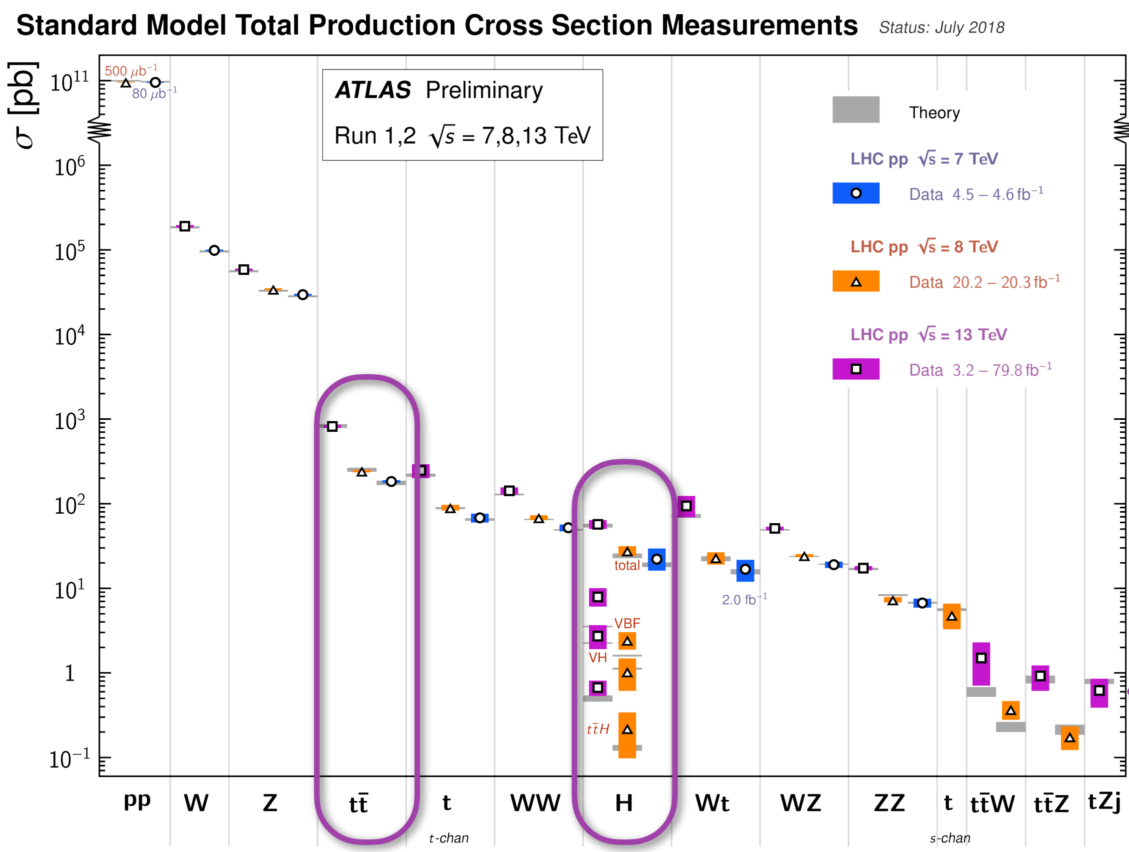
Search for 4-tops production at the LHC

Yang Qin (Quake)

- Latest ATLAS and CMS results
- Current ATLAS analysis
- Modelling the ttbar background

Introduction

- 4-tops production has extremely small cross-section in SM
  - 9.2 fb at NLO QCD precision at 13 TeV [arXiv:1405.0301 [hep-ph]]
  - 30% scale and 6% PDF uncertainties



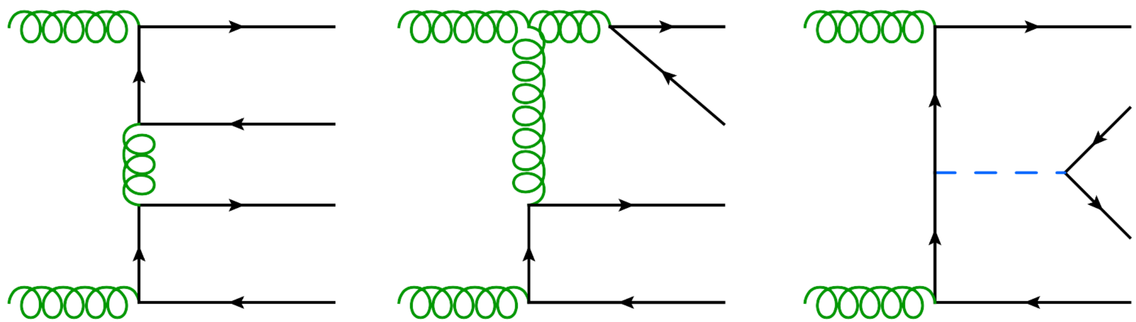
[CMS recent observation]

4tops

Introduction

- At LO, 72 (12) gg (qq') initiated diagrams

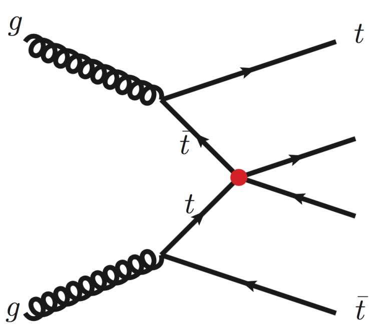
  [arXiv:1611.05032 [hep-ph]](arXiv:1611.05032 [hep-ph])

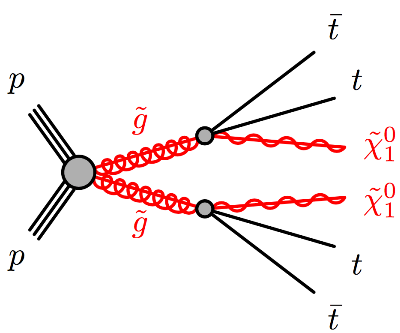


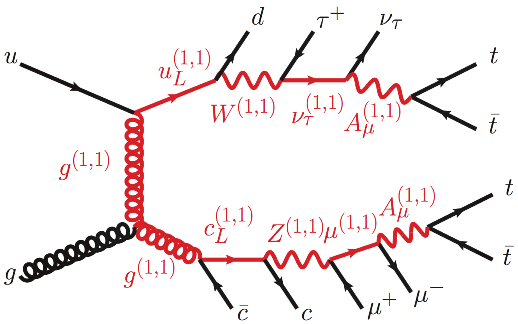
Leading: $\mathcal{O}(\alpha_S^4)$

Sub-leading: $\mathcal{O}(\alpha_S^2 y_t^4)$, $\mathcal{O}(\alpha_S^2 \alpha^2)$

- Sensitive to top-Yukawa coupling (y_t)
  - non-SM value of y_t can change dramatically the production via an off-shell Higgs
- Many BSM theories may enhance the production cross-section
  - 2HDM: heavy / pseudo scalar Higgs (H/A)
  - EFT: contact interaction
  - SUSY
  - 2 universal extra dimensions (2UDM)







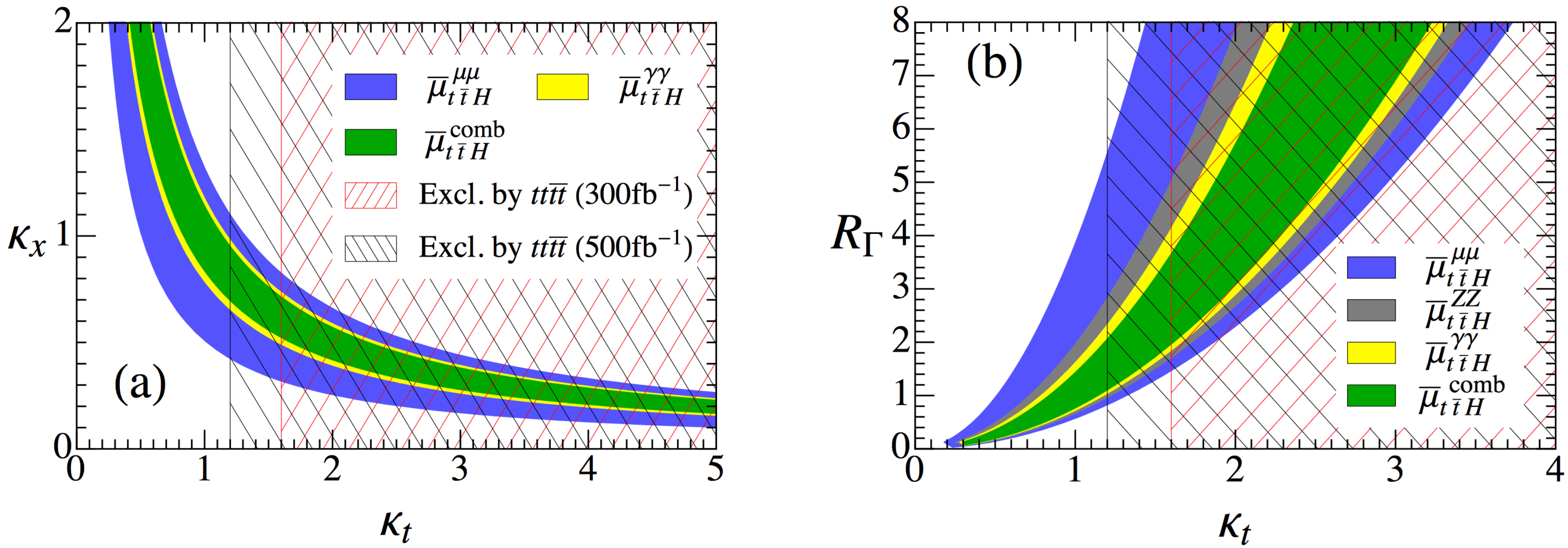
Introduction

- Example pheno study constraining top Yukawa coupling and Higgs width using ttH and 4-tops production (14 TeV) https://arxiv.org/abs/1602.01934

$$\sigma(pp \to t\bar{t}H \to t\bar{t}xx) = \sigma^{\rm SM}(pp \to t\bar{t}H \to t\bar{t}xx) \times \kappa_t^2 \kappa_x^2 \frac{\Gamma_H^{\rm SM}}{\Gamma_H}$$

$$\kappa_t = y_t / y_t^{\rm SM}$$
$$\kappa_x = y_x / y_x^{\rm SM}$$

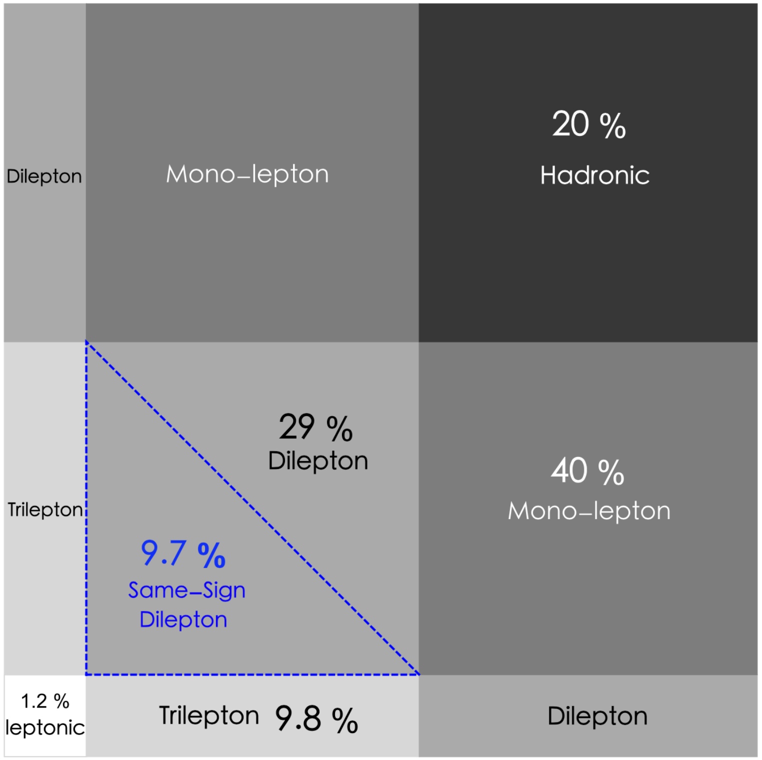
$$\mu_{t\bar{t}H}^{xx} \equiv \frac{\sigma}{\sigma^{\rm SM}} = \frac{\kappa_t^2 \kappa_x^2}{R_\Gamma} \qquad \text{with} \qquad R_\Gamma \equiv \frac{\Gamma_H}{\Gamma_H^{\rm SM}}$$



Latest ATLAS and CMS results

- CMS results with 2015+2016 data 35.9 fb^{-1} EPJC 78 (2018) 140
  - SS dilepton and ≥3 lepton channels
- ATLAS results with 2015+2016 data 36.1 fb^{-1}
  - SS dilepton and multilepton channels (**SS/ML**)
    arXiv:1807.11883 [hep-ex] Submitted to JHEP
  - single-lepton and OS dilepton channels (**1L/OS**), results combined with SS/ML channel
    arXiv:1811.02305 [hep-ex] Submitted to PRD
    - Compared to the SS/ML channel, larger branching fraction (>50%), but much larger background
    - Not the most sensitive channel, but could still improve sensitivity in the combination

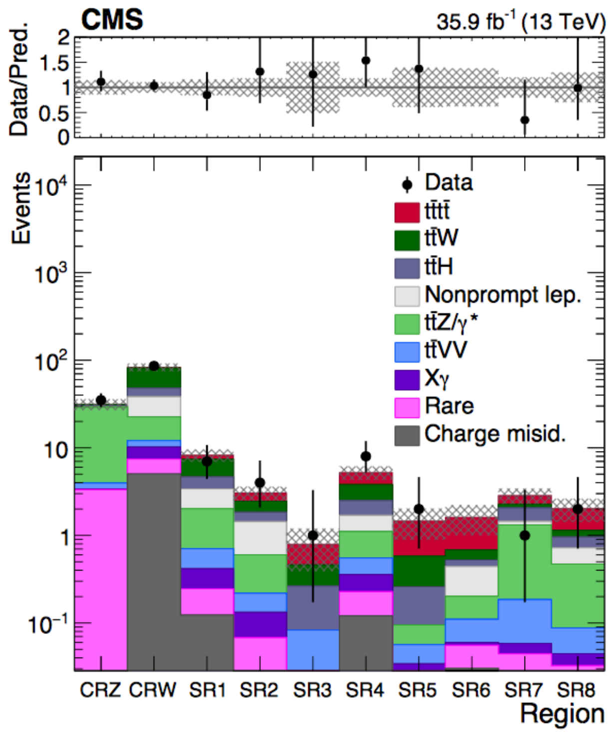
4tops branching fraction arXiv:1611.05032 [hep-ph]



* undecayed τ included

Latest CMS results

- CMS results with 2015+2016 data 35.9 fb^{-1} EPJC 78 (2018) 140
  - SS dilepton and ≥3 lepton channels
  - Dilepton (8 GeV) + H_T (300 GeV) trigger
    - 92% ee/eμ efficiency, 95% μμ efficiency
  - Selecting events with H_T > 300 GeV, p_T^{miss} > 50 GeV, low mass QCD veto, Z veto in ≥3 lepton channel (for ttZ)
    - 1.5% signal acceptance after baseline selection
  - Signal/control regions defined according to N_{lep}, N_{jets}, $N_{b\text{-}jets}$



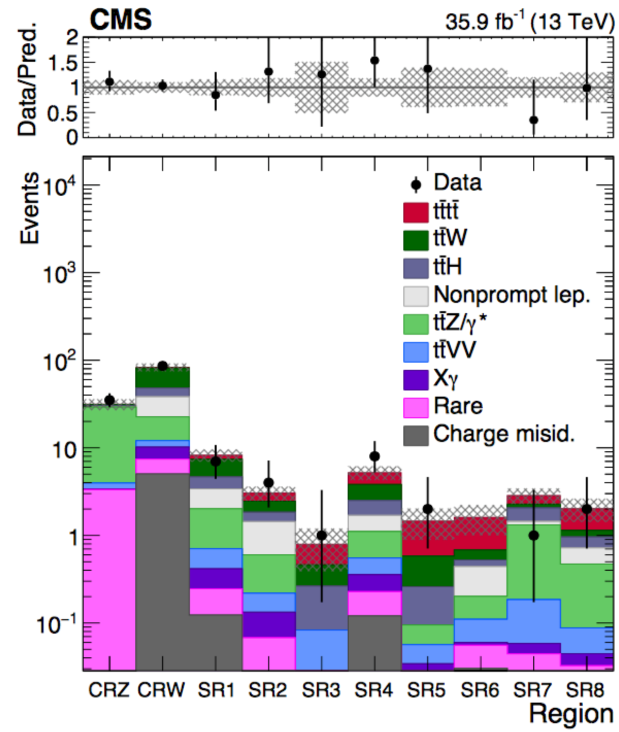
N_ℓ	N_b	N_{jets}	Region
2	2	≤5	CRW
		6	SR1
		7	SR2
		≥8	SR3
	3	5, 6	SR4
		≥7	SR5
	≥4	≥5	SR6
≥3	2	≥5	SR7
	≥3	≥4	SR8
Inverted Z veto			CRZ

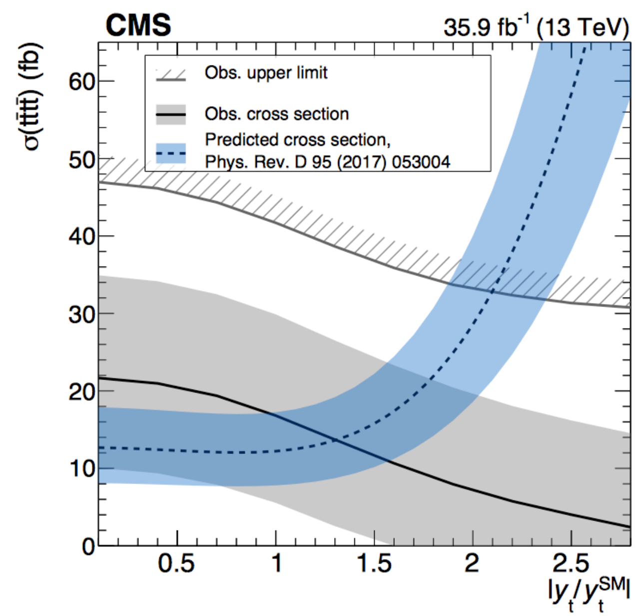
	SM background	tt̄tt̄	Total	Observed
CRZ	31.7 ± 4.6	0.4 ± 0.3	32.1 ± 4.6	35
CRW	83.7 ± 8.8	1.9 ± 1.2	85.6 ± 8.6	86
SR1	7.7 ± 1.2	0.9 ± 0.6	8.6 ± 1.2	7
SR2	2.6 ± 0.5	0.6 ± 0.4	3.2 ± 0.6	4
SR3	0.5 ± 0.3	0.4 ± 0.2	0.8 ± 0.4	1
SR4	4.0 ± 0.7	1.4 ± 0.9	5.4 ± 0.9	8
SR5	0.7 ± 0.2	0.9 ± 0.6	1.6 ± 0.6	2
SR6	0.7 ± 0.2	1.0 ± 0.6	1.7 ± 0.6	0
SR7	2.3 ± 0.5	0.6 ± 0.4	2.9 ± 0.6	1
SR8	1.2 ± 0.3	0.9 ± 0.6	2.1 ± 0.6	2

Latest CMS results

- CMS results with 2015+2016 data 35.9 fb^{-1} EPJC 78 (2018) 140
  - SS dilepton and ≥3 lepton channels
  - 95% CL upper limit on cross-section 41.7 fb (~4.5 x σ_{SM})
  - Observed (expected) significance 1.6 (1.0) σ
  - 95% CL limit on $| y_t / y_t^{SM} | < 2.1$

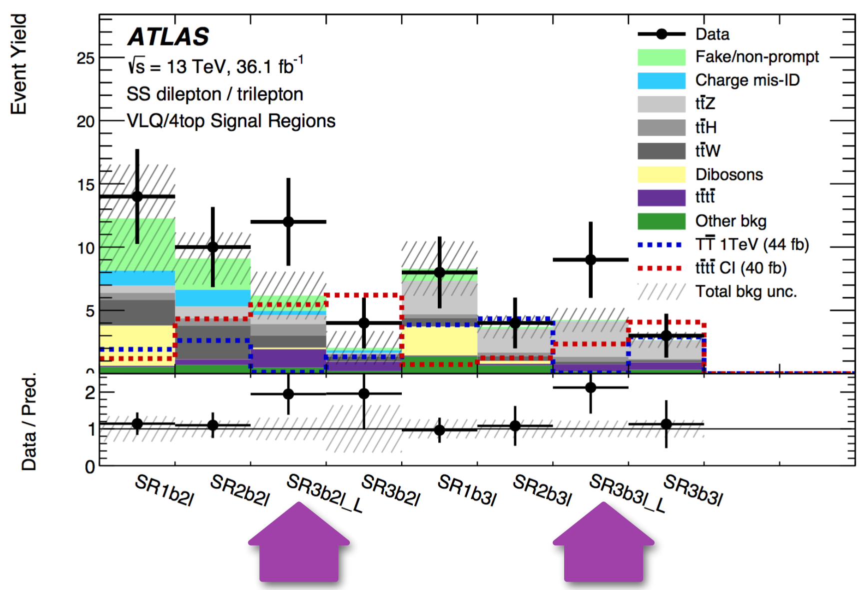
Source	Uncertainty (%)
Integrated luminosity	2.5
Pileup	0–6
Trigger efficiency	2
Lepton selection	4–10
Jet energy scale	1–15
Jet energy resolution	1–5
b tagging	1–15
Size of simulated sample	1–10
Scale and PDF variations	10–15
ISR/FSR (signal)	5–15
$t\bar{t}H$ (normalization)	50
Rare, $X\gamma$, $t\bar{t}VV$ (norm.)	50
$t\bar{t}Z/\gamma^*$, $t\bar{t}W$ (normalization)	40
Charge misidentification	20
Nonprompt leptons	30–60





Latest ATLAS results

- ATLAS results with 2015+2016 data 36.1 fb^{-1}
  - SS/ML
    - combination of high pT single-lepton and low pT dilepton triggers - 95% trigger efficiency
    - $m_{ee} > 15$ GeV and $| m_{ee} - 91 | > 10$ GeV for SS events
    - events categorised according to N_{lep} and $N_{b\text{-jets}}$ (N_{jets}), each category split into signal/validation regions using H_T and E_T^{miss}
    - main background
      - tt+V/H - based on MC simulation and profiled in the likelihood fit
      - non-prompt leptons - data driven technique



Region name	N_j	N_b	N_ℓ	Lepton charges	Kinematic criteria
VR1b2ℓ	≥ 1	1	2	++ or −−	$400 < H_T < 2400$ GeV or $E_T^{miss} < 40$ GeV
SR1b2ℓ	≥ 1	1	2	++ or −−	$H_T > 1000$ GeV and $E_T^{miss} > 180$ GeV
VR2b2ℓ	≥ 2	2	2	++ or −−	$H_T > 400$ GeV
SR2b2ℓ	≥ 2	2	2	++ or −−	$H_T > 1200$ GeV and $E_T^{miss} > 40$ GeV
VR3b2ℓ	≥ 3	≥ 3	2	++ or −−	$400 < H_T < 1400$ GeV or $E_T^{miss} < 40$ GeV
SR3b2ℓ_L	≥ 7	≥ 3	2	++ or −−	$500 < H_T < 1200$ GeV and $E_T^{miss} > 40$ GeV
SR3b2ℓ	≥ 3	≥ 3	2	++ or −−	$H_T > 1200$ GeV and $E_T^{miss} > 100$ GeV
VR1b3ℓ	≥ 1	1	3	any	$400 < H_T < 2000$ GeV or $E_T^{miss} < 40$ GeV
SR1b3ℓ	≥ 1	1	3	any	$H_T > 1000$ GeV and $E_T^{miss} > 140$ GeV
VR2b3ℓ	≥ 2	2	3	any	$400 < H_T < 2400$ GeV or $E_T^{miss} < 40$ GeV
SR2b3ℓ	≥ 2	2	3	any	$H_T > 1200$ GeV and $E_T^{miss} > 100$ GeV
VR3b3ℓ	≥ 3	≥ 3	3	any	$H_T > 400$ GeV
SR3b3ℓ_L	≥ 5	≥ 3	3	any	$500 < H_T < 1000$ GeV and $E_T^{miss} > 40$ GeV
SR3b3ℓ	≥ 3	≥ 3	3	any	$H_T > 1000$ GeV and $E_T^{miss} > 40$ GeV

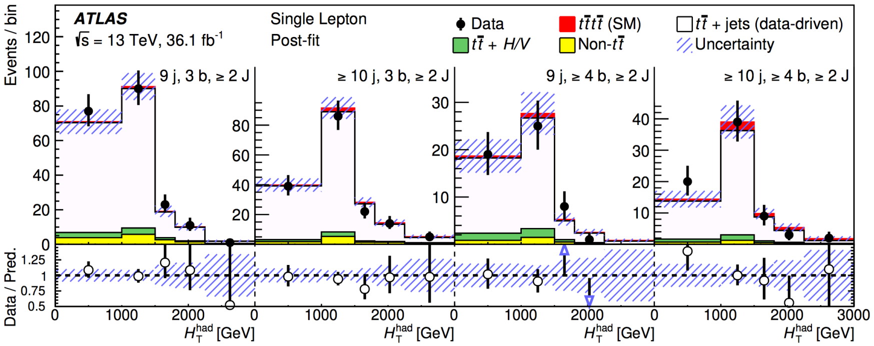
Latest ATLAS results

- ATLAS results with 2015+2016 data 36.1 fb^{-1}
  - 1L/OS

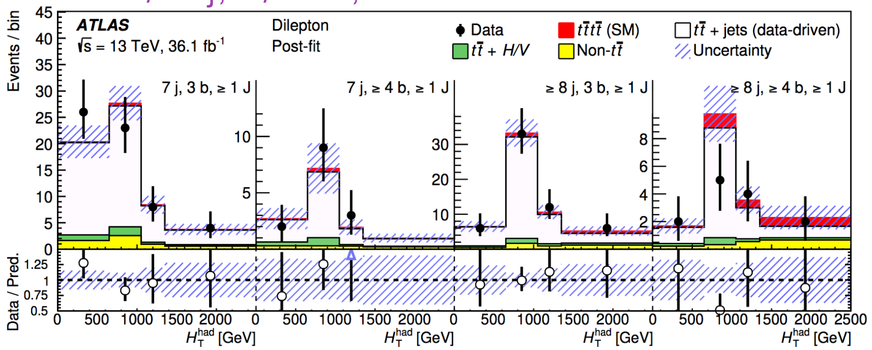
Requirement	Single-lepton	Dilepton
	$E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV	$m_{\ell\ell} > 50$ GeV
	$E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}}^{W} > 60$ GeV	$\lvert m_{\ell\ell} - 91\ \mathrm{GeV}\rvert > 8$ GeV

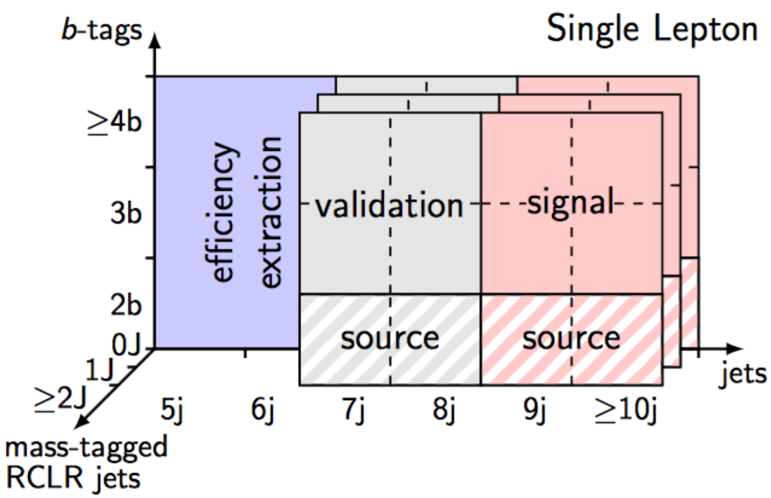
  - events categorised according to N$_{\text{jets}}$, N$_{\text{b-jets}}$ and N$_{\text{RC-jets}}$
  - main background from ttbar+jets
    - **ttbar+HF jets** in signal regions - data-driven method

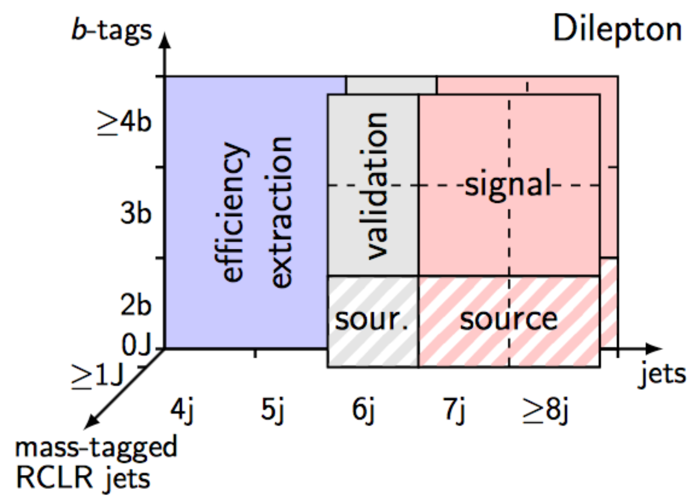
1L 9/≥10j, 3/≥4b, ≥2J



OS 7/≥8j, 3/≥4b, ≥1J







Latest ATLAS results

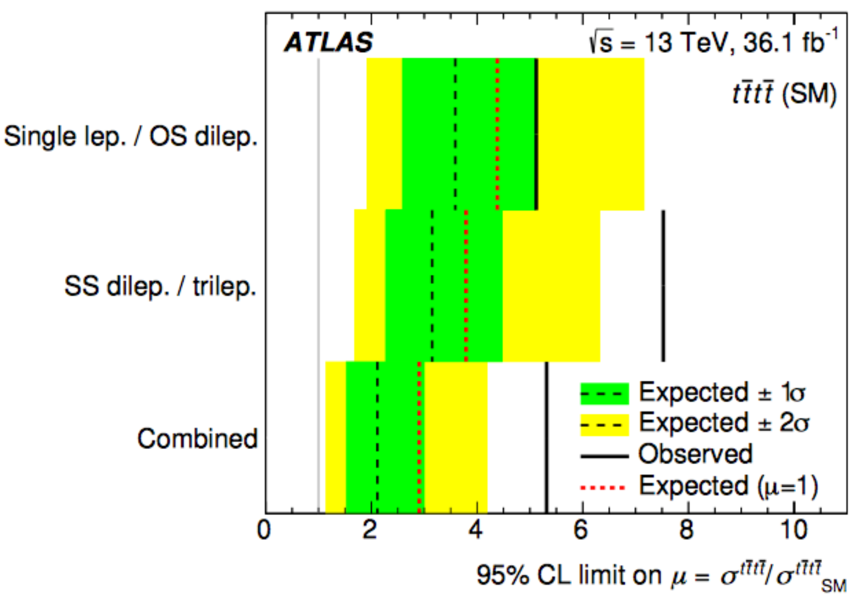
- ATLAS results with 2015+2016 data 36.1 fb^{-1}
  - Uncertainties
    - 1L/OS:
      - systematic dominated
      - primarily tt+jets modelling
    - SS/ML:
      - limited by stat.
      - dominant systematic - fake background and other background modelling

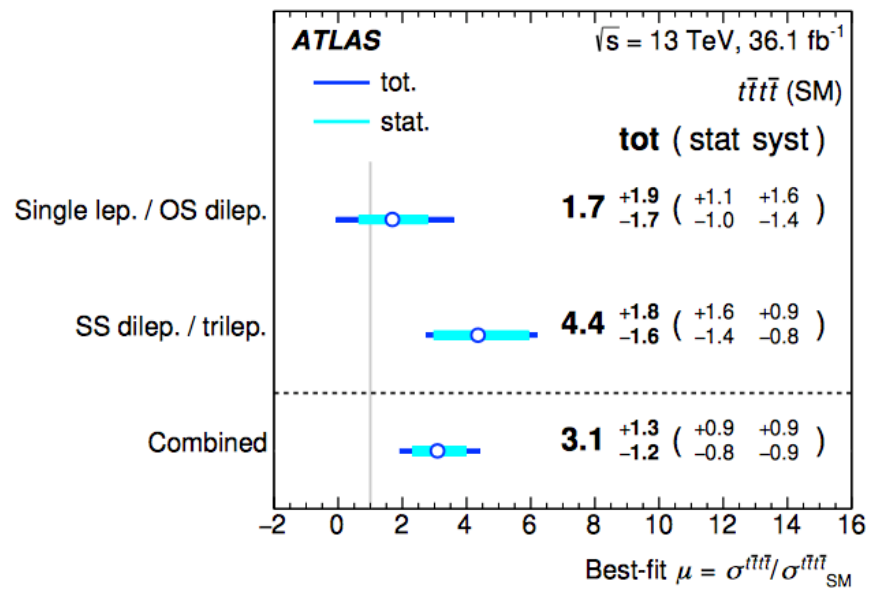
Uncertainty source	$\pm\Delta\mu$	
$t\bar{t}$+jets modeling	+1.2	−0.96
Background-model statistical uncertainty	+0.91	−0.85
Jet energy scale and resolution, jet mass	+0.38	−0.16
Other background modeling	+0.26	−0.20
b-tagging efficiency and mis-tag rates	+0.33	−0.10
JVT, pileup modeling	+0.18	−0.073
$t\bar{t}+H/V$ modeling	+0.053	−0.055
Luminosity	+0.050	−0.026
Total systematic uncertainty	+1.6	−1.4
Total statistical uncertainty	+1.1	−1.0
Total uncertainty	+1.9	−1.7

Uncertainty source	SR1b2ℓ [%]	SR2b2ℓ [%]	SR3b2ℓ_L [%]	SR3b2ℓ [%]	SR1b3ℓ [%]	SR2b3ℓ [%]	SR3b3ℓ_L [%]	SR3b3ℓ [%]
Jet energy resolution	3	1	5	6	3	5	3	4
Jet energy scale	3	3	9	6	3	5	11	6
b-tagging efficiency	5	3	6	7	3	4	9	9
Lepton ID efficiency	2	1	1	1	3	3	2	3
Pile-up reweighting	5	2	3	3	3	5	1	6
Luminosity	1	1	2	2	2	2	2	2
Fake/non-prompt	20	12	13	8	7	2	3	1
Charge mis-ID	2	3	1	2	-	-	-	-
Cross-section × acceptance	25	13	22	32	32	26	21	24

Latest ATLAS results

- ATLAS results with 2015+2016 data 36.1 fb^{-1}
  - Combined results
    - Observed (expected) 95% CL upper limit on cross-section 5.3 (2.1) x σ_{SM}
    - Observed (expected) significance 2.8 (1.0) σ
      - SS/ML: 3.0 (0.8) σ
      - 1L/OS: 1.0 (0.6) σ



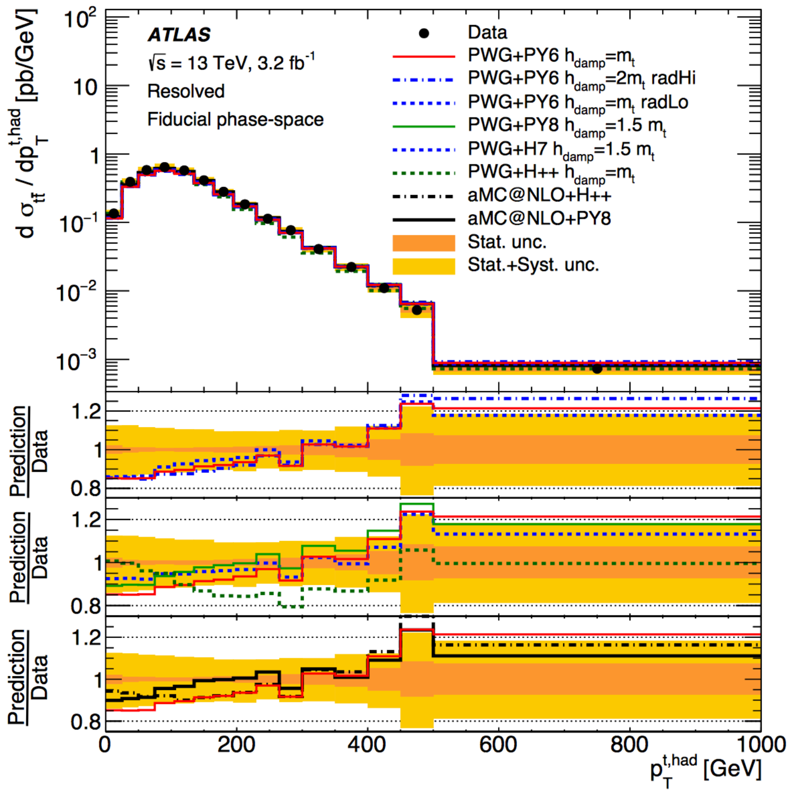


Current ATLAS analysis overview

- Based on the previous analyses
    - working towards full Run2 dataset
    - a combined effort from SS/ML and 1L/OS
- At the moment focusing on improving the background modelling
    - SS/ML channel mainly focusing on non-prompt lepton contribution
        - Matrix Method (data-driven)
        - template fit method (MC-based)
    - 1L/OS channel mainly focusing on modelling ttbar+jets - ideally also reduce the relevant modelling uncertainties
- Techniques to extract the signal to be reviewed - previous studies used simply cut-and-count strategy
    - use boosted decision tree to reconstruct the 4-tops system - mainly the jet combinatorics
    - investigate kinematic variables sensitive to the 4-tops signal topology

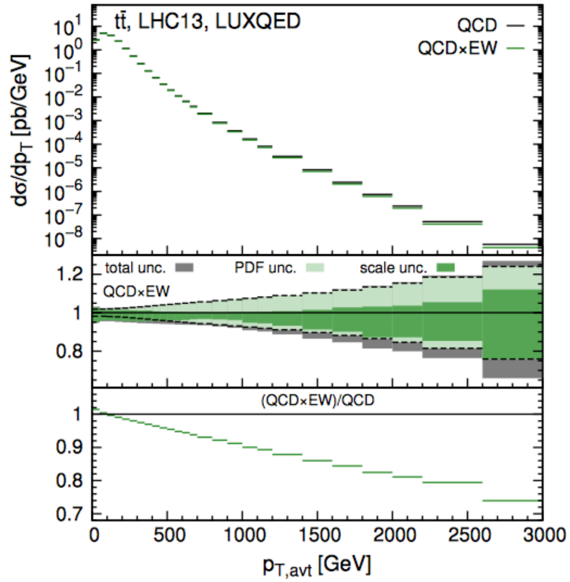
Modelling of ttbar background

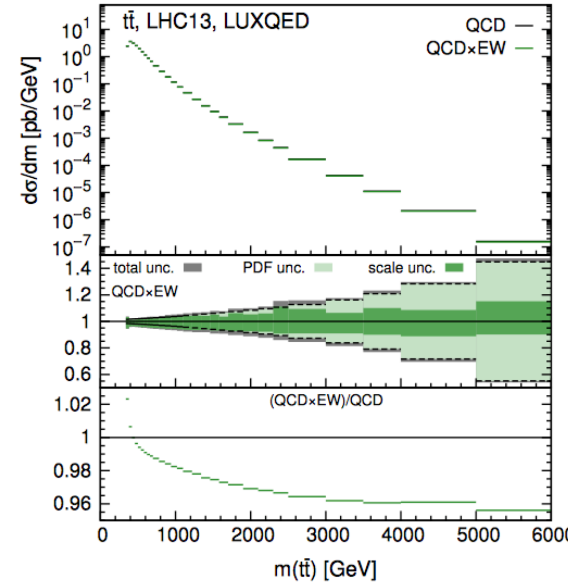
- Most important background - ttbar+jets
- Infamous mismodelling of ttbar with current simulations
  - smaller discrepancy for lower energy regime
    - often can be tuned to data
  - we work in an extreme phase space
    - high energy scale
      - higher order corrections
    - high jet and b-jet multiplicities
      - Sherpa+OpenLoops NLO ttbb prediction with multi-leg
- Improvement from the theory side
  - NNLO QCD + NLO EW corrections [arXiv:1705.04105 [hep-ph]]
  - NNLO + NNLL' [arXiv: 1803.07623 [hep-ph]]
- Experimentalists' approaches
  - Data-driven method based tag-rate-function (ttTRF), assisted with corrections from MC
  - MC-based method, assisted with data-driven corrections



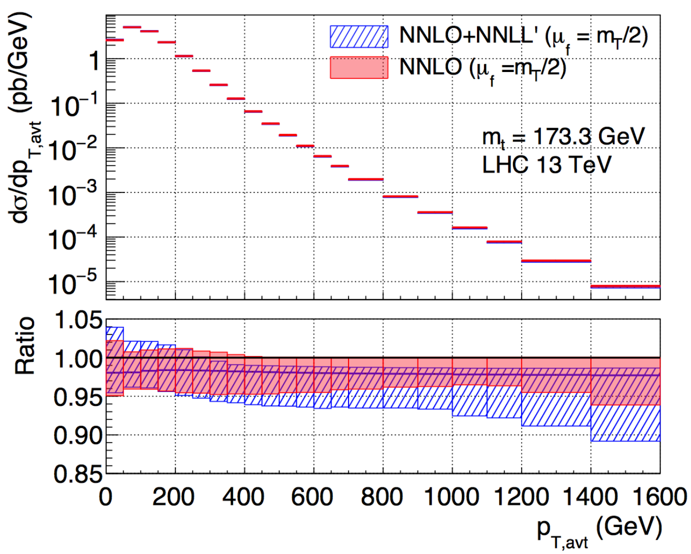
Higher order corrections

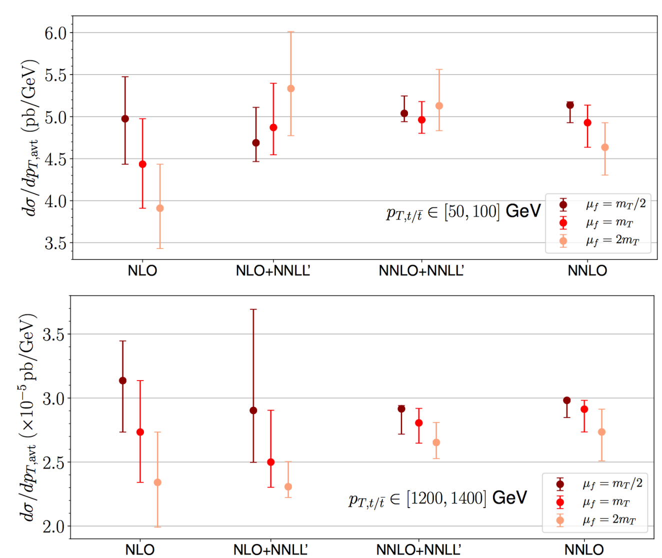
- Reweighting to NNLO QCD + NLO EW prediction on ttbar production





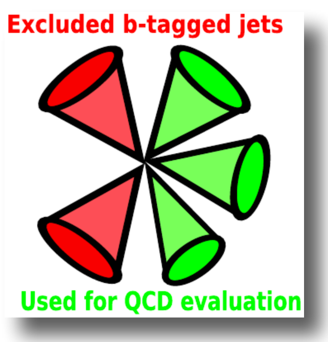
- NNLO+NNLL'
  - important for boosted tops
  - reduces dependence on scale choice

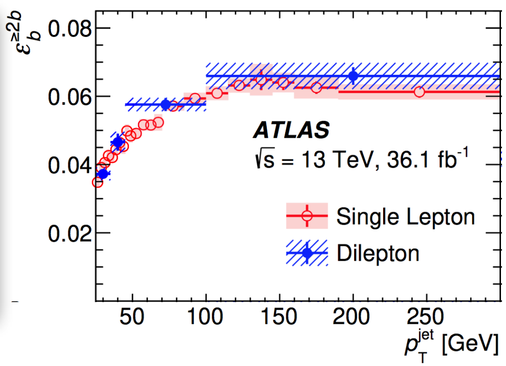


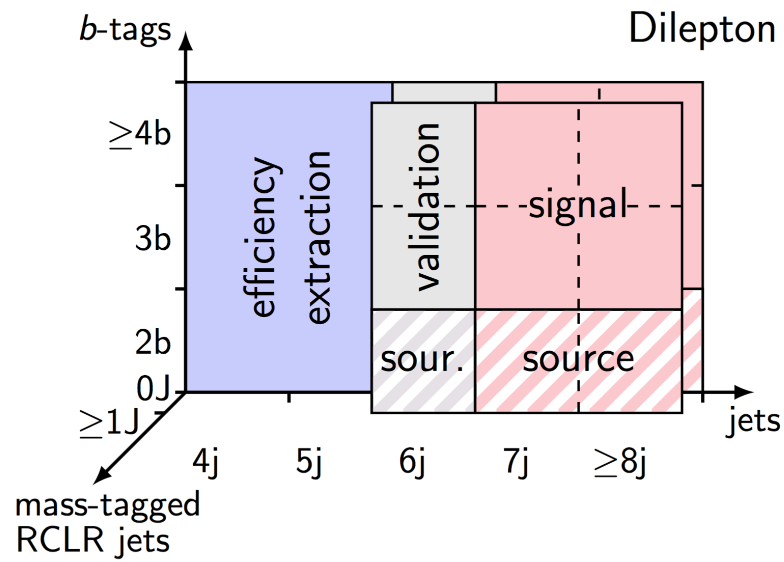


Data-driven ttbar estimate

- Data-driven method based on the tag-rate-function (ttTRF)
  - Extract b-tagging efficiencies from **data** in lower N_{jets} regions
    - inclusive efficiency regardless of the jet truth flavour







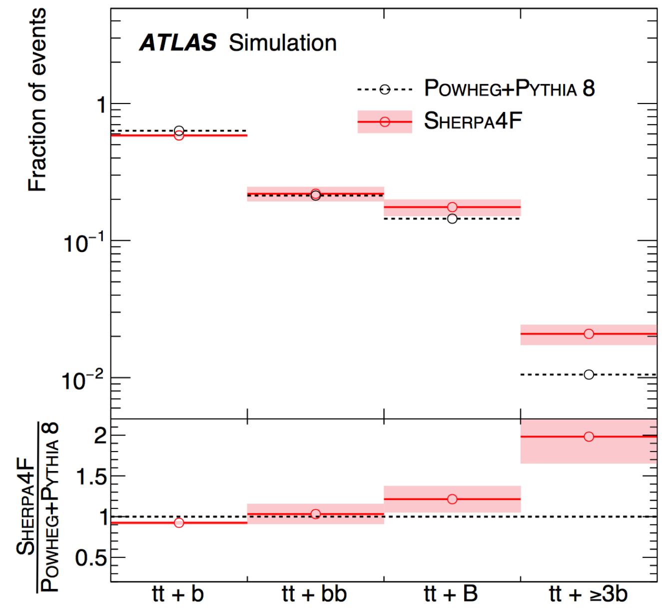
  - Using the tag-rate-function (TRF) to promote data events in 2b regions (source regions) to higher b-tag regions
  - Apply correction factors based on MC
    - Apply the same procedure to ttbar MC as in data
    - Take the non-closure between TRF and pure MC prediction as correction factor
    - Two effects taken into account
      - Residual kinematic dependence
      - Dependence of the efficiencies due to the extrapolation to higher N_{jets}

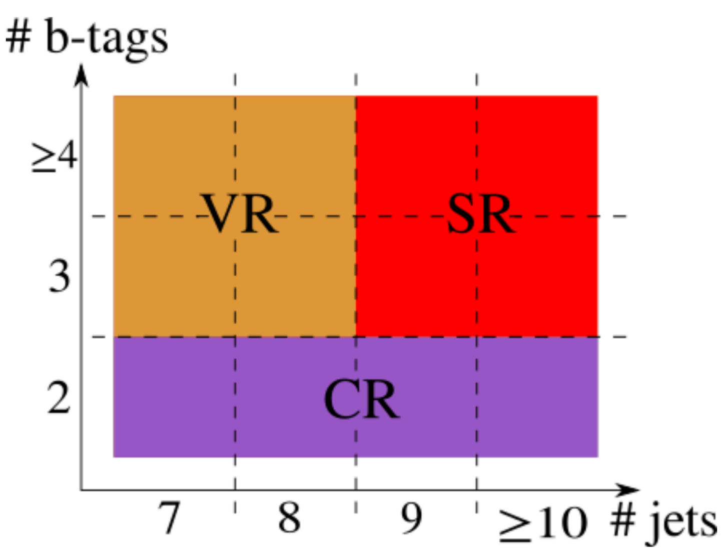
MC-based method

- Originated from [ATLAS ttH(bb) analysis]
- Baseline: 4FS Sherpa+OpenLoops (NLO ttbb production at ME level)
  - sophisticated ttbar categorisation based on the type and number of extra jets
  - reweighting to Sherpa
- Data-driven corrections to diminish the kinematic mismodelling
  - Derive parameterised reweighting factors in 2 b-tags control regions (CR)

$$w \propto \frac{\text{Data} - \text{MC}^{\text{non}-t\bar{t}}}{\text{MC}^{t\bar{t}}}$$

  - Apply higher b-tags regions: validation regions (VR) and eventually signal regions (SR)
- Systematic uncertainties from Sherpa and the reweighting
  - profiling in the likelihood fit





Summary

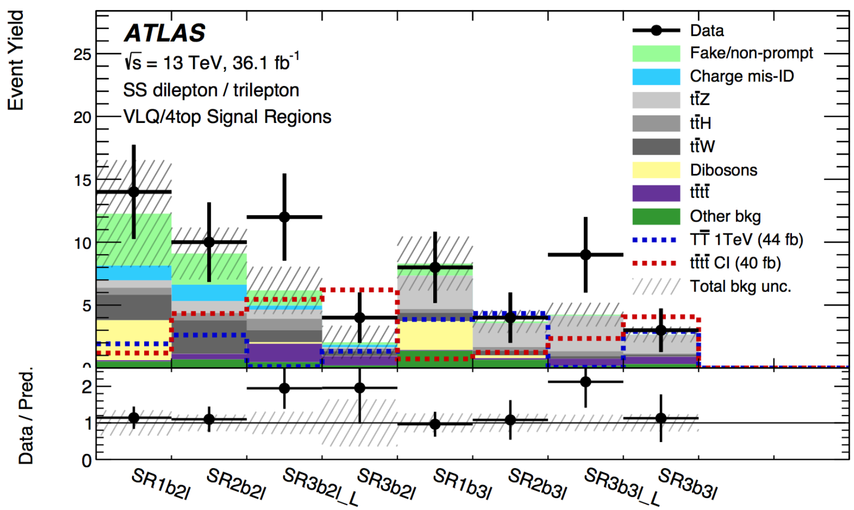
- 4-tops production has small cross-section in SM, but an interesting and important process for testing the SM and probing BSM signals
- ATLAS and CMS have both shown increasing interest in this process as the amount of LHC data ramps up
- ATLAS is working towards a combination of all sensitive channels, 1L/OS + SS/ML
    - Current focusing on improving background modelling
    - In the 1L/OS channel, ttbar background modelling is critical
    - Potential improvement are being studied: data-driven and MC-based techniques
- Expect a much increased sensitivity with
    - full Run2 data
    - improved background modelling
    - MVA signal extraction

MAY THE TOPS BE WITH YOU

Thanks for your attention

BACKUP

Latest ATLAS results

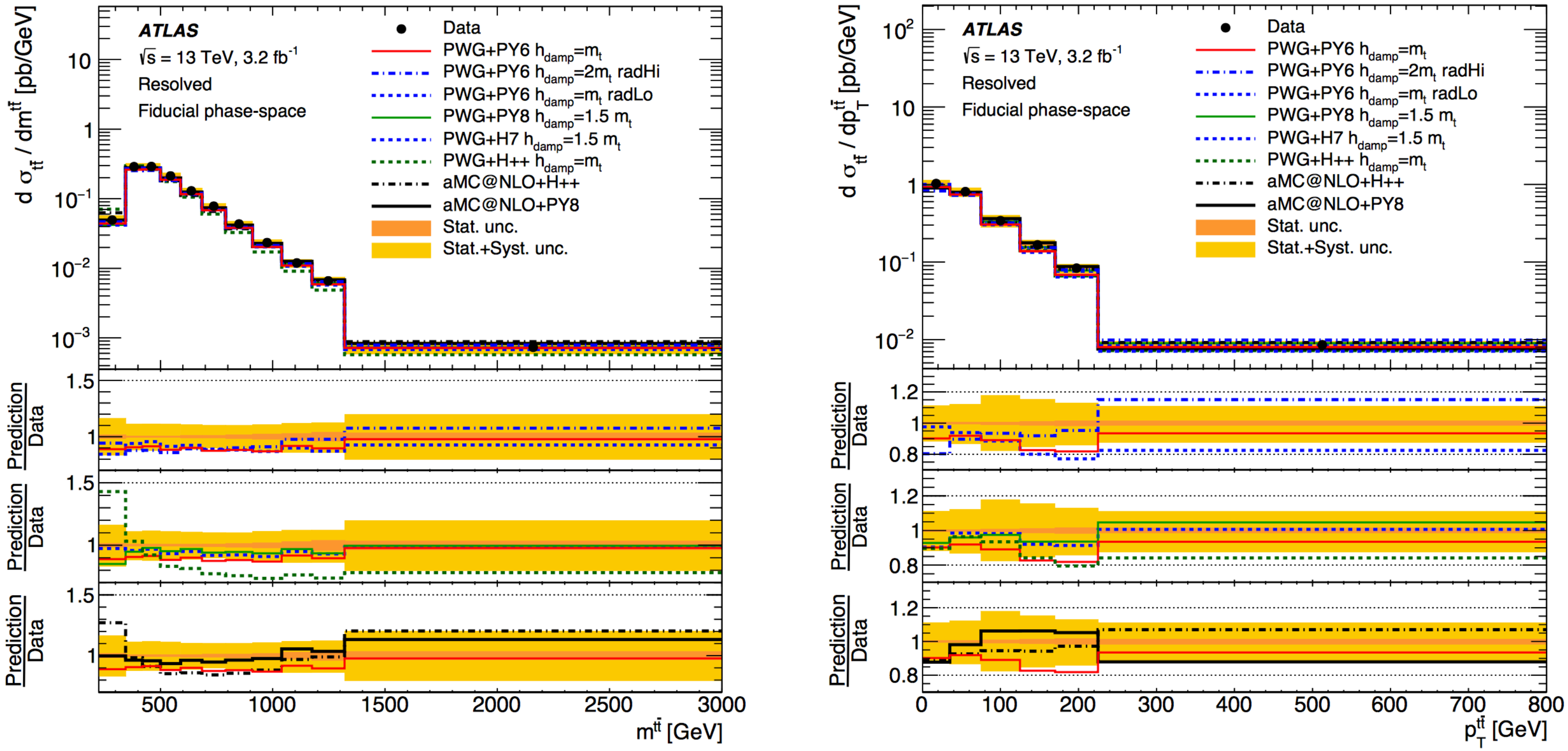


Source	SR1b2ℓ	SR2b2ℓ	SR3b2ℓ_L	SR3b2ℓ
$t\bar{t}W$	$2.04 \pm 0.14 \pm 0.49$	$2.68 \pm 0.15 \pm 0.55$	$0.95 \pm 0.11 \pm 0.31$	$0.40 \pm 0.06 \pm 0.10$
$t\bar{t}Z$	$0.58 \pm 0.08 \pm 0.10$	$0.95 \pm 0.11 \pm 0.17$	$0.72 \pm 0.11 \pm 0.19$	$0.11 \pm 0.05\ ^{+0.13}_{-0.10}$
Dibosons	$3.2 \pm 1.5 \pm 2.4$	< 0.5	$0.13 \pm 0.13\ ^{+0.27}_{-0.00}$	< 0.5
$t\bar{t}H$	$0.56 \pm 0.07 \pm 0.07$	$0.57 \pm 0.10 \pm 0.09$	$0.91 \pm 0.11 \pm 0.22$	$0.19 \pm 0.05 \pm 0.07$
$t\bar{t}t\bar{t}$	$0.10 \pm 0.01 \pm 0.05$	$0.44 \pm 0.03 \pm 0.23$	$1.46 \pm 0.05 \pm 0.74$	$0.75 \pm 0.04 \pm 0.38$
Other bkg	$0.52 \pm 0.07 \pm 0.14$	$0.68 \pm 0.09 \pm 0.24$	$0.47 \pm 0.08 \pm 0.18$	$0.20 \pm 0.04 \pm 0.06$
Fake/non-prompt	$4.1\ ^{+1.6}_{-1.4} \pm 2.4$	$2.5\ ^{+1.0}_{-0.9} \pm 1.1$	$1.2\ ^{+0.9}_{-0.7} \pm 0.6$	$0.20\ ^{+0.46}_{-0.20} \pm 0.16$
Charge mis-ID	$1.17 \pm 0.10 \pm 0.27$	$1.29 \pm 0.10 \pm 0.28$	$0.32 \pm 0.04 \pm 0.09$	$0.21 \pm 0.04 \pm 0.04$
Total bkg	$12.3\ ^{+2.2}_{-2.1} \pm 3.4$	$9.1\ ^{+1.2}_{-1.1} \pm 1.2$	$6.2\ ^{+1.0}_{-0.8} \pm 1.2$	$2.0\ ^{+0.5}_{-0.2} \pm 0.3$
Data yield	14	10	12	4
BSM significance	0.31	0.25	1.7	1.1
SM $t\bar{t}t\bar{t}$ significance	0.33	0.38	2.1	1.6

Source	SR1b3ℓ	SR2b3ℓ	SR3b3ℓ_L	SR3b3ℓ
$t\bar{t}W$	$0.66 \pm 0.08 \pm 0.20$	$0.38 \pm 0.05 \pm 0.11$	$0.21 \pm 0.05 \pm 0.09$	$0.15 \pm 0.04 \pm 0.05$
$t\bar{t}Z$	$2.66 \pm 0.15 \pm 0.43$	$1.90 \pm 0.14 \pm 0.42$	$2.80 \pm 0.17 \pm 0.58$	$1.47 \pm 0.14 \pm 0.28$
Dibosons	$2.3 \pm 0.7 \pm 1.7$	$0.22 \pm 0.16 \pm 0.27$	< 0.5	< 0.5
$t\bar{t}H$	$0.30 \pm 0.04 \pm 0.04$	$0.28 \pm 0.05 \pm 0.05$	$0.38 \pm 0.06 \pm 0.07$	$0.10 \pm 0.03 \pm 0.02$
$t\bar{t}t\bar{t}$	$0.06 \pm 0.01 \pm 0.03$	$0.13 \pm 0.02 \pm 0.06$	$0.58 \pm 0.04 \pm 0.29$	$0.59 \pm 0.03 \pm 0.30$
Other bkg.	$1.37 \pm 0.13 \pm 0.45$	$0.65 \pm 0.10 \pm 0.27$	$0.17 \pm 0.09 \pm 0.10$	$0.31 \pm 0.07 \pm 0.11$
Fake/non-prompt	$1.0\ ^{+0.6}_{-0.5} \pm 0.6$	$0.14\ ^{+0.31}_{-0.12} \pm 0.09$	$0.00\ ^{+0.38}_{-0.00}\ ^{+0.09}_{-0.00}$	$0.03\ ^{+0.15}_{-0.02} \pm 0.00$
Total bkg	$8.3\ ^{+0.9}_{-0.8} \pm 1.8$	$3.7\ ^{+0.6}_{-0.3} \pm 0.4$	$4.2\ ^{+0.4}_{-0.2} \pm 0.7$	$2.7 \pm 0.2 \pm 0.5$
Data yield	8	4	9	3
BSM significance	−0.09	0.14	1.8	0.19
SM $t\bar{t}t\bar{t}$ significance	−0.07	0.21	2.1	0.6

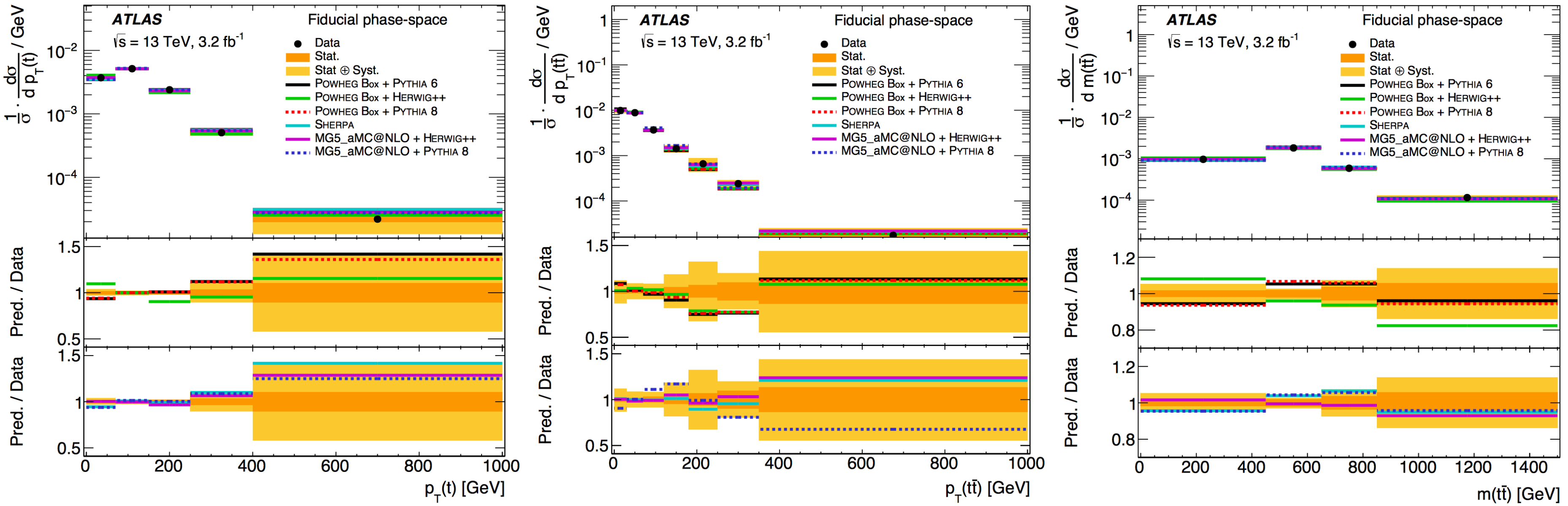
Modelling of ttbar background

- Infamous mismodelling of ttbar with current simulations on the market
-

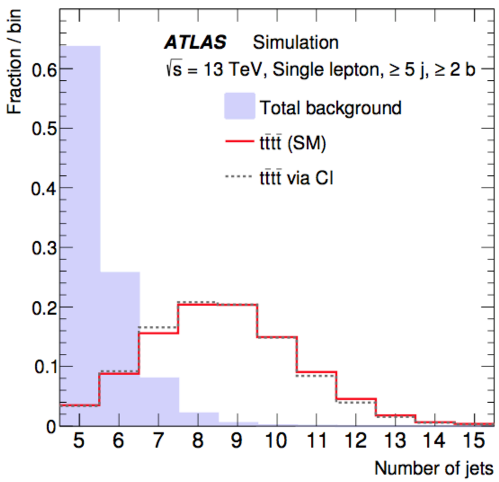


Modelling of ttbar background

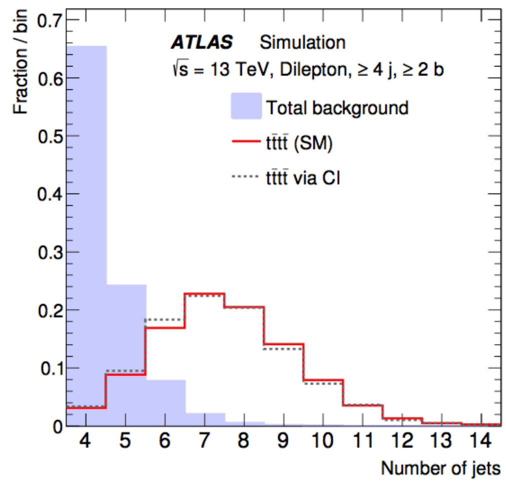
- Infamous mismodelling of ttbar with current simulations on the market
-



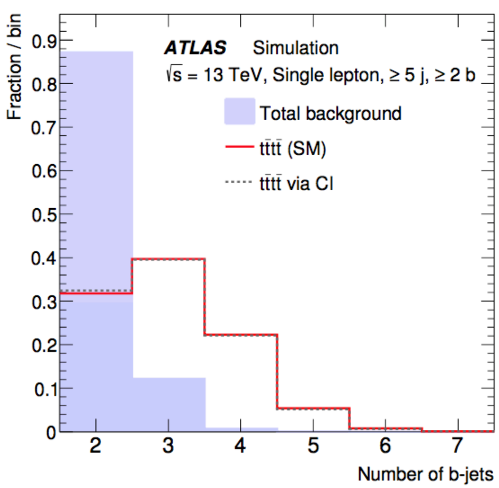
1L/OS


ATLAS Simulation
√s = 13 TeV, Single lepton, ≥ 5 j, ≥ 2 b
Total background
tt̄tt̄ (SM)
tt̄tt̄ via CI
Fraction / bin
Number of jets

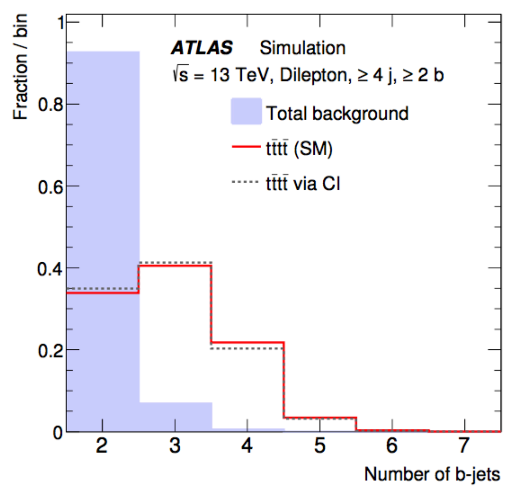
(a)


ATLAS Simulation
√s = 13 TeV, Dilepton, ≥ 4 j, ≥ 2 b
Total background
tt̄tt̄ (SM)
tt̄tt̄ via CI
Fraction / bin
Number of jets

(b)


ATLAS Simulation
√s = 13 TeV, Single lepton, ≥ 5 j, ≥ 2 b
Total background
tt̄tt̄ (SM)
tt̄tt̄ via CI
Fraction / bin
Number of b-jets

(c)


ATLAS Simulation
√s = 13 TeV, Dilepton, ≥ 4 j, ≥ 2 b
Total background
tt̄tt̄ (SM)
tt̄tt̄ via CI
Fraction / bin
Number of b-jets

(d)

Data-driven ttbar estimate - efficiencies