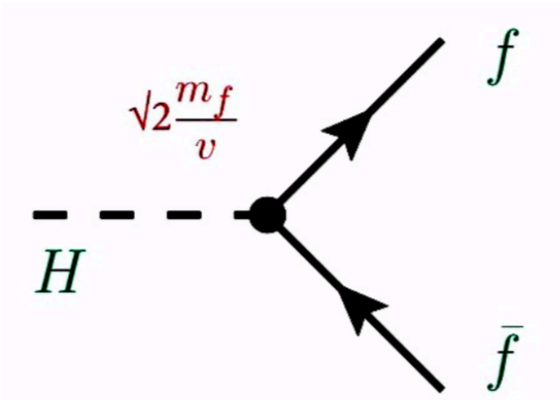
Measurement of ttH production in multileptonic final states at the ATLAS experiment using $36\ fb^{-1}$ data

Rohin Narayan
Southern Methodist University, Dallas

Phys. Rev. D 97 (2018) 072003

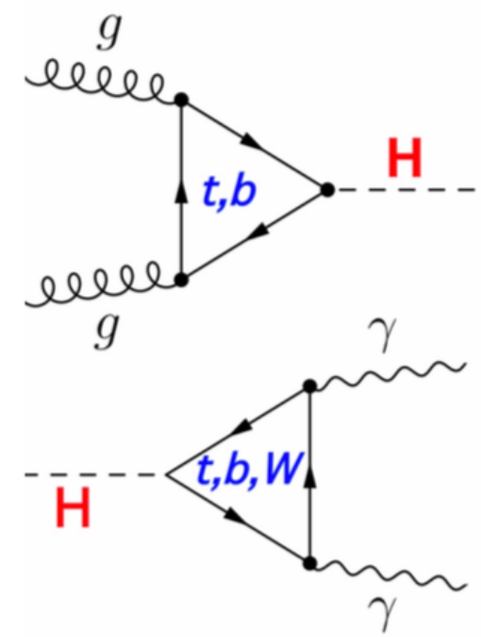
Motivation

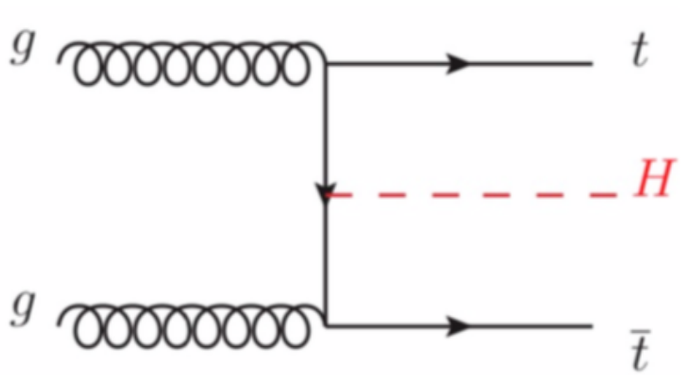


- Fermion masses are generated through Yukawa interaction
- Heaviest SM particle (top) expected to have largest Yukawa coupling (y_t) to the Higgs field.

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

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





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

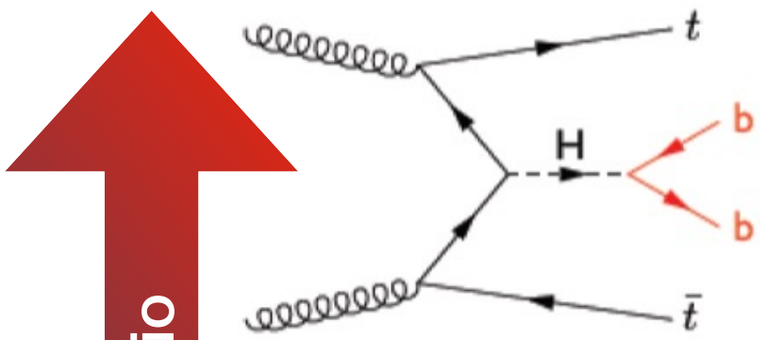
Experimental challenges

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

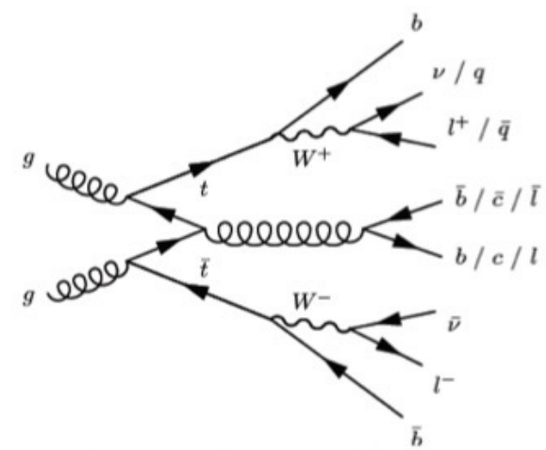
Signal

Background

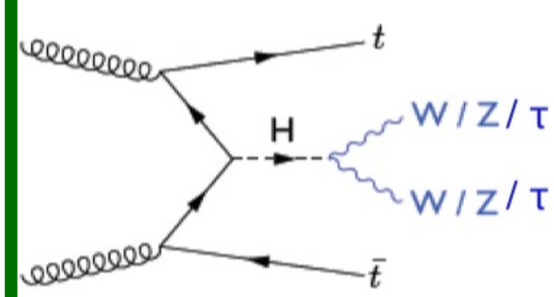
Higgs Branching ratio



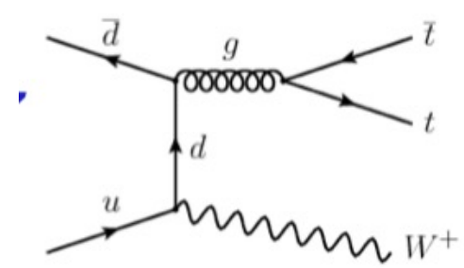
ttH(bb)

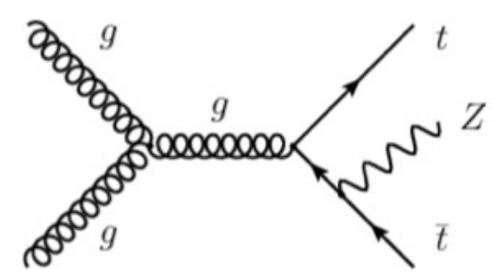


tt+(HF) jets (irreducible)

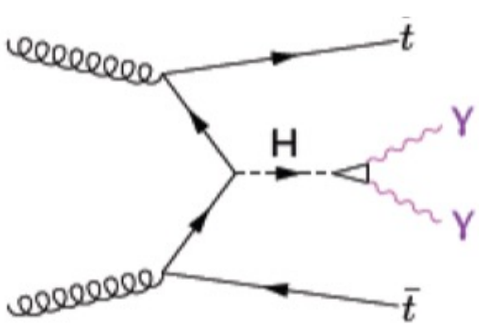


ttH-multileptons

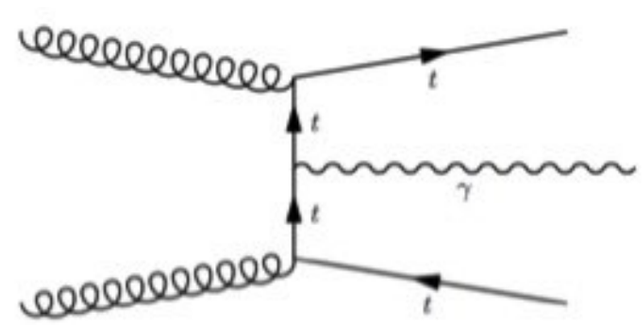




ttV



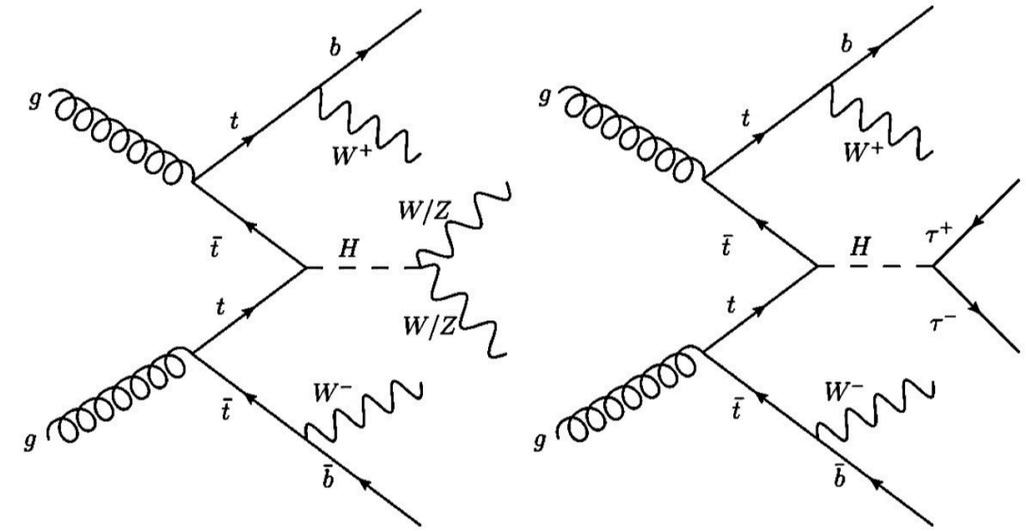
ttH(γγ)



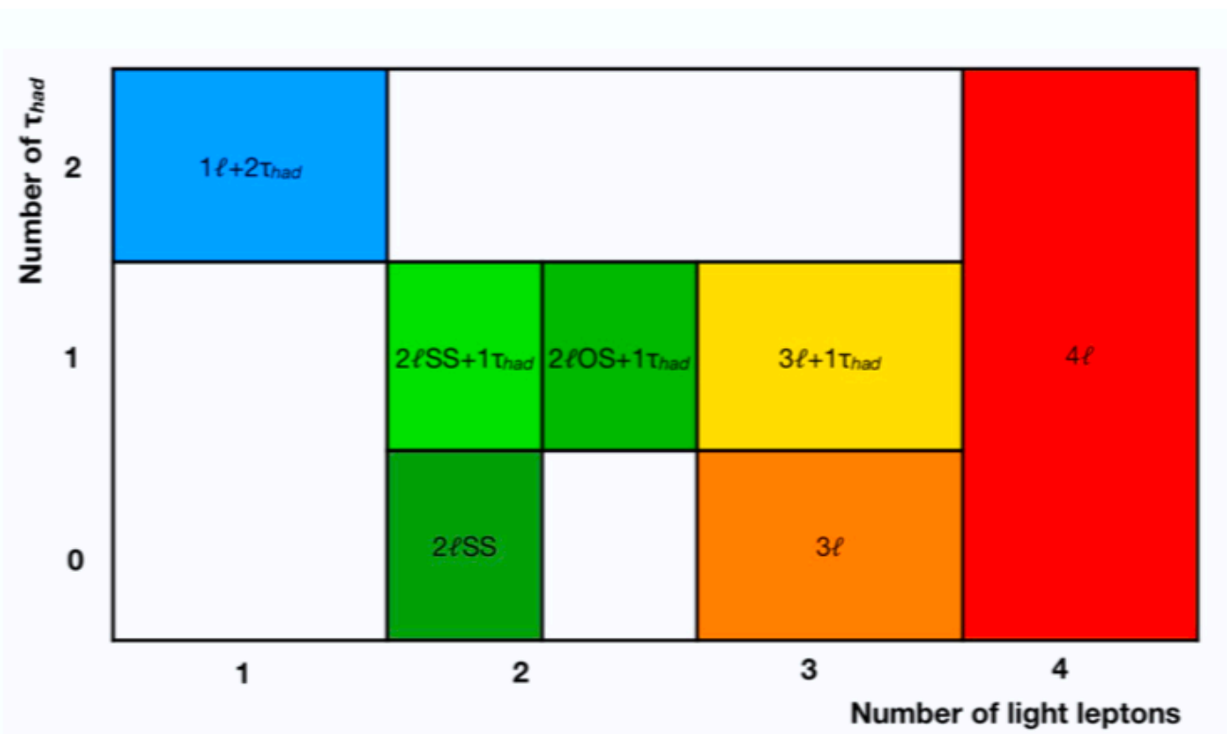
ttγ

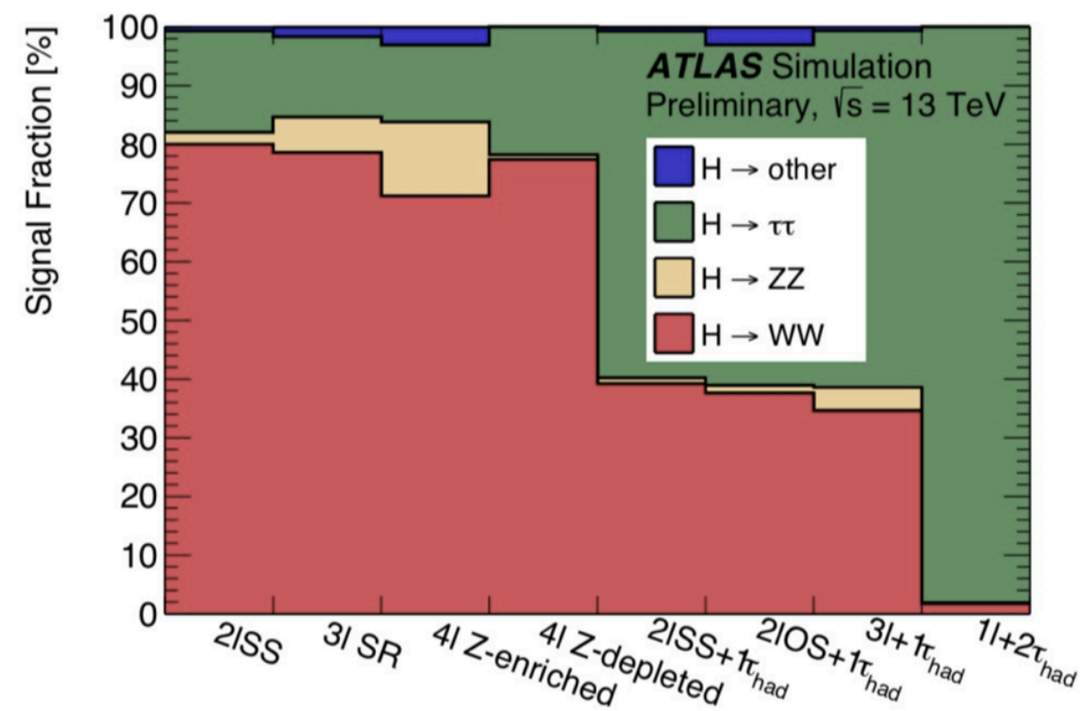
Analysis strategy

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



Decay modes in sub channels



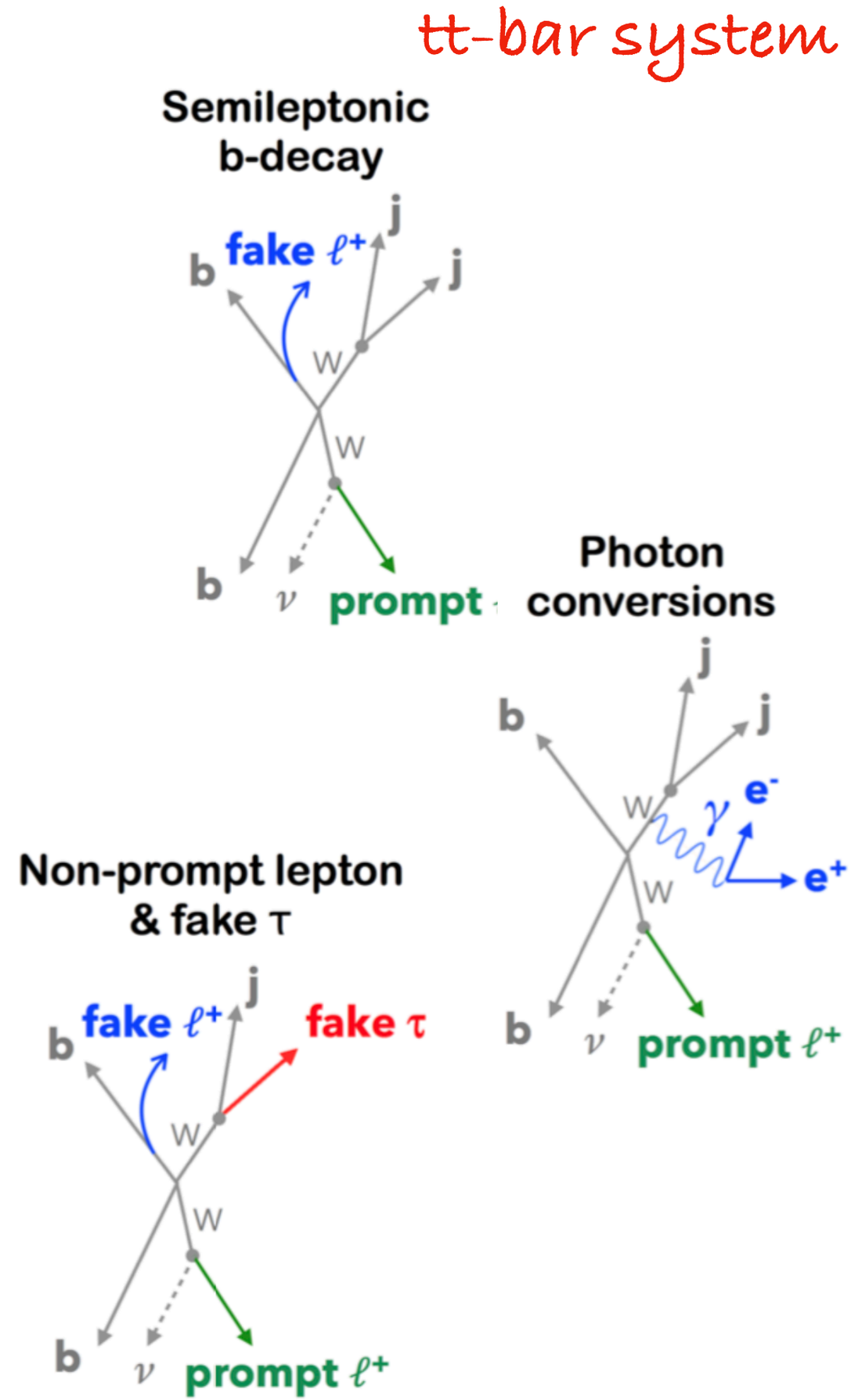


- Common jet selection $N_{\text{jet}} \geq 2$ and $N_{\text{bjet}} \geq 1$
- Optimized lepton selection in each category
  - A BDT algorithm to suppress fakes from B-hadron decay

- Light leptons channels dominated by $H \to WW^*$
- Tau channels dominated by $H \to \tau^+\tau^-$

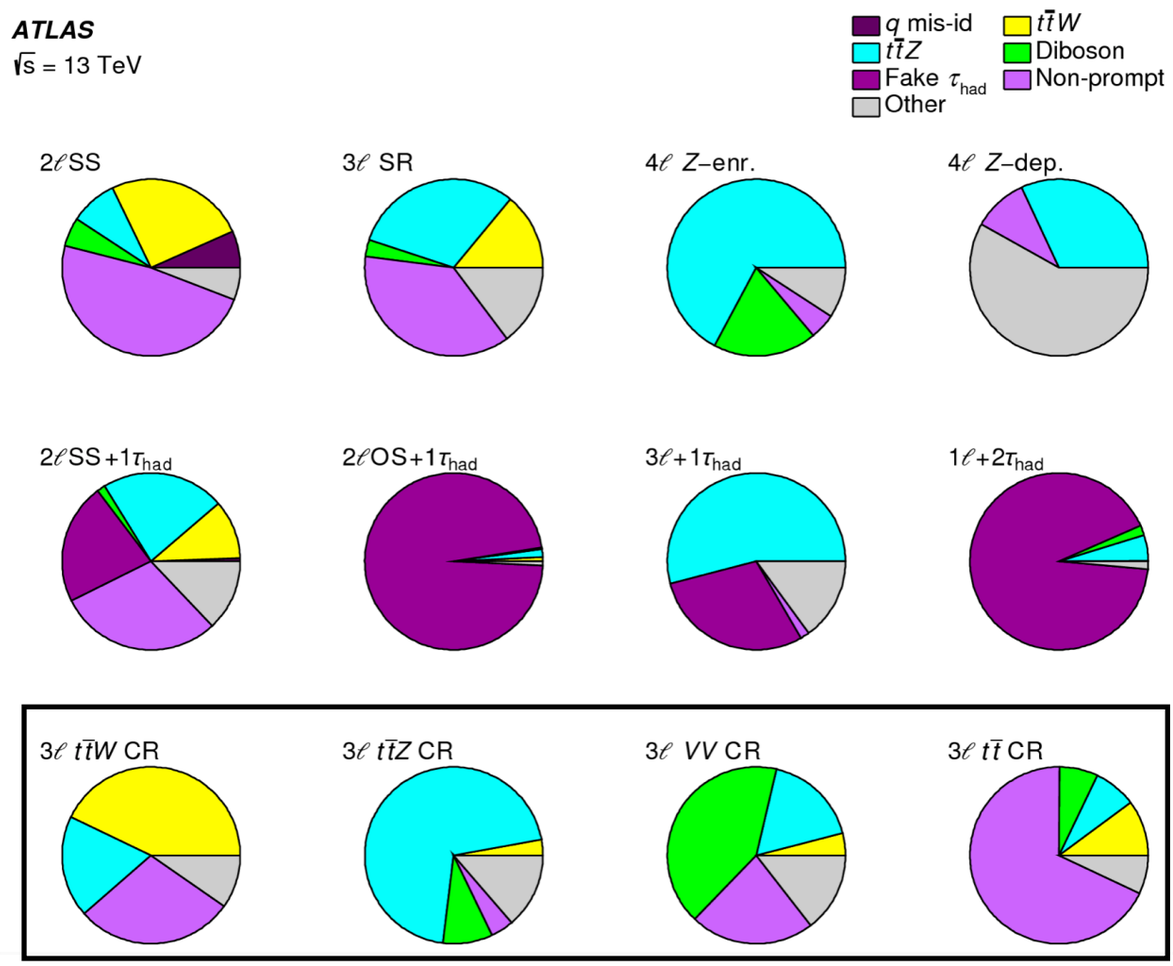
Background Estimation

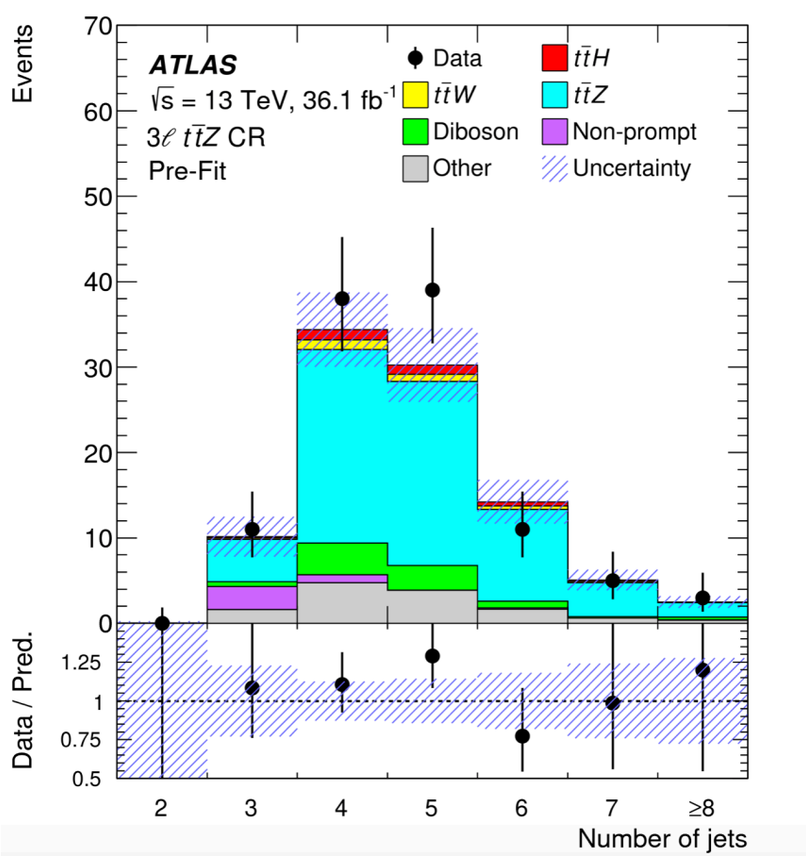
- Prompt-leptons and hadronic taus
  - Estimated with MonteCarlo
- Electron charge mis-identification for 2lSS and 2lSS+1τ-had
  - Datadriven estimation: Charge-misid rates in $Z \rightarrow e^+e^-/\mu^+\mu^-$
- Light-lepton and non-prompt lepton fakes
  - Source: Semileptonic b-decay, photon conversions
- Hadronic-tau fakes
  - Source: light flavor jets and mis-reconstructed electrons.

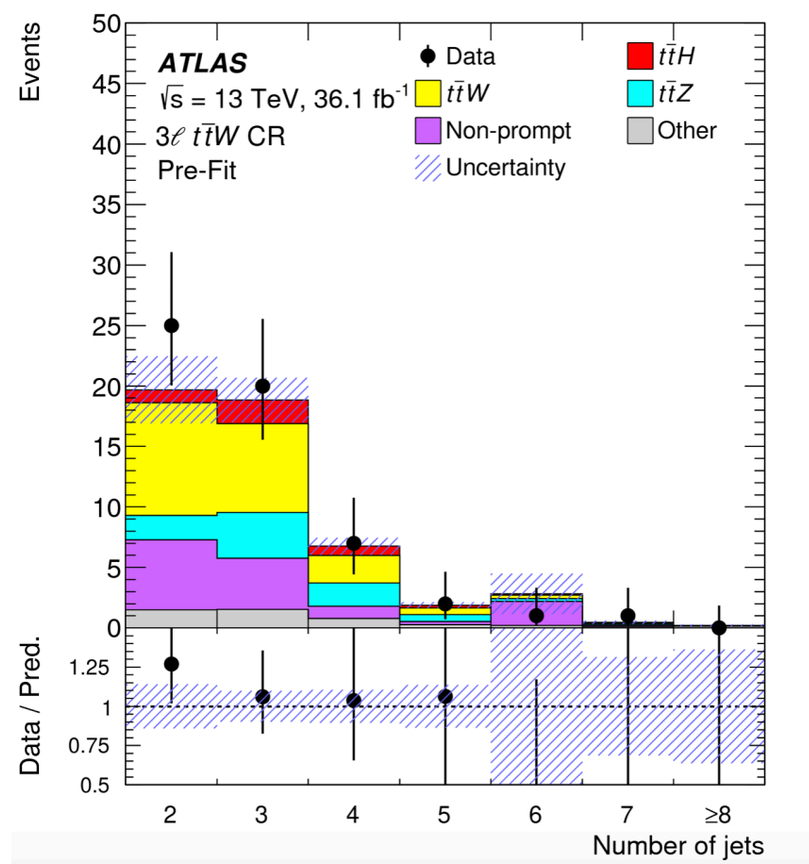


Background: Prompt leptons

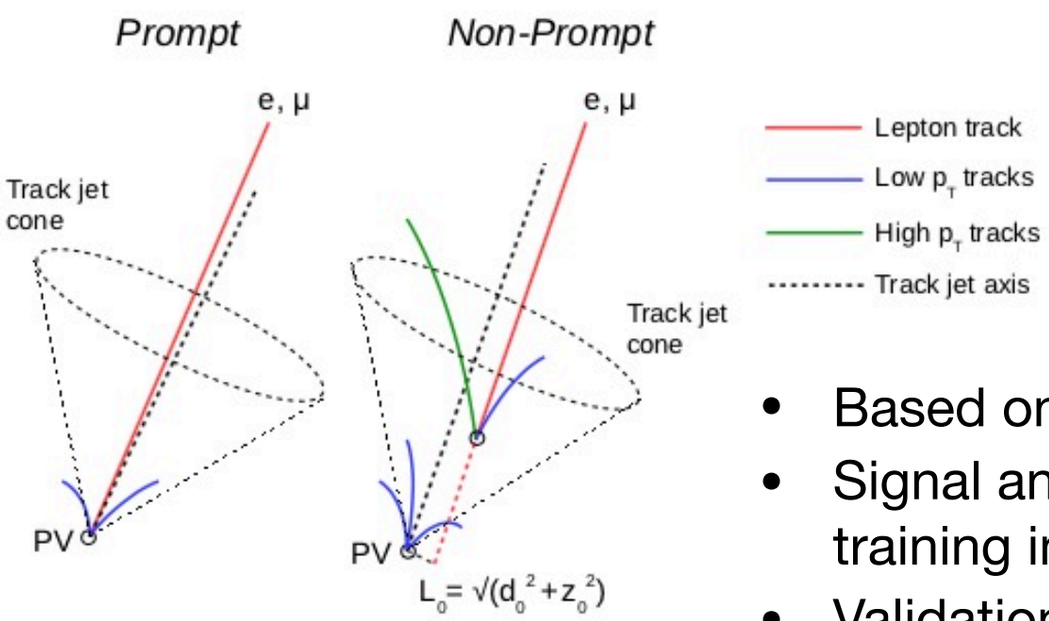
- Largest prompt backgrounds: $t\bar{t}W$, $t\bar{t}Z$ and VV
- Modeled using NLO MC generators (MG5aMC@NLO, Sherpa)
  - Systematics: Cross-Section uncertainties, scale-variations, alternate generator comparisons
- Validated in dedicated validation regions







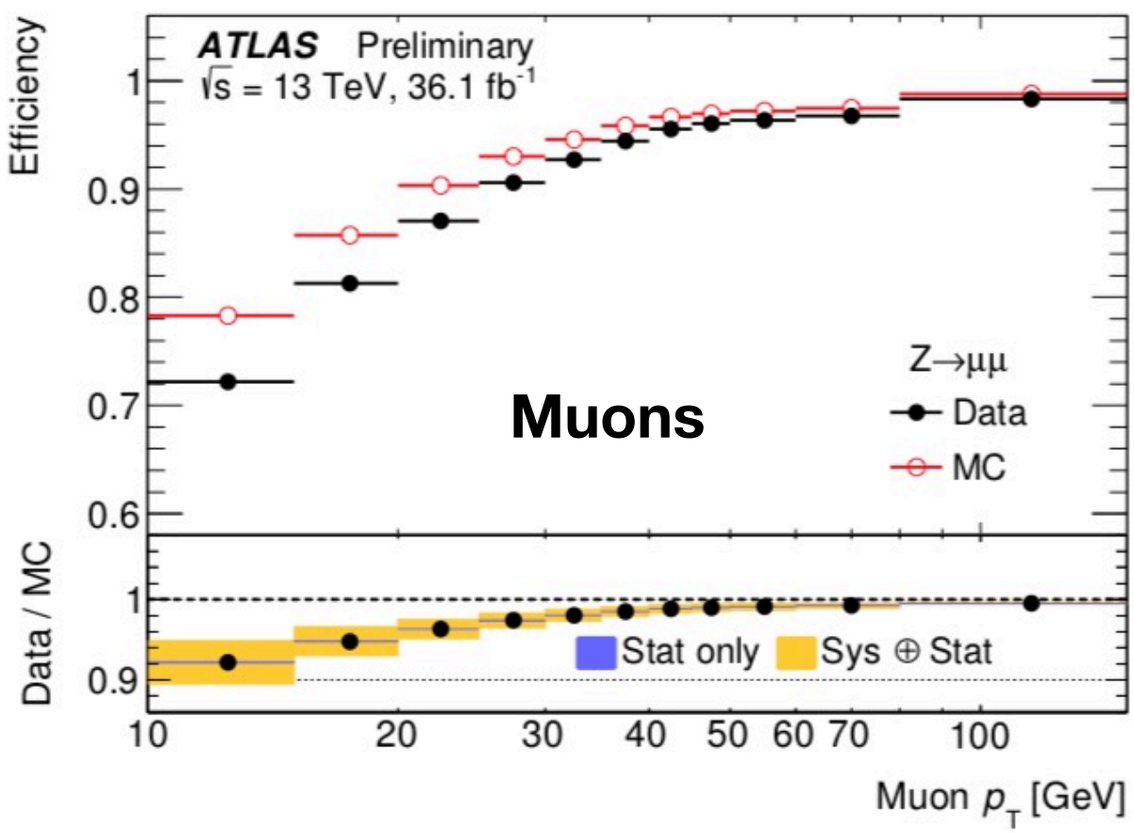
Background: Non prompt leptons

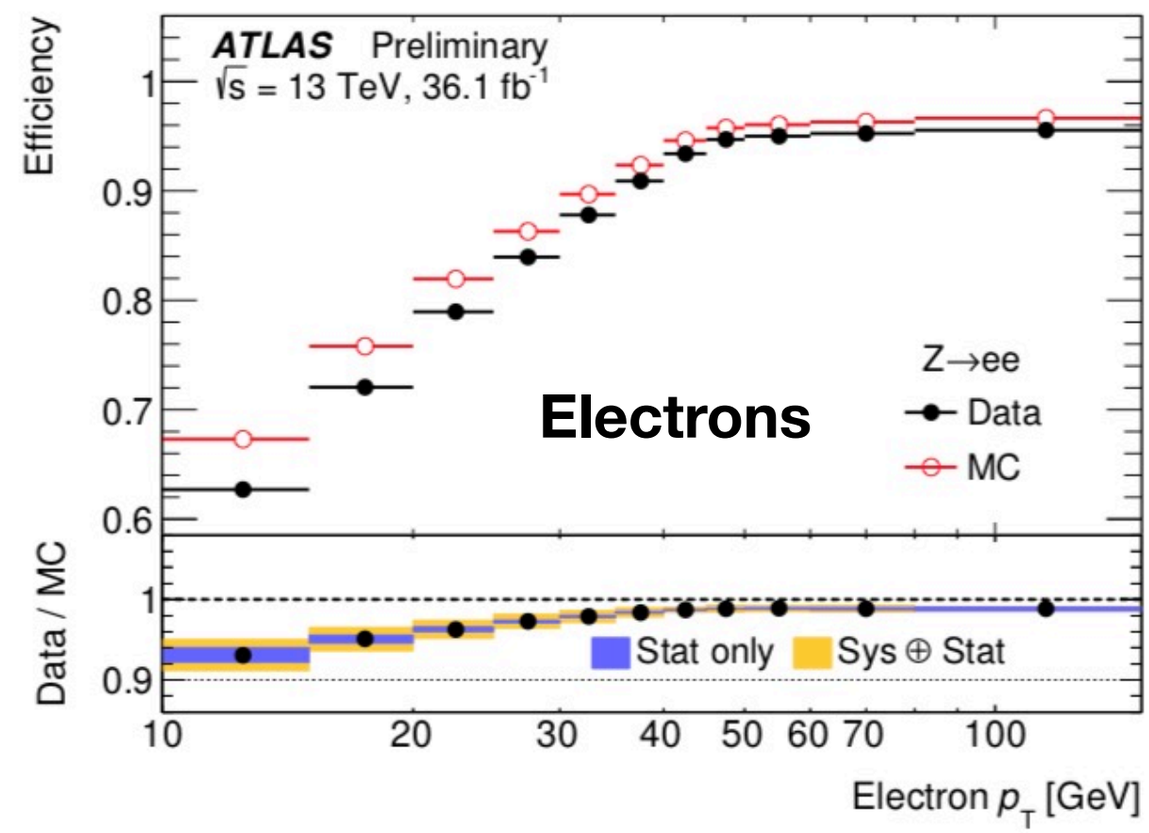


- **A dedicated BDT algorithm to suppress Heavy flavor lepton fakes**

- Based on lepton track associated to track-jets
- Signal and Background classification and training in MonteCarlo
- Validation in data using Z tag & Probe

Variable
N_{track} in track jet
IP2 $\log(P_b/P_{\text{light}})$
IP3 $\log(P_b/P_{\text{light}})$
N_{TrkAtVtx} SV + JF
$p_T^{\text{lepton}}/p_T^{\text{track jet}}$
$\Delta R(\text{lepton, track jet})$
$p_T\text{VarCone30}/p_T$
$E_T\text{TopoCone30}/p_T$

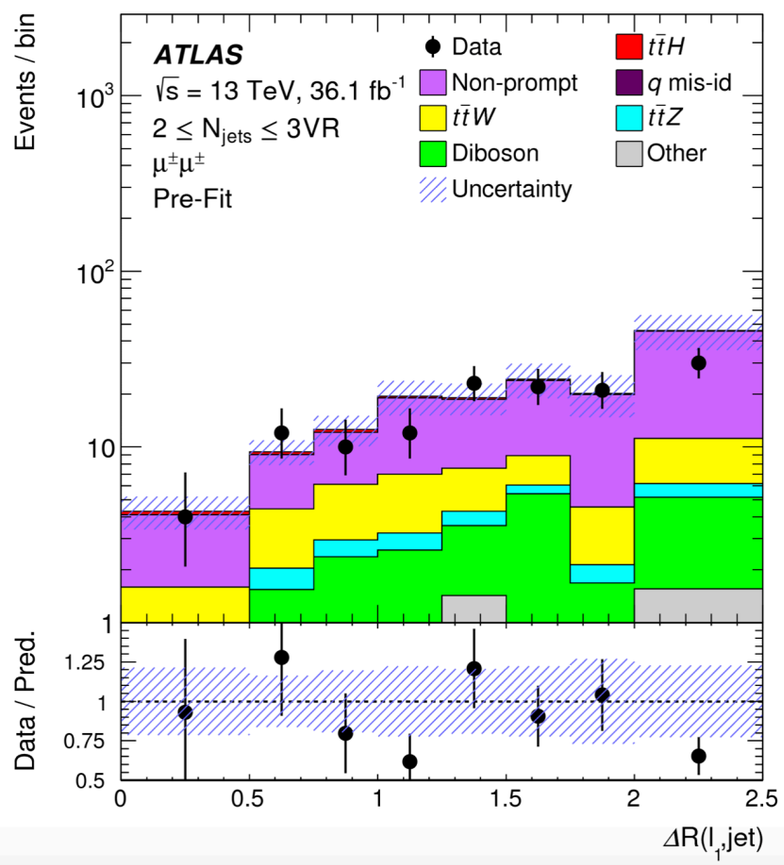


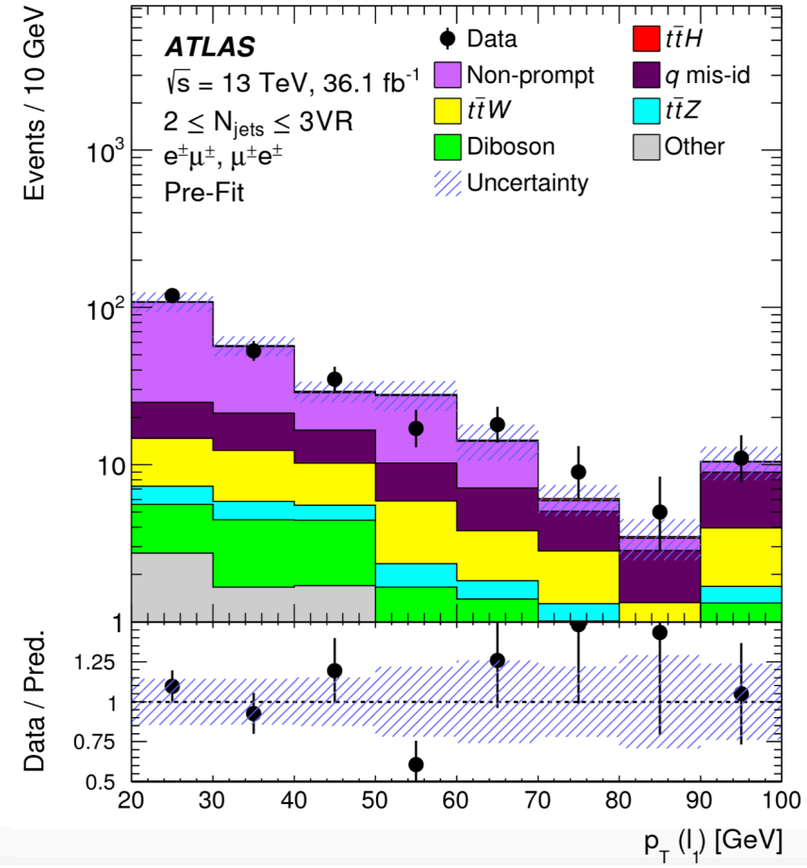


Background: Light lepton fakes

2l/3l

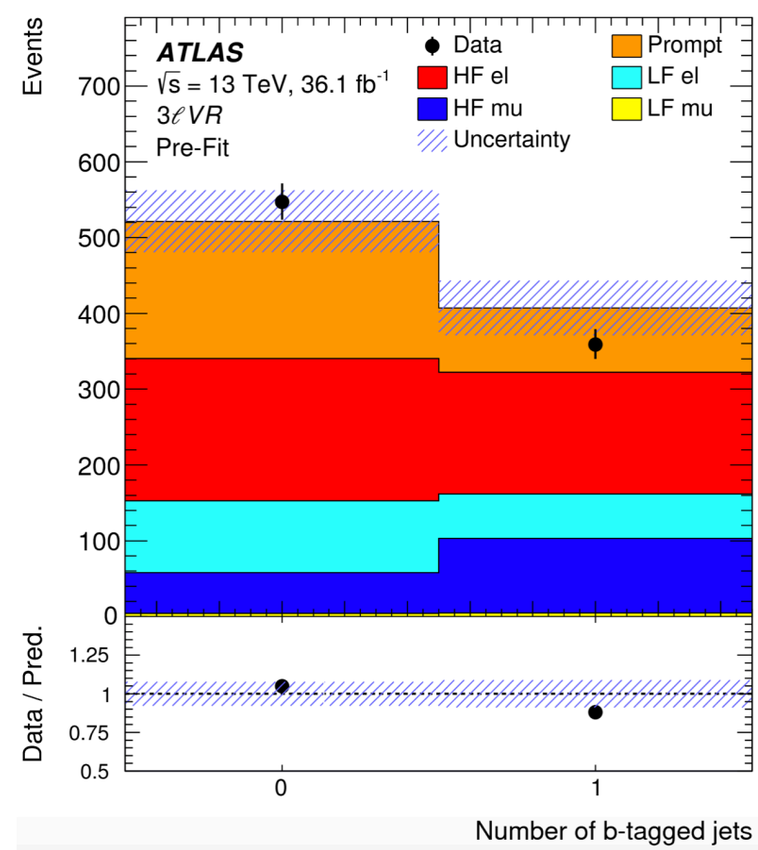
- Fully data driven method using matrix method.
- Real and fake efficiencies are measured in data in a loose control region and extrapolated to the signal region.
- Method validated in various validation regions





4l

- Semi-data driven method in 4l-channel.
- MC split into prompt, heavy and light
- Scale factor determined for each of the components in 3l low-njet CR and extrapolated to 4l region



Background: Tau Fakes

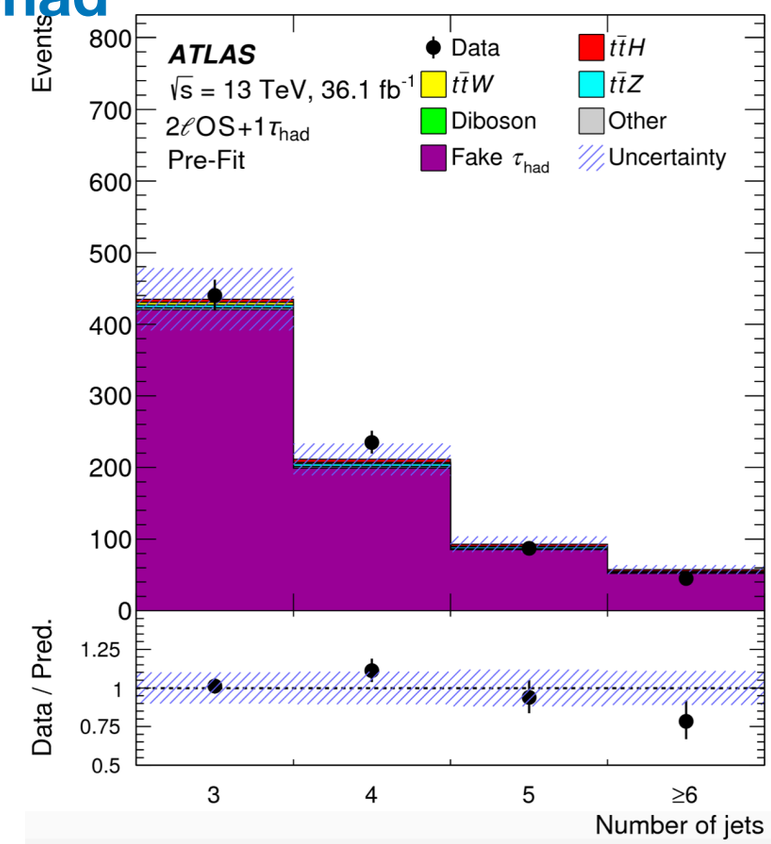
- Relevant for all SR's which has a hadronic tau.
  - Light lepton fakes are negligible except in 2lSS+1τ-had

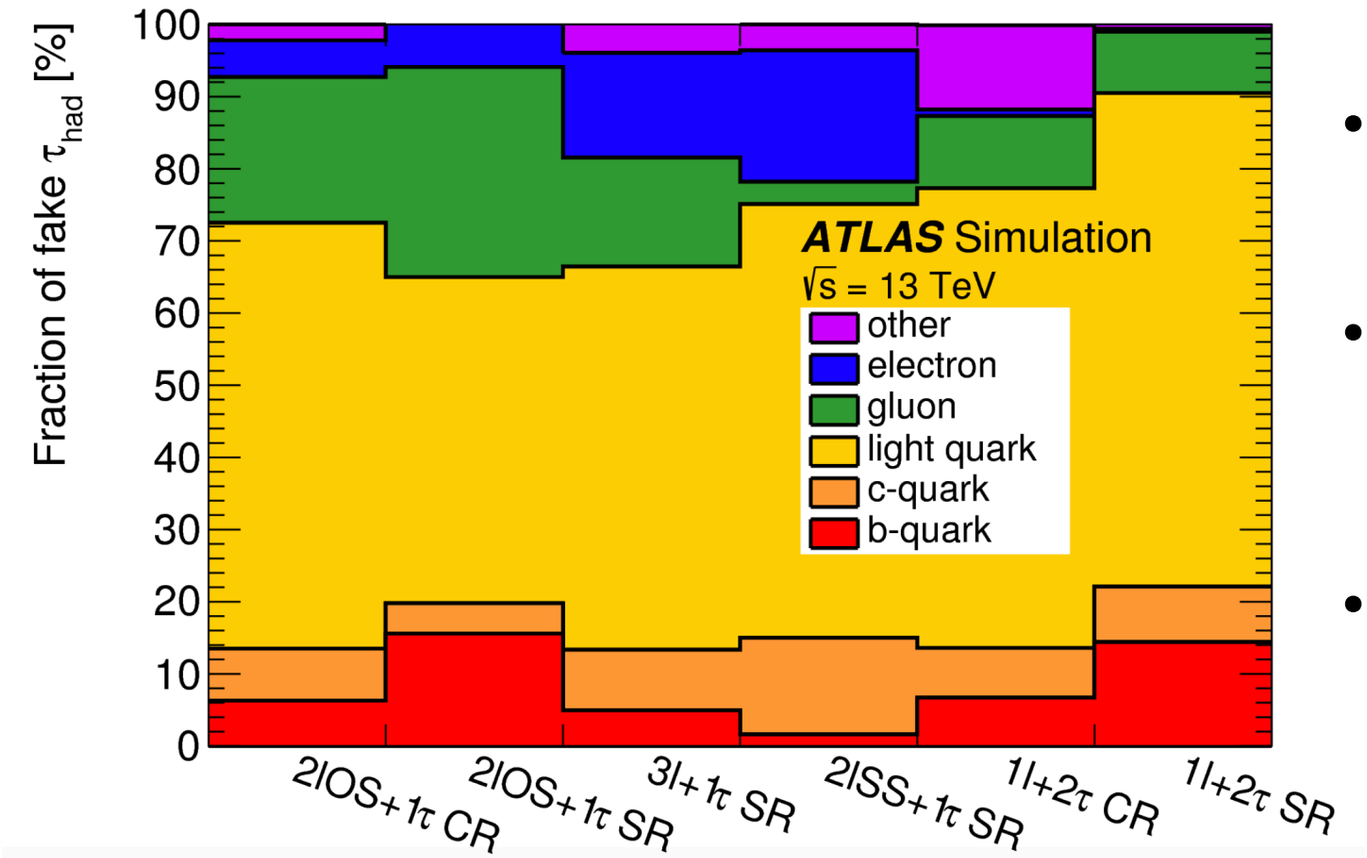
1l + 2τ-had

- Contains two OS τ.
- Fakes are estimated in a SS CR and extrapolated to SR (MC based)

2lOS+1τ-had

- Fake factor method: Measured in a control region inverting part of τ identification.
- Good data/MC agreement.

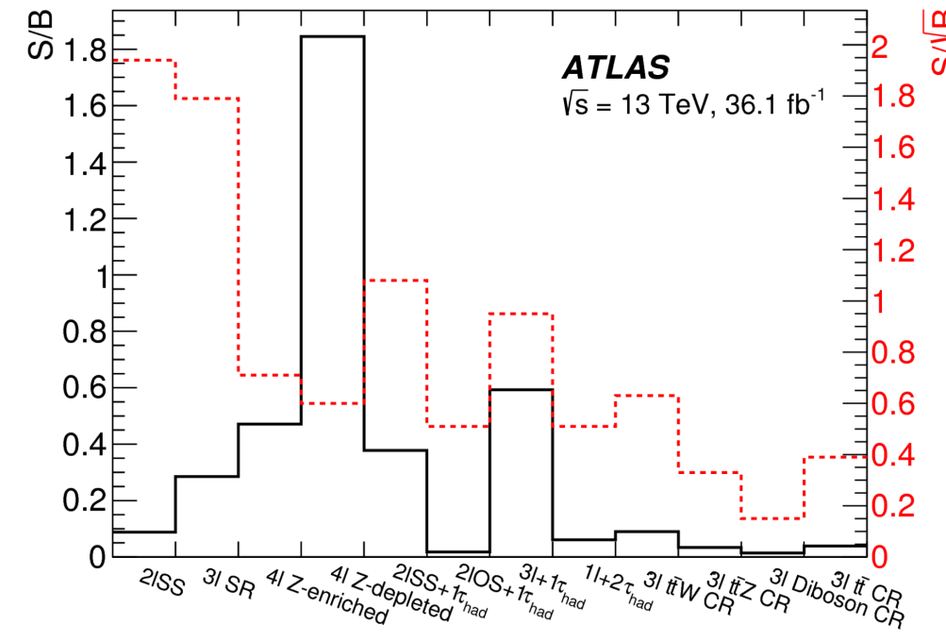




- Quite similar composition of fakes in 2lOS+1τ-had, 2lSS+1τ-had, 3l+1τ-had
- Data-driven estimates from 2lOS+1τ-had used as a correction factor for τ-had fakes in those channels.
- Light-lepton fakes in 2lSS+1τ-had estimated using fake factor method.

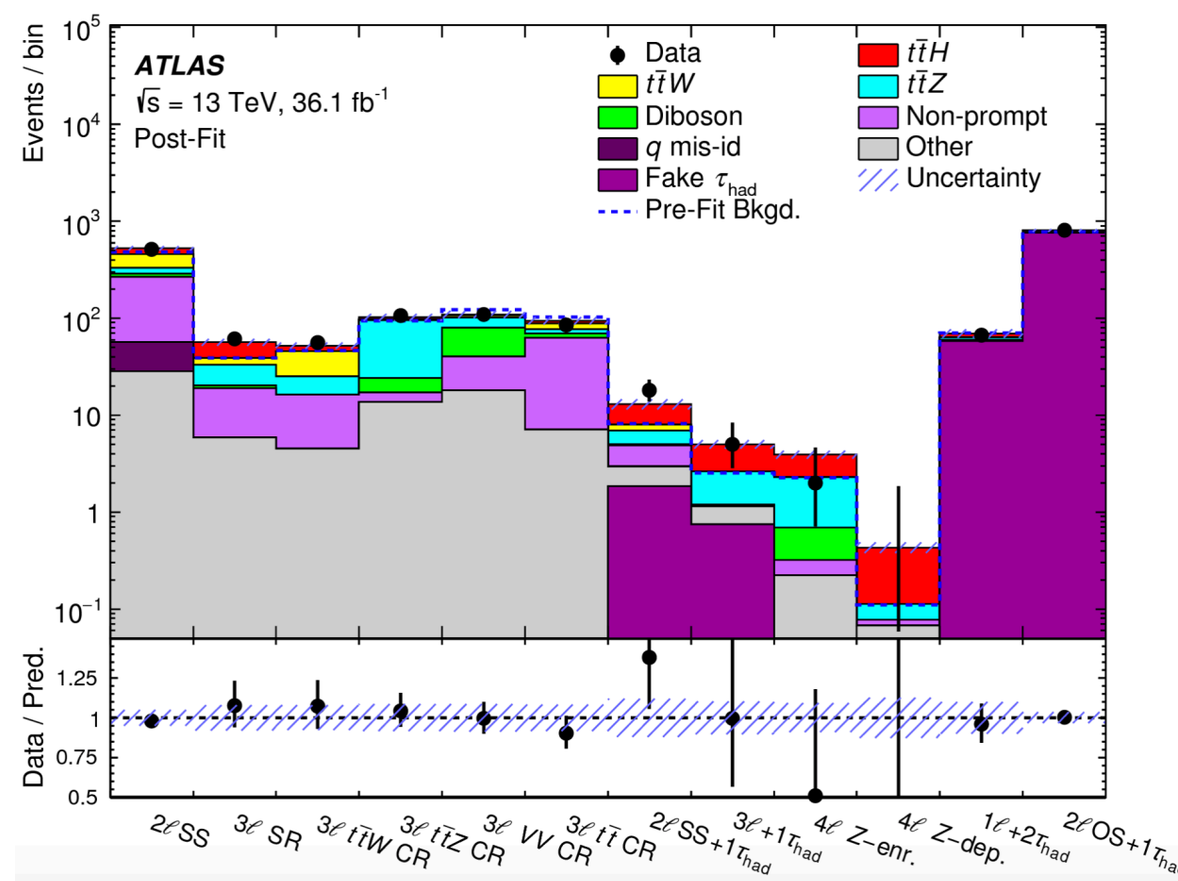
Fit model

- Except for 3l+1τ_{had} all sub channels use MVA techniques to further separate signal from background
  - 2lSS0τ: 2 BDT's: against ttV and ttbar
  - 3l0τ: 5 dimensional multinomial BDT to build categories enriched in ttH, ttW, ttZ, ttbar, diboson
  - 2lss+1τ: 1 BDT against ttbar
  - Cut and count cross check analysis in high sensitivity sub-channels.

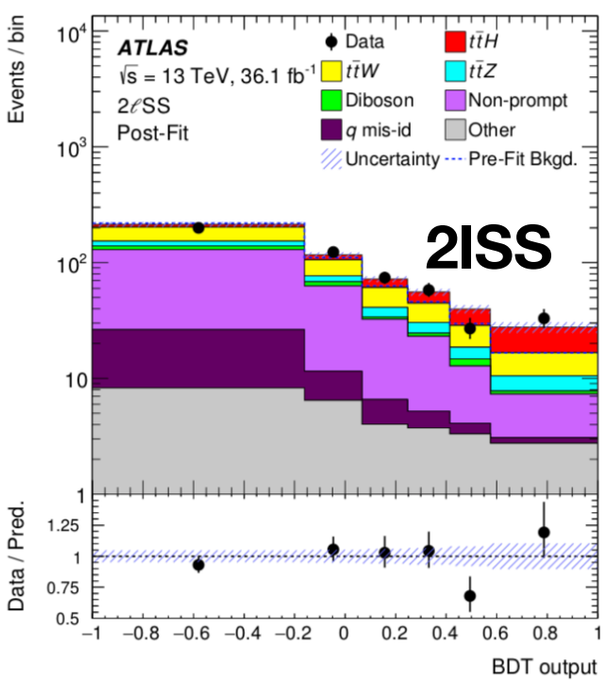


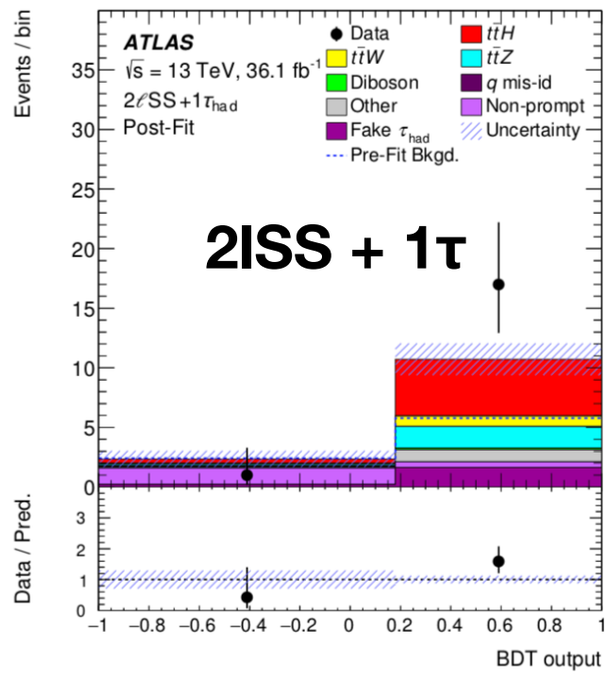
	2ℓSS	3ℓ	4ℓ	1ℓ+$2\tau_{\text{had}}$	2ℓSS+$1\tau_{\text{had}}$	2ℓOS+$1\tau_{\text{had}}$	3ℓ+$1\tau_{\text{had}}$
BDT trained against	Fakes and $t\bar{t}V$	$t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, VV	$t\bar{t}Z$ / -	$t\bar{t}$	all	$t\bar{t}$	-
Discriminant	2×1D BDT	5D BDT	Event count	BDT	BDT	BDT	Event count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

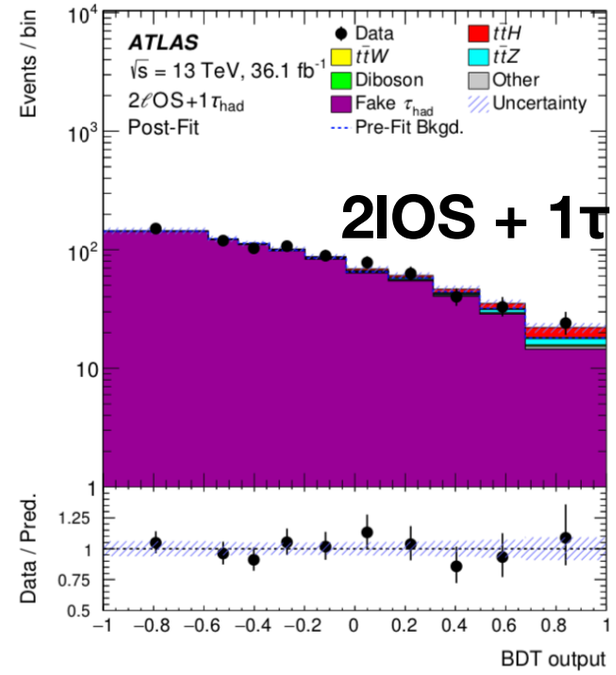
- Simultaneous fit in 12 regions (CR+SR)
- Single bin used in 3l CR's as well as low stat SR's
  - 3l+1τ
  - 4l (z-enriched,z-depeleted)
- BDT shape information used in 5 SR's

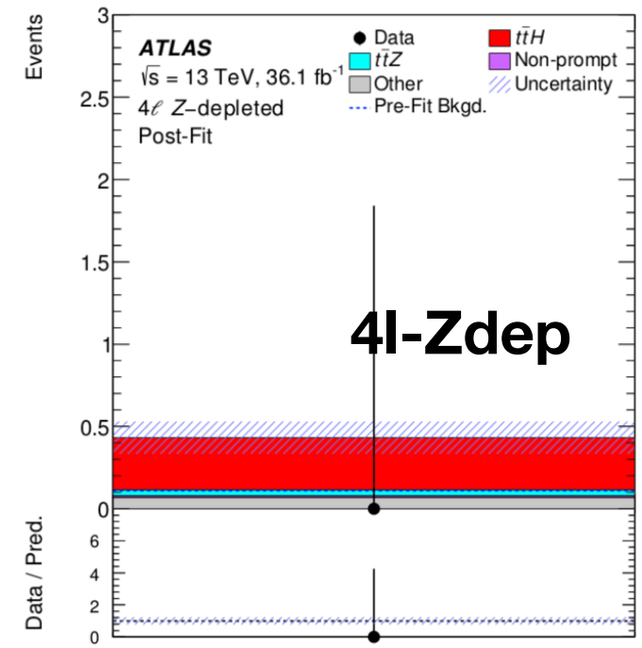


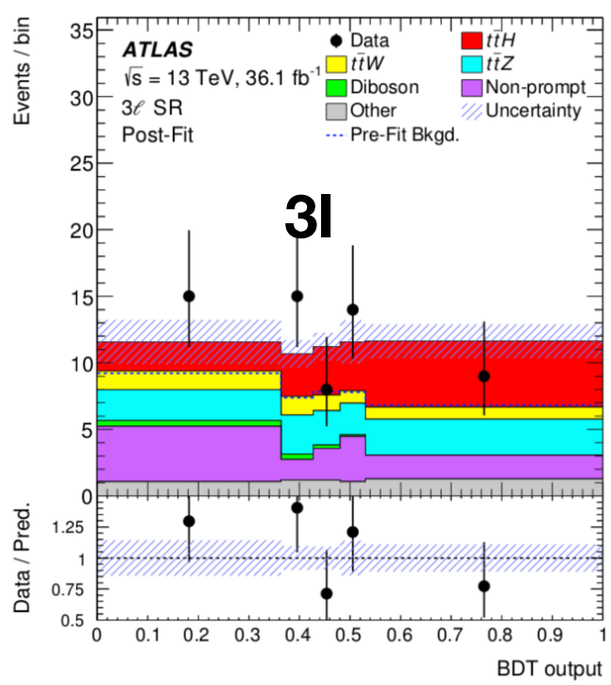
Post-Fit: Signal regions

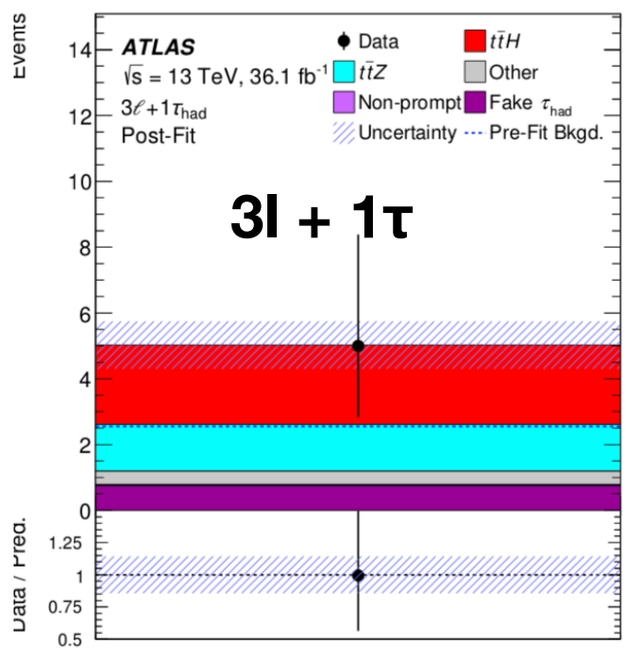


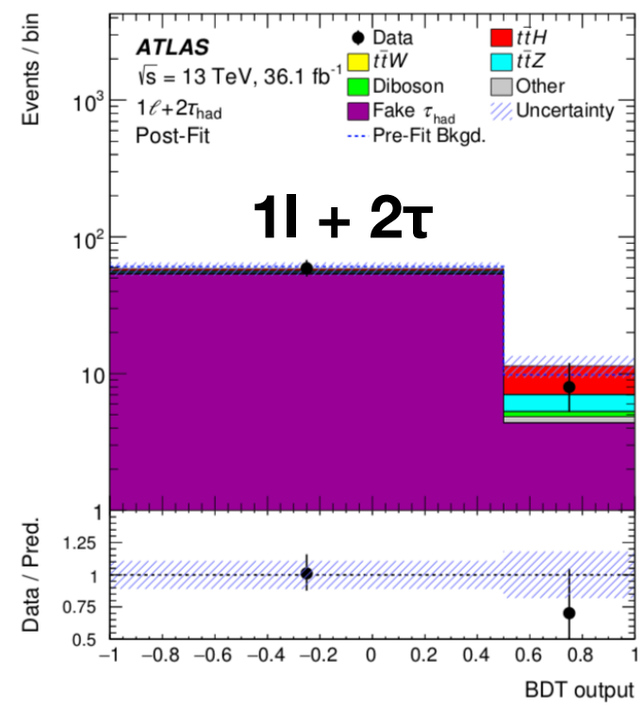


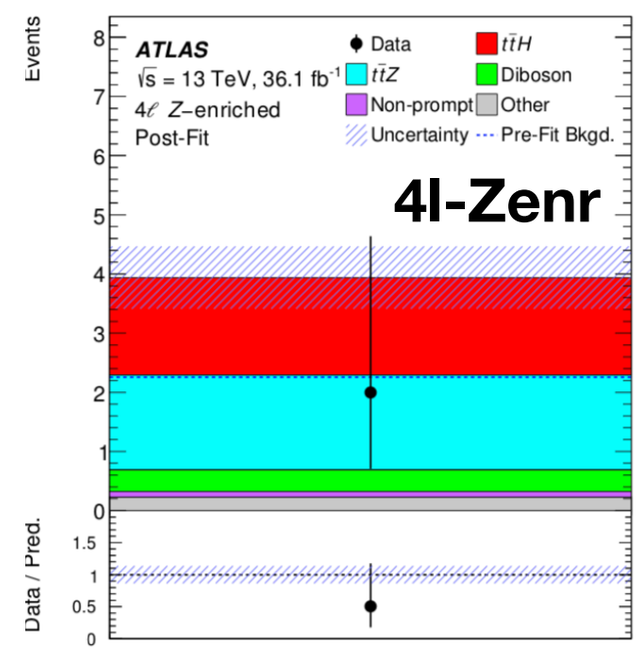




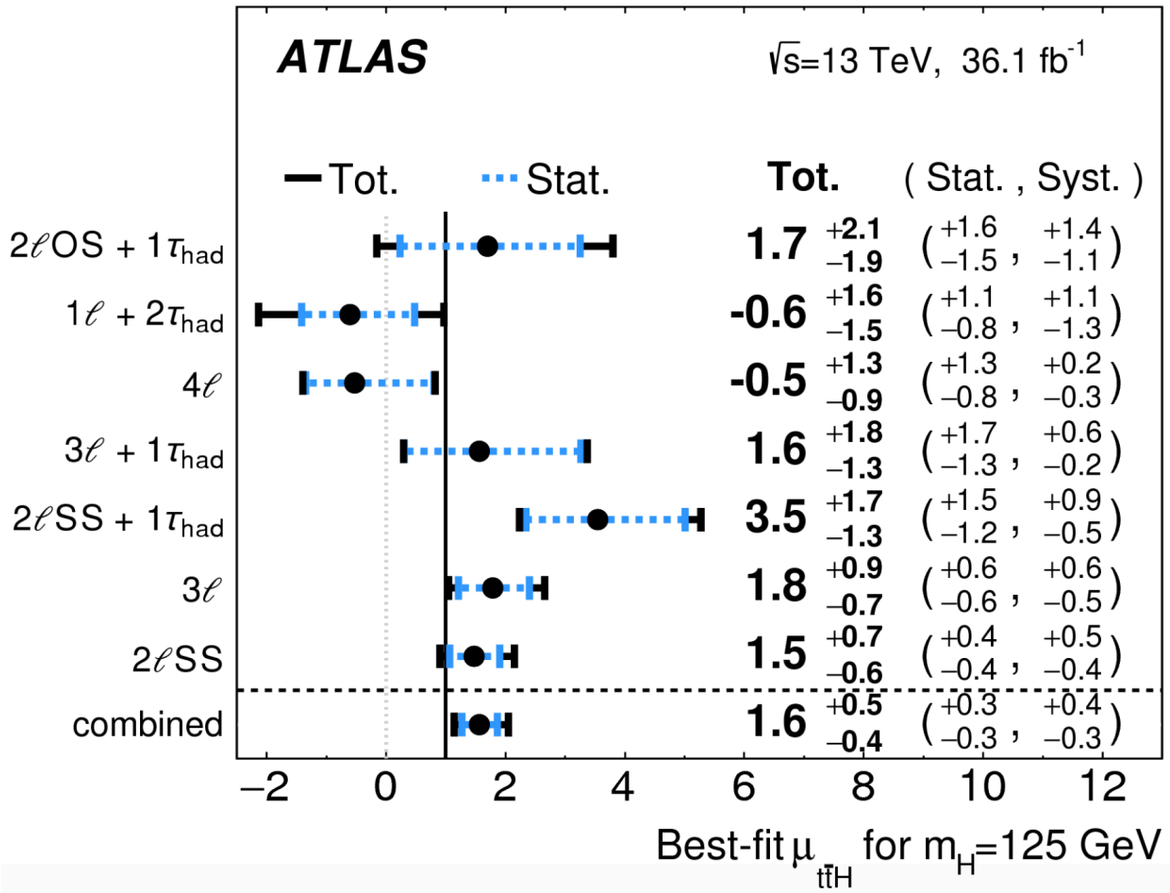








ttHML: Results



- Measured $\mu = 1.6^{+0.5}_{-0.4}$ with a significance with respect to background only hypothesis **4.1σ (expct. 2.8σ)**
- Cross section extrapolated to inclusive phase space

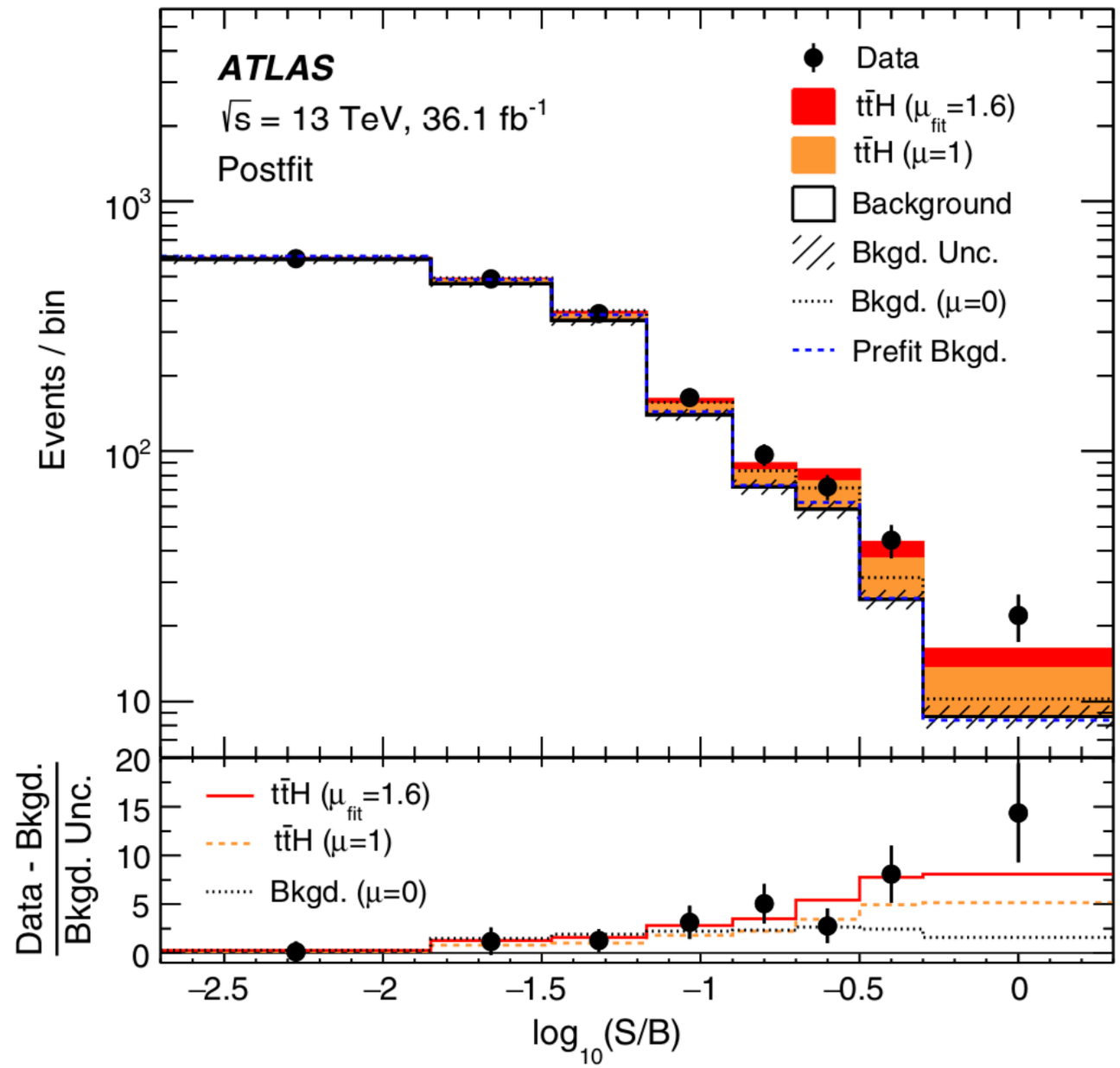
$$\sigma(t\bar{t}H) = 790^{+150}_{-150}(stat)^{+170}_{-150}(syst)\,fb$$

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

- Leading systematic uncertainty from ttH modeling followed by jet-energy scale and non-prompt light lepton estimates.

Conclusion

- Presented results of ttH-multileptons analysis measured at the ATLAS experiment using $36fb^{-1}$ of data collected in 2015-16
- Leading systematics that affect the analysis are ttH, ttV-modeling (theory), jet-energy scale, light-lepton fake estimation (experimental)
- In combination with other ttH-channels ATLAS observed ttH-production in pp collisions at the LHC
- The analysis is being updated to include newer data collected in latest data taking periods.



Backup

Object definitions and Selections

L = Loose
L* = Loose + isolated
T = Tight
M = Medium
T* = Tight + Charge mis-assignment veto

Leptons

- pT > 10 GeV; central η

	e					μ			
	L	L†	L*	T	T*	L	L†	L*	T
Isolation	No	Yes				No	Yes		
Non-prompt lepton BDT	No		Yes			No		Yes	
Identification	Loose			Tight		Loose			
Charge mis-assignment veto	No				Yes	N/A			
Tranverse impact parameter significance $\lvert d_0\rvert/\sigma_{d_0}$	< 5					< 3			
Longitudinal impact parameter $\lvert z_0 \sin\theta\rvert$	< 0.5 mm								

Trigger

- A combination of single and di-leptonic trigger

light & b-jets

pT > 25 GeV, |eta| < 2.5

b-tag WP: MV2c10 70%.

Hadronic Taus

Medium/Tight ID to reject jets

pT > 25 GeV

BDT to reject el faking taus

b-jet veto

tau vertex is PV

Overlap Removal

Keep	Remove	Cone size (Δ R)
electron	electron (low p_T)	0.1
muon	electron	0.1
electron	jet	0.3
jet	muon	min(0.4, 0.04 + 10[GeV]/p_T (muon))
electron	tau	0.2
muon	tau	0.2
tau	jet	0.3

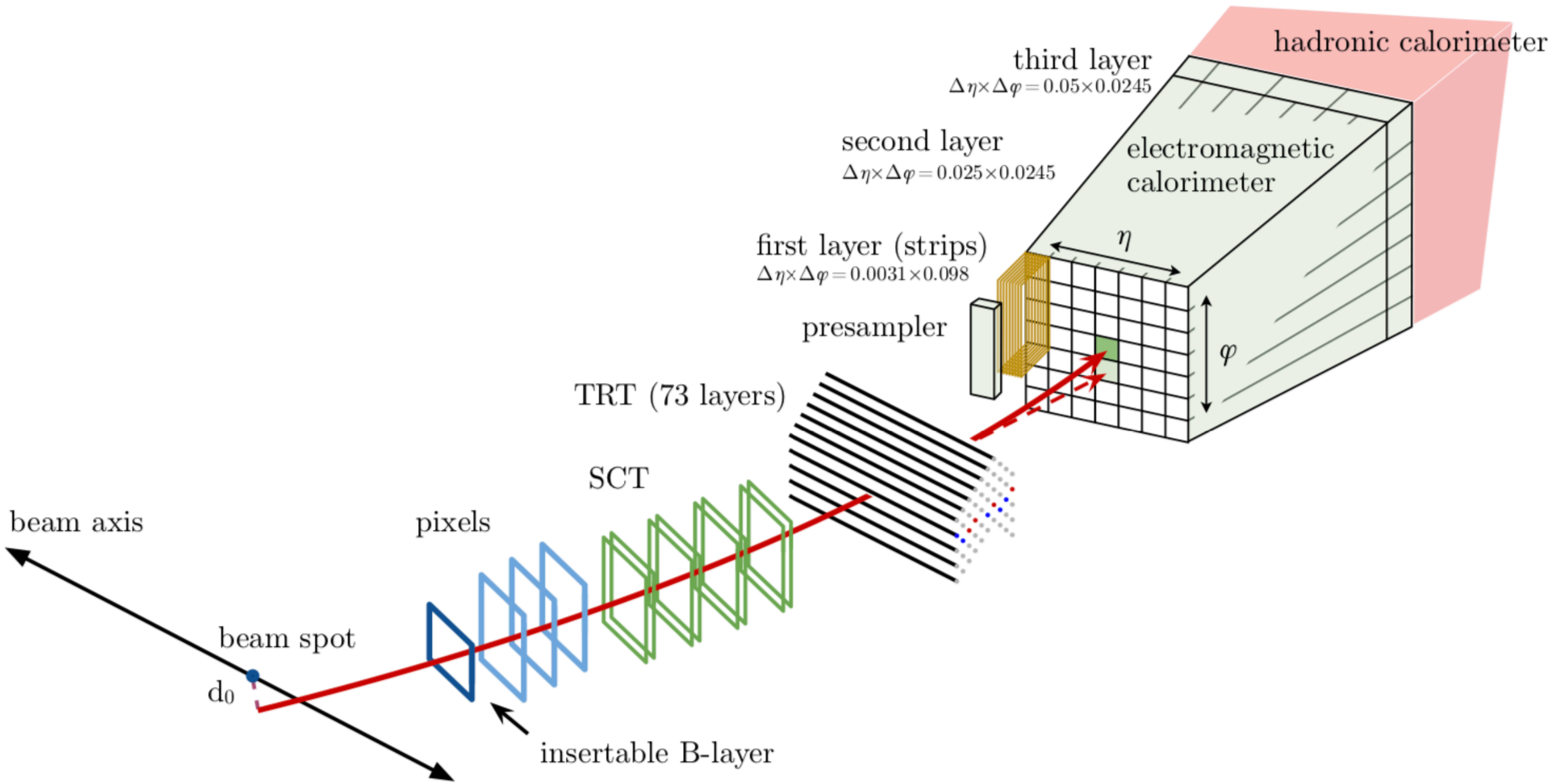
$t\bar{t}H$ ATLAS combination results

Phys. Lett. B 784 (2018) 173

	Run-1		Run-2			
Year	**2011**	**2012**	**2015**	**2016**	**2017**	**2018**
Collision energy	7 TeV	8 TeV	13 TeV			
$t\bar{t}H$-ML	$24.8\,\mathrm{fb}^{-1}$		$36\,\mathrm{fb}^{-1}$		✗	✗
$t\bar{t}H(b\bar{b})$	$24.8\,\mathrm{fb}^{-1}$		$36\,\mathrm{fb}^{-1}$		✗	✗
$t\bar{t}H(\gamma\gamma)$	$24.8\,\mathrm{fb}^{-1}$		$79.8\,\mathrm{fb}^{-1}$			✗
$t\bar{t}H(Z,Z^*)$	$24.8\,\mathrm{fb}^{-1}$		$79.8\,\mathrm{fb}^{-1}$			✗

- $\mu_{t\bar{t}H} = 1.32 \pm 0.18(\text{stat}) \pm^{+0.21}_{-0.19}$ (syst)
- Observed significance $6.3\ \sigma$ (expected $5.1\ \sigma$)

Electron reconstruction



Matrix method fakes

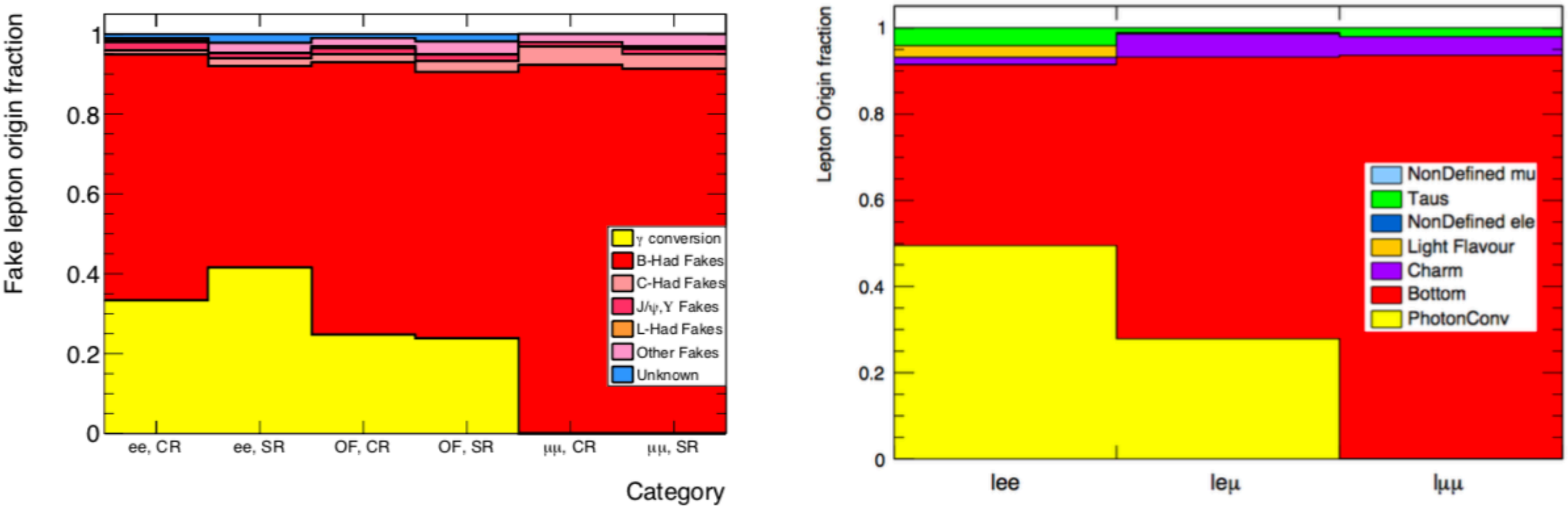
- TT_i : event with both leptons passing tight selection (Tot. events: N^{TT}).
- $T\not{T}_i$: event with leading lepton passing tight selection and subleading lepton failing tight selection (Tot. events: $N^{T\not{T}}$).
- $\not{T}T_i$: event with leading lepton failing tight selection and subleading lepton passing tight selection (Tot. events: $N^{\not{T}T}$).
- $\not{T}\not{T}_i$: event with both leptons failing tight selection (Tot. events: $N^{\not{T}\not{T}}$).

- rr_i : event with both leptons being real (Tot. events: N^{rr}).
- rf_i : event with leading lepton being real and subleading lepton being fake (Tot. events: N^{rf}).
- fr_i : event with leading lepton being fake and subleading lepton being real (Tot. events: N^{fr}).
- ff_i : event with both leptons being fake (Tot. events: N^{ff}).

$$\begin{pmatrix} N^{TT} \\ N^{T\not{T}} \\ N^{\not{T}T} \\ N^{\not{T}\not{T}} \end{pmatrix} = \begin{pmatrix} \varepsilon_{r,1}\varepsilon_{r,2} & \varepsilon_{r,1}\varepsilon_{f,2} & \varepsilon_{f,1}\varepsilon_{r,2} & \varepsilon_{f,1}\varepsilon_{f,2} \\ \varepsilon_{r,1}\not{\varepsilon}_{r,2} & \varepsilon_{r,1}\not{\varepsilon}_{f,2} & \varepsilon_{f,1}\not{\varepsilon}_{r,2} & \varepsilon_{f,1}\not{\varepsilon}_{f,2} \\ \not{\varepsilon}_{r,1}\varepsilon_{r,2} & \not{\varepsilon}_{r,1}\varepsilon_{f,2} & \not{\varepsilon}_{f,1}\varepsilon_{r,2} & \not{\varepsilon}_{f,1}\varepsilon_{f,2} \\ \not{\varepsilon}_{r,1}\not{\varepsilon}_{r,2} & \not{\varepsilon}_{r,1}\not{\varepsilon}_{f,2} & \not{\varepsilon}_{f,1}\not{\varepsilon}_{r,2} & \not{\varepsilon}_{f,1}\not{\varepsilon}_{f,2} \end{pmatrix} \begin{pmatrix} N^{rr} \\ N^{rf} \\ N^{fr} \\ N^{ff} \end{pmatrix}$$

Fake fraction

- Various sources of light-lepton fakes in 2l and 3l channel



(a) 2ℓ (b) 3ℓ