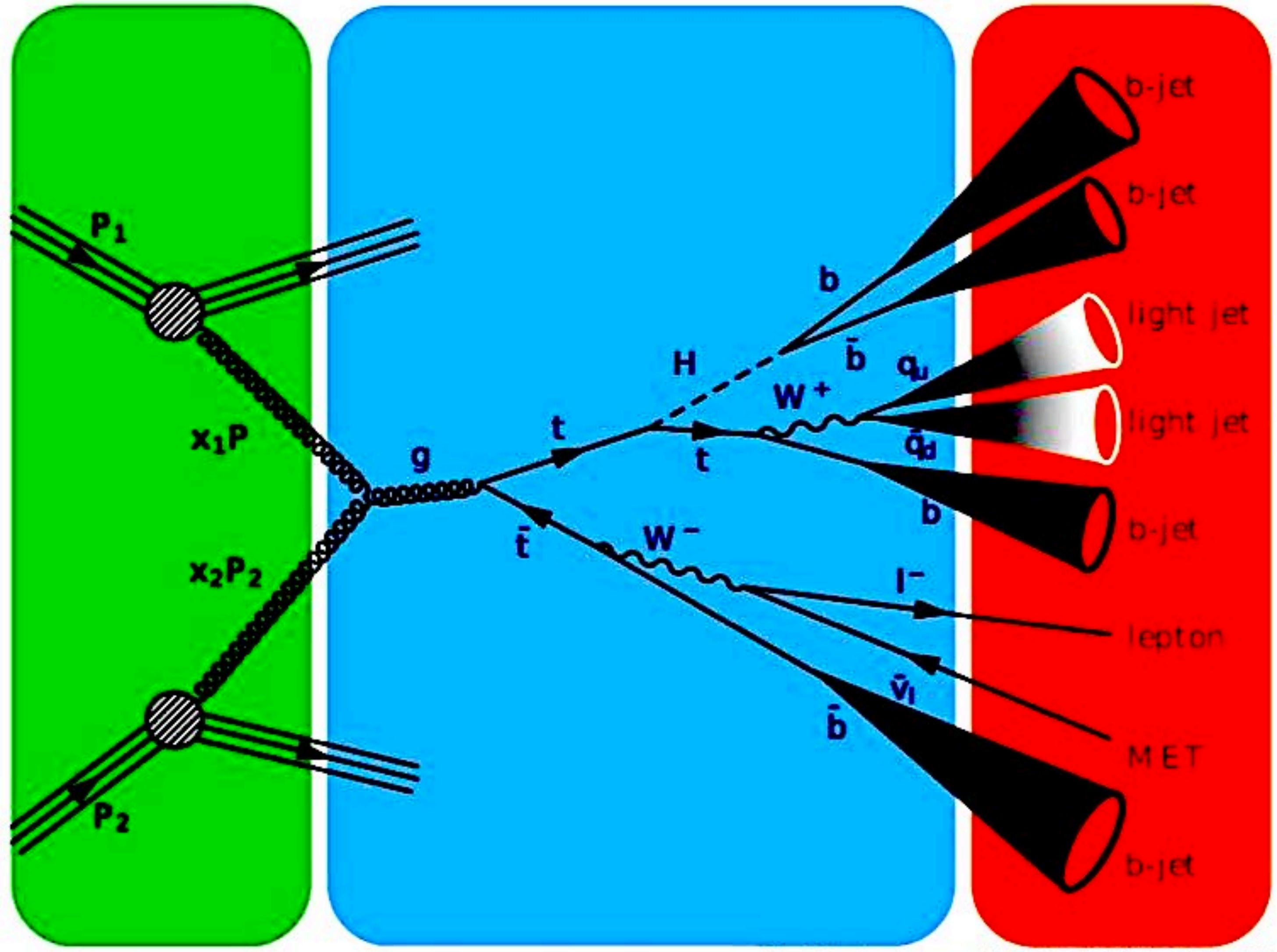


ttHML: Uncertainties

Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modeling (cross section)	+0.20	−0.09
Jet energy scale and resolution	+0.18	−0.15
Non-prompt light-lepton estimates	+0.15	−0.13
Jet flavor tagging and τ_{had} identification	+0.11	−0.09
$t\bar{t}W$ modeling	+0.10	−0.09
$t\bar{t}Z$ modeling	+0.08	−0.07
Other background modeling	+0.08	−0.07
Luminosity	+0.08	−0.06
$t\bar{t}H$ modeling (acceptance)	+0.08	−0.04
Fake τ_{had} estimates	+0.07	−0.07
Other experimental uncertainties	+0.05	−0.04
Simulation sample size	+0.04	−0.04
Charge misassignment	+0.01	−0.01
Total systematic uncertainty	+0.39	−0.30

ttH(bb): Matrix Element method

- Calculate the likelihood of each event to originate from ttH or tt + bb
- Used in most powerful signal region 6j SR1



Olaf Nackenhorst, CERN-THESIS-2015-186

parton distribution function
accounts for production mechanism

transfer function
maps detector response to diagram

$$L_i = \sum_{\text{flavor}} \int \frac{f_1\left(x_1, Q^2\right) f_2\left(x_2, Q^2\right)}{|\vec{q}_1|\,|\vec{q}_2|} \left|\mathcal{M}_i\left(\mathbf{Y}\right)\right|^2 T(\mathbf{X};\mathbf{Y})\, \mathrm{d}\Phi_n \mathrm{d}\mathbf{Y}$$

matrix element
describes signal or background process