$t\bar{t}b\bar{b}$+$t\bar{t}t\bar{t}$ measurements at the LHC

ATLAS and CMS results

Tae Jeong Kim (Hanyang University) on behalf of ATLAS and CMS collaborations

May 8 in 2024 for SM@LHC at Rome, Italy

$t\bar{t}b\bar{b}$ cross section measurement

- $t\bar{t}b\bar{b}$ is the main background for the $t\bar{t}H(b\bar{b})$ signature
- a good test of the NLO QCD theory
- Measurements of $t\bar{t}$ + heavy flavor are challenging
    - Huge combinatorics from multiple b-jets
    - Identification of b-jets origin: top quark, gluon splitting, Higgs, other?
    - large theory uncertainty due to the presence of two very different scales (top quark mass, b quark mass)

Differential cross section measurement

$t\bar{t}b\bar{b}$ process

- Which b jets? - Need to identify two b jets at the generator level
- How do we find those b jets at the reconstruction level?
- Additional b jets can mean
  - Two b jets with the highest p_T - targeting b jets from Higgs
  - Two b jets with the closest angle between them - targeting b jets from gluon splitting
  - Two b jets not from a top quark →

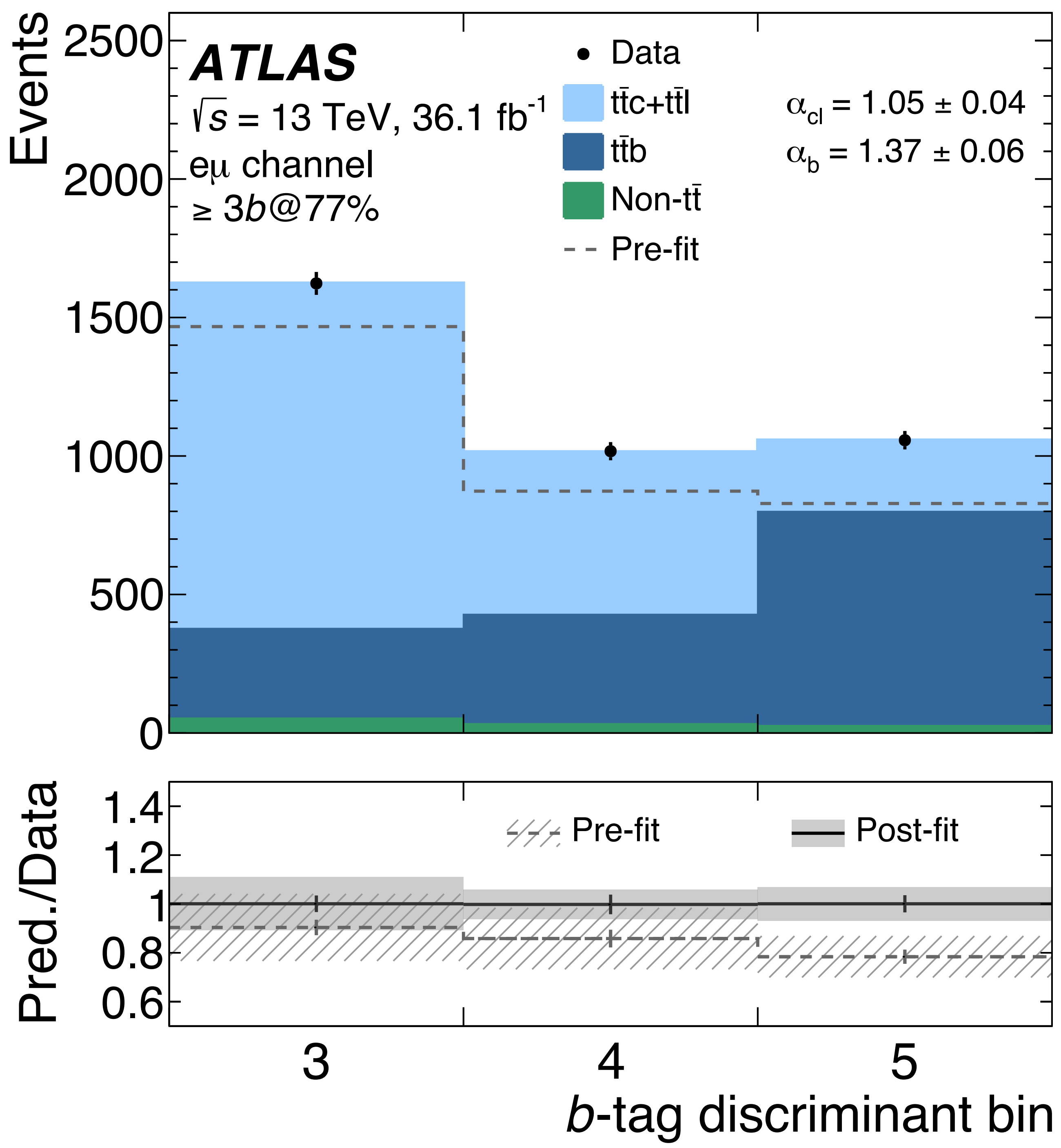
- Remove systematic uncertainty on the theory dependence
- Consistent with the reconstruction level definition

- more accurate and sensitive to modeling of gluon splitting
- DNN at the reconstruction level

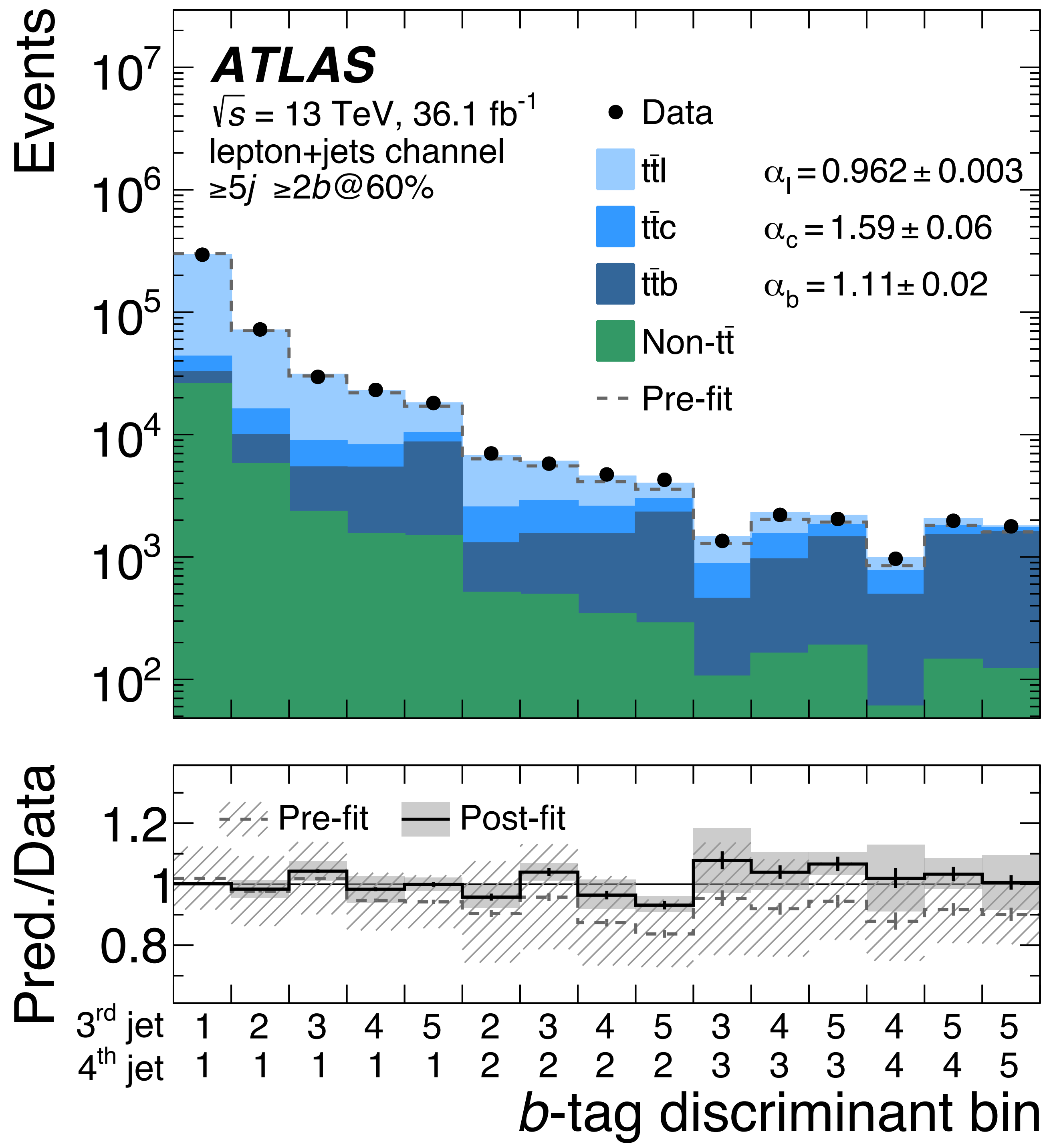
[JHEP 04 (2019) 046](https://doi.org/)

Signal extraction

- Binned maximum-likelihood fitting to extract $t\bar{t}b, t\bar{t}c, t\bar{t}l$
- Unfolded back to particle level by correcting detector resolution, efficiency and acceptance
- Iterative Bayesian unfolding technique in RooUnfold Package



Template of third highest b-tagged jets in $e\mu$ channel

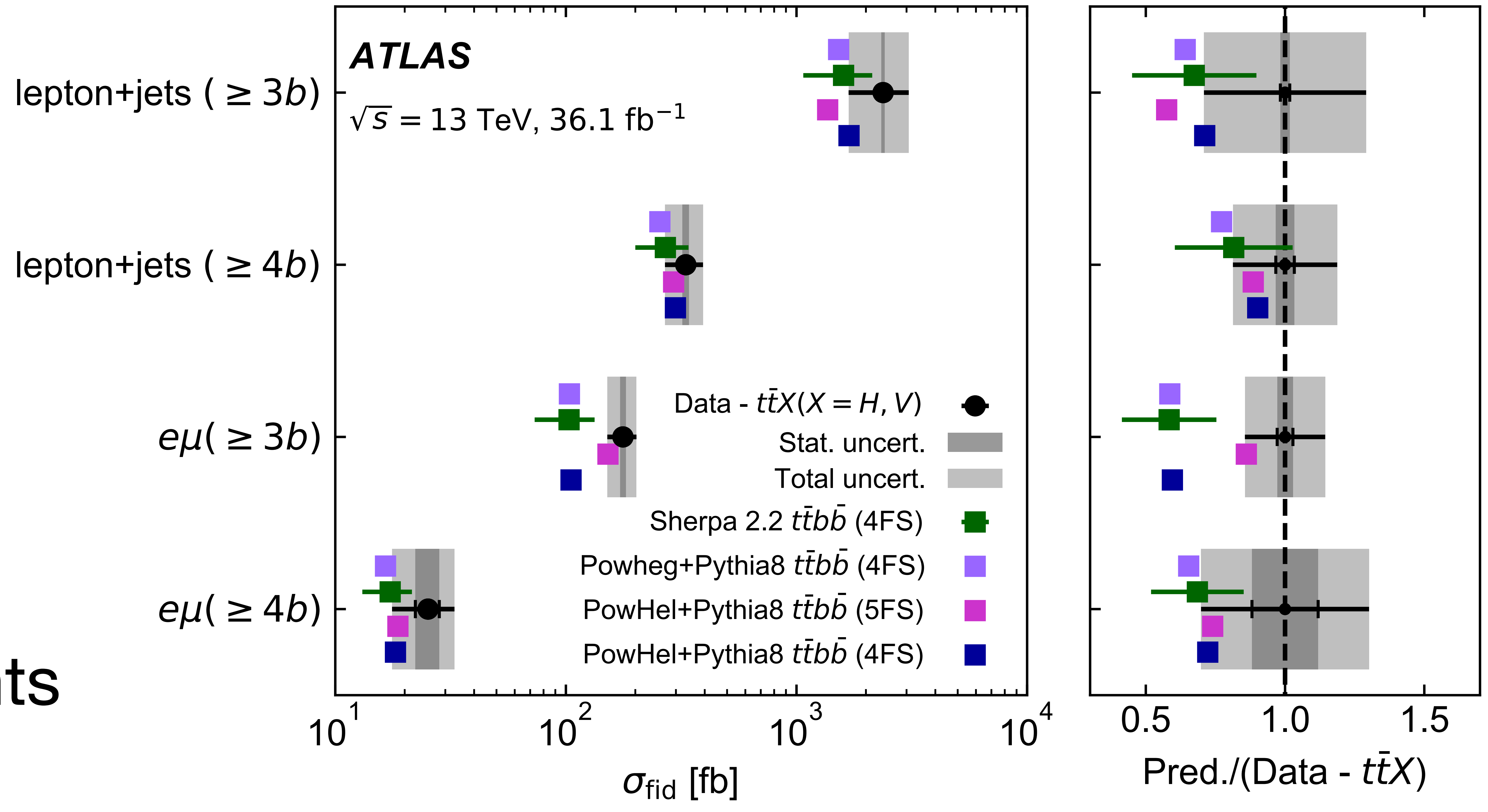


Template of third and fourth highest b-tagged jets in lepton+jets channel

JHEP 04 (2019) 046

Inclusive cross section measurement

- $t\bar{t}H$ and $t\bar{t}V$ contributions are subtracted to facilitate the comparison with theory
- $e\mu$ channel shows more precise measurement
- Observed generally lower prediction than the measurements



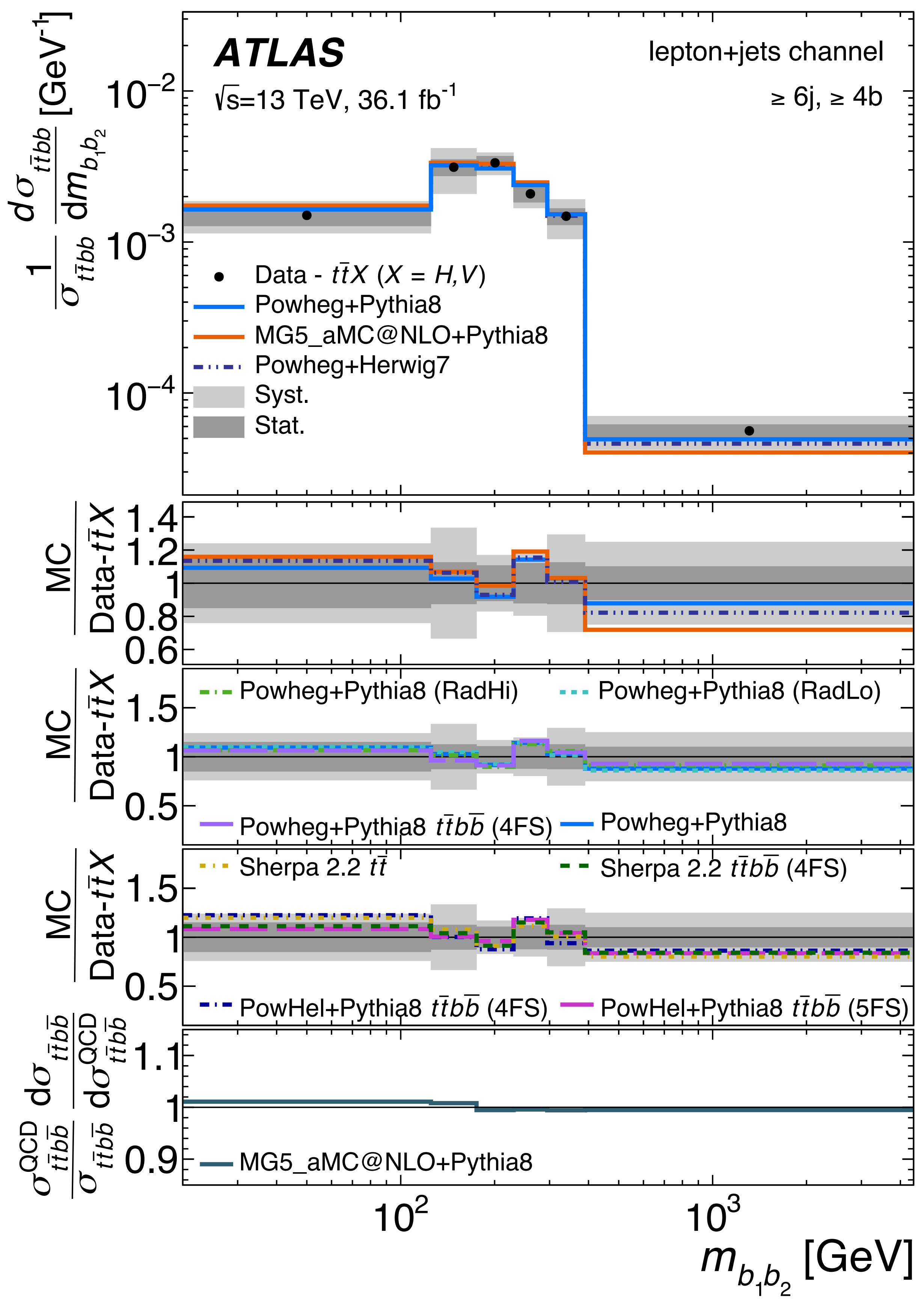
Visible Phase Space

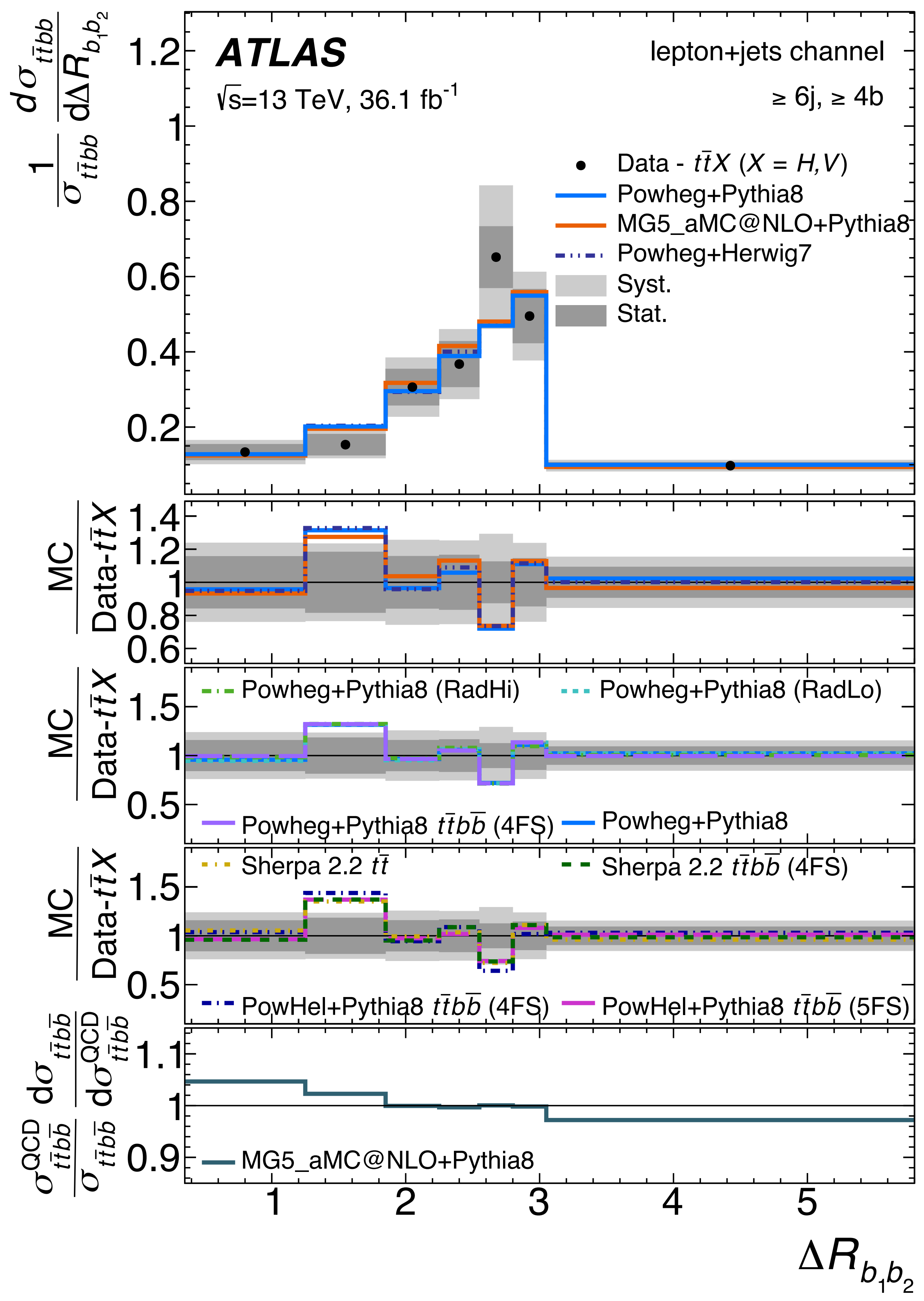
JHEP 04 (2019) 046

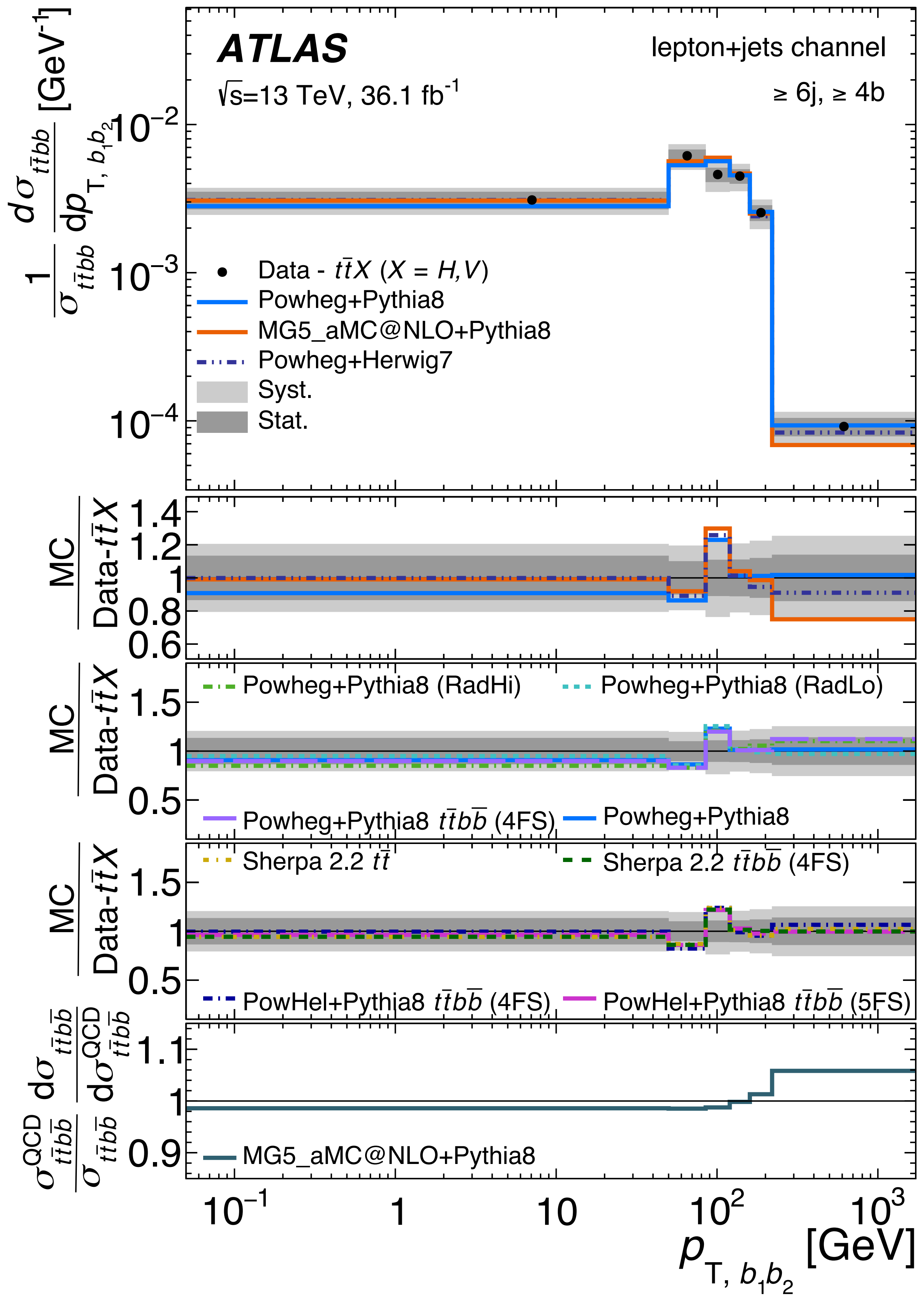
Differential cross section

Extra b jets using the highest p_T b jets

- $\geq$6j , $\geq$4b phase space:measurements are generally consistent with predictions





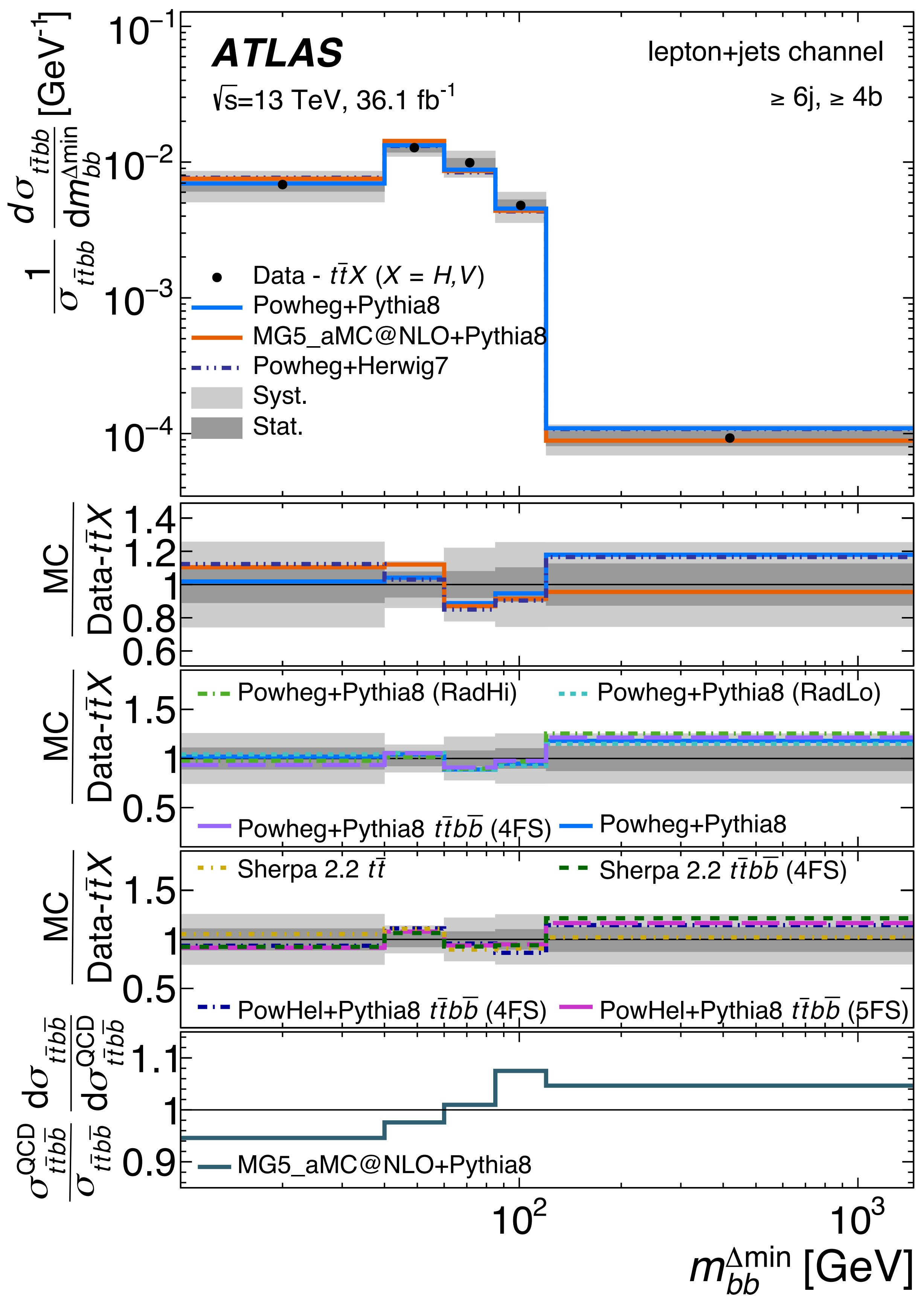


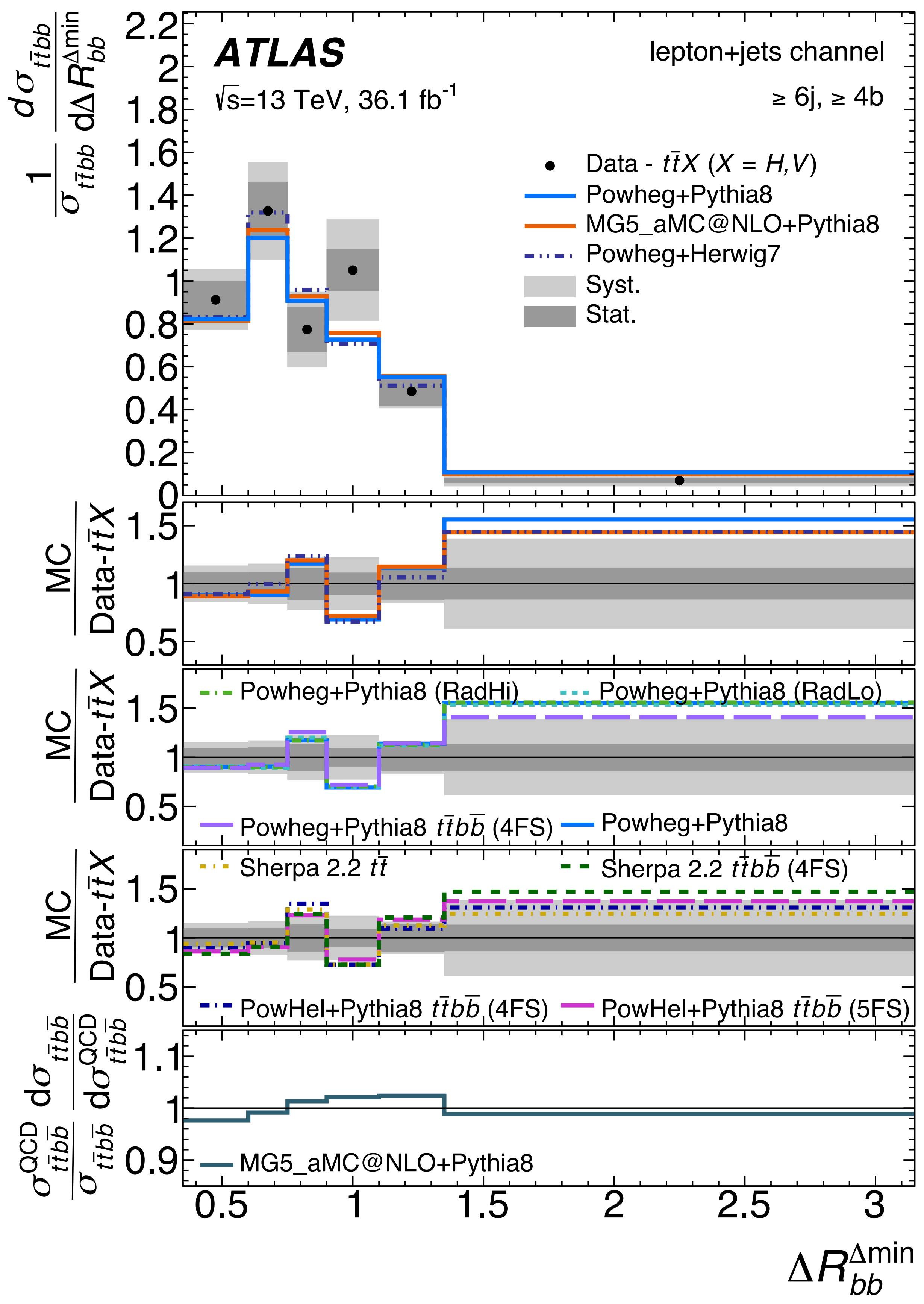
JHEP 04 (2019) 046

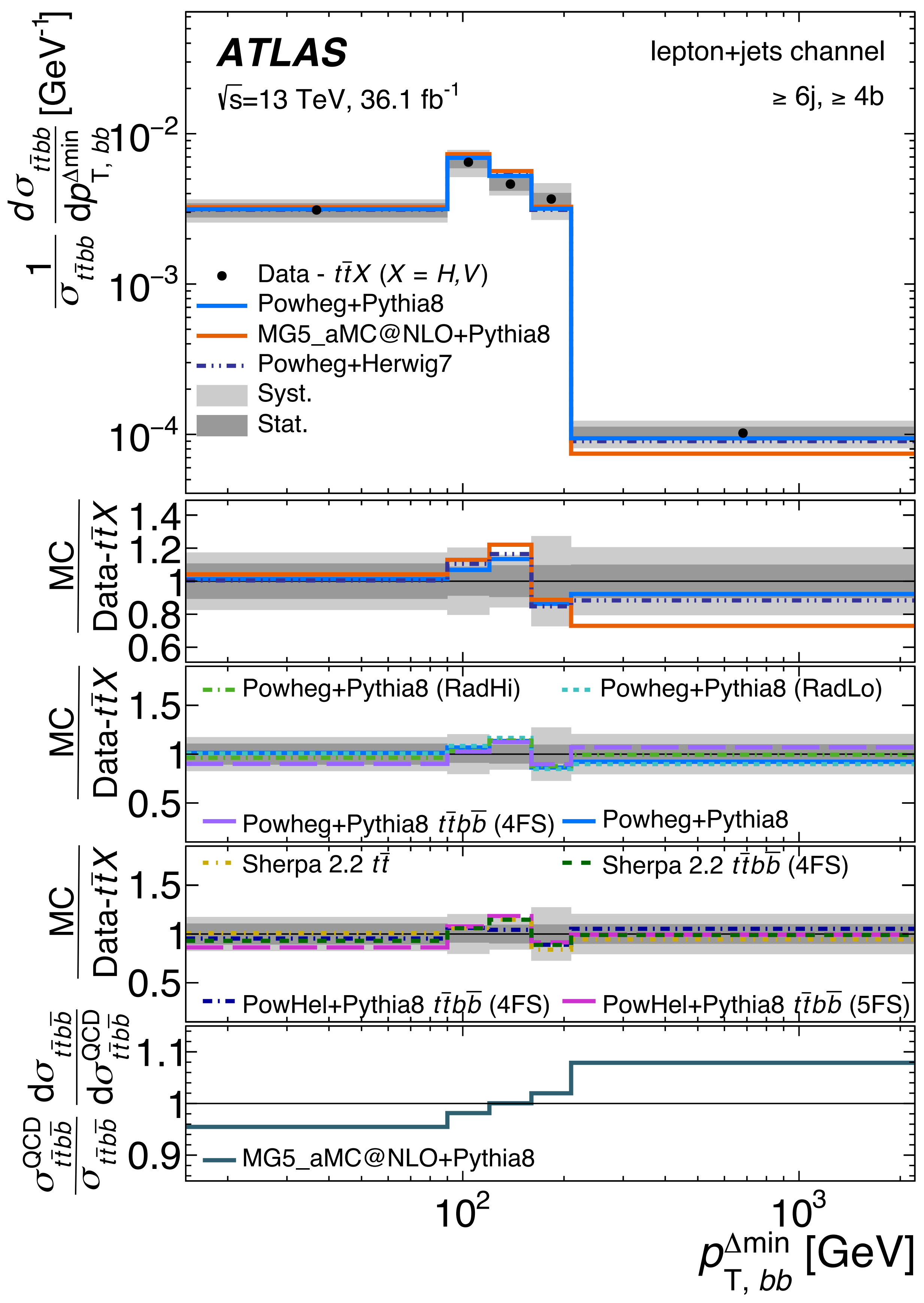
Differential cross section

Extra b jets using the closest b jets

- ≥6j , ≥4b phase space:measurements are generally consistent with predictions





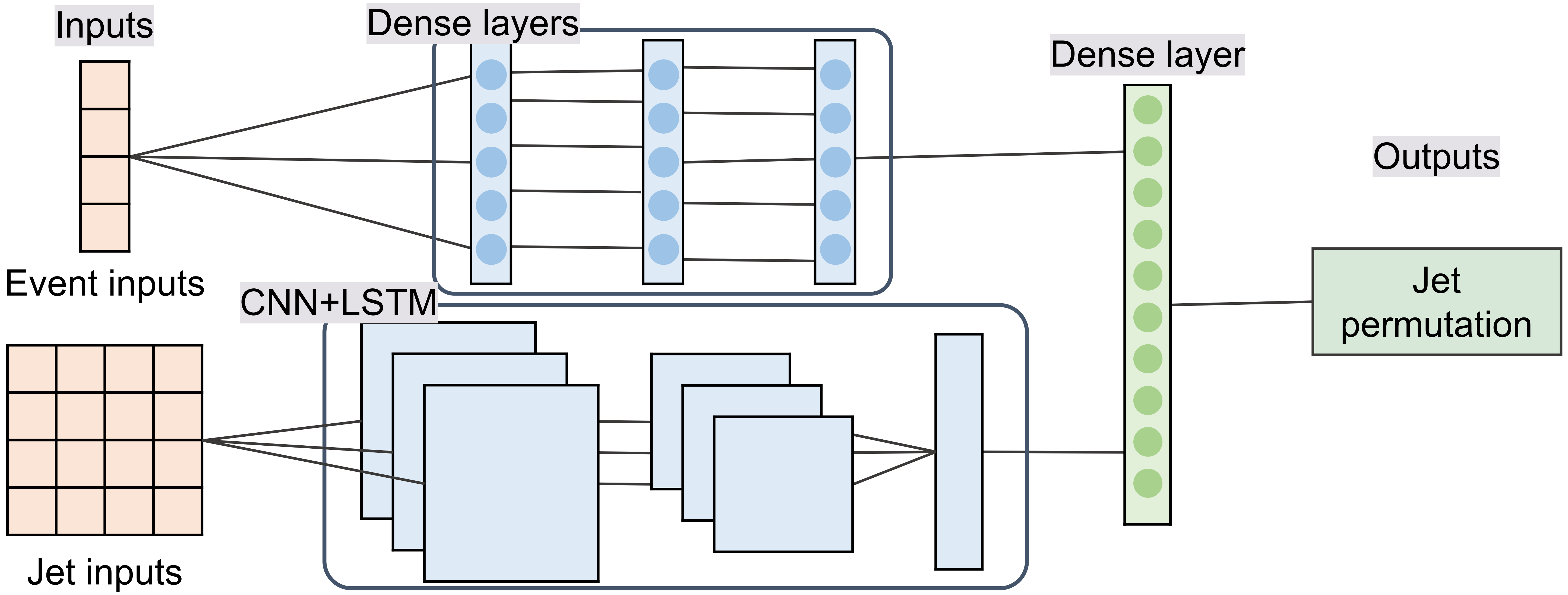


arXiv:2309.14442

Accepted in JHEP

DNN for additional b jets

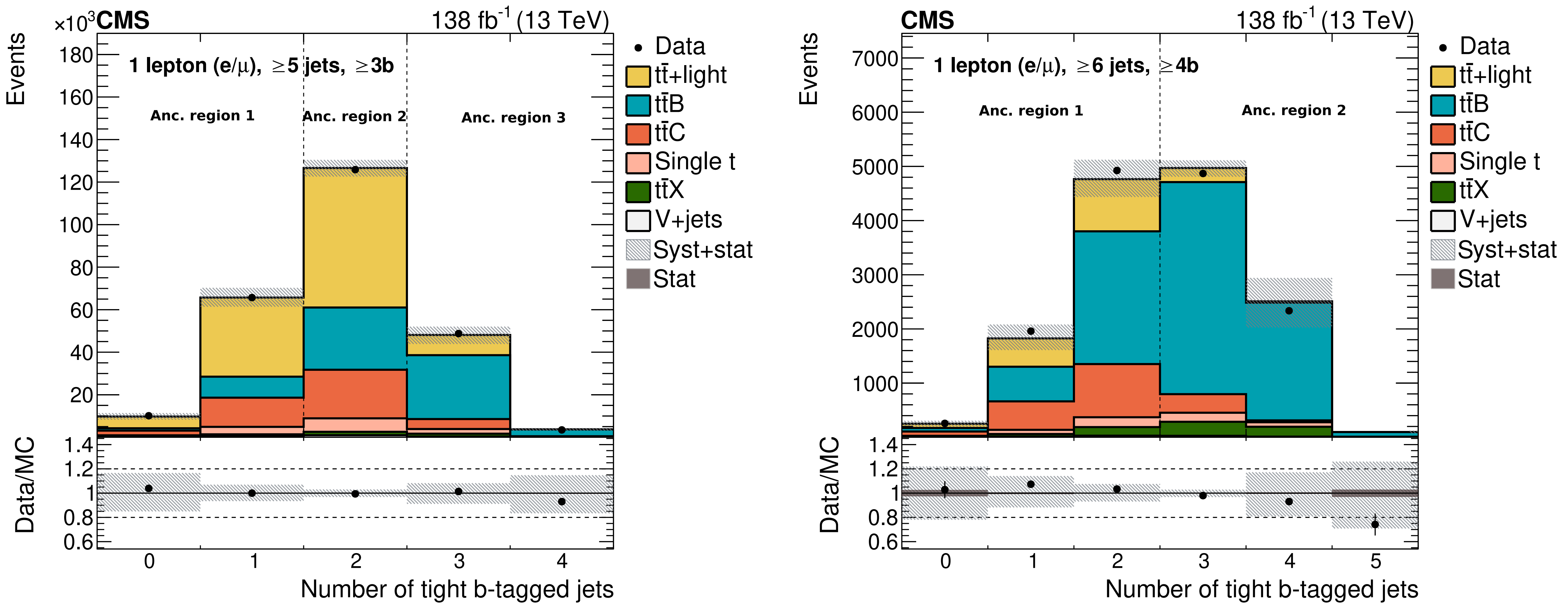
To identify b jet not from top quark in 6j4b category



- Combinations of four highest p_T b-jets : 6 combinations (6 output nodes)
- Jet specific variables (jet p_T, jet η, b-tag, m_{bl}, Δ_{bl}) for CNN+LSTM and global event variables (sum of jet p_T, b-tagged multiplicity, lepton specific variables, etc…) for Dense layers : both are connected later
- The highest output of the pair per event is selected : correct assignment of b-jets ~ 49%

Unfolding

[arXiv:2309.14442](https://arxiv.org/abs/2309.14442)



- b-jet multiplicity as Ancillary variables
    - Divide signal and background regions
- Unfolded to the particle level by removing detector effect and acceptance
- Maximum likelihood fitting is performed

$$L(\vec{\mu},\vec{\alpha}) = \left[\prod_{e,i} \mathrm{Poi}\left(D_{e,i} \middle| S_{e,i}(\vec{\mu},\vec{\alpha}) + \sum_{p\in \mathrm{bkg.}} N^{p}_{e,i}(\vec{\alpha})\right)\right] \mathcal{N}(\vec{\alpha})$$

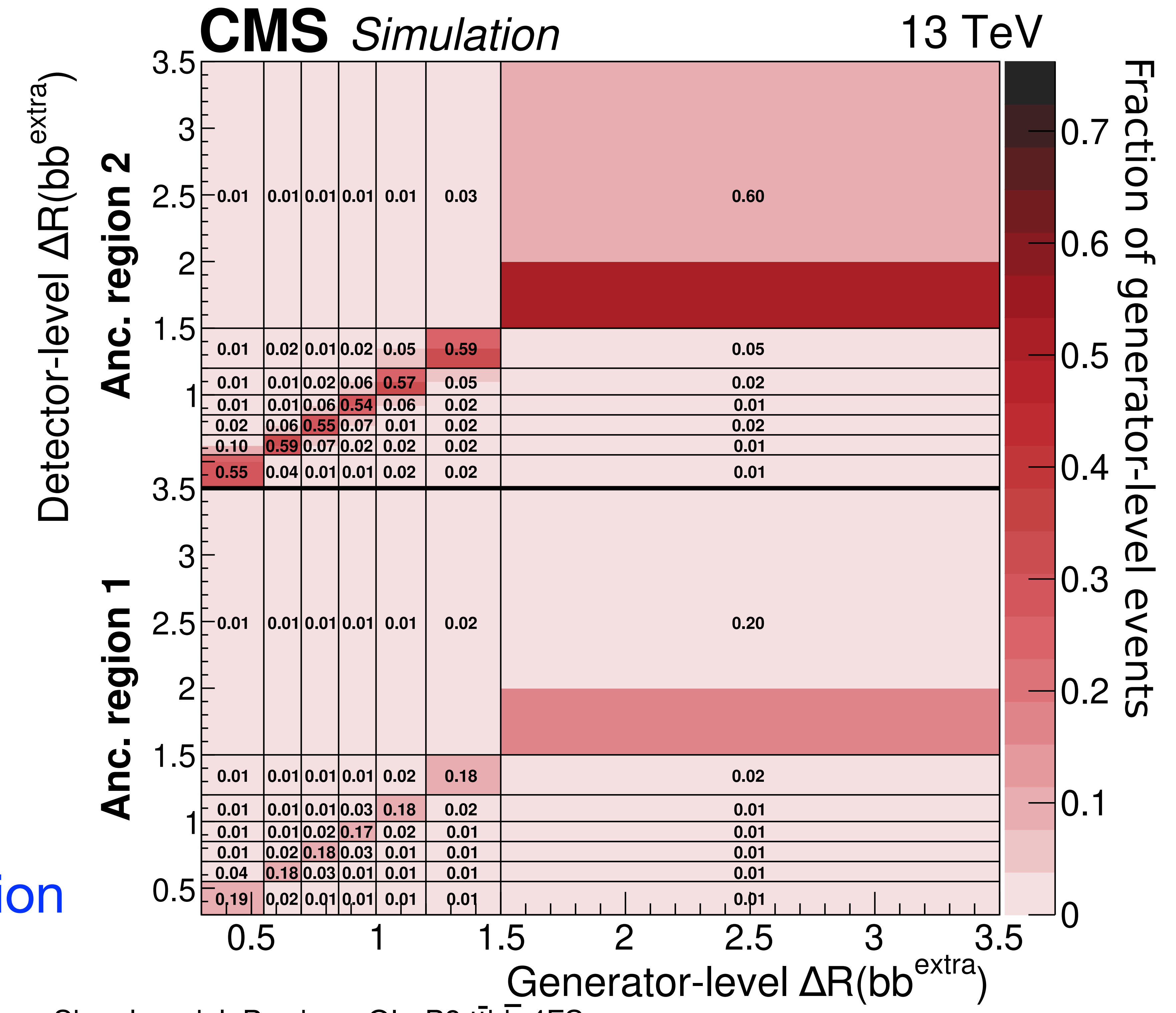
$$S_{e,i}(\vec{\mu},\vec{\alpha}) = \mu_{\mathrm{fid}} \sum_{j=1}^{n} \mu_j M^{e}_{ij}(\vec{\alpha}).$$

i = detector-level bin, j = generator-level bin

M_{ij} is expected event using the response matrix

μ_{fid} = signal-strength modify for the inclusive cross section

μ_j = parameters for the fraction of signal in each i bin



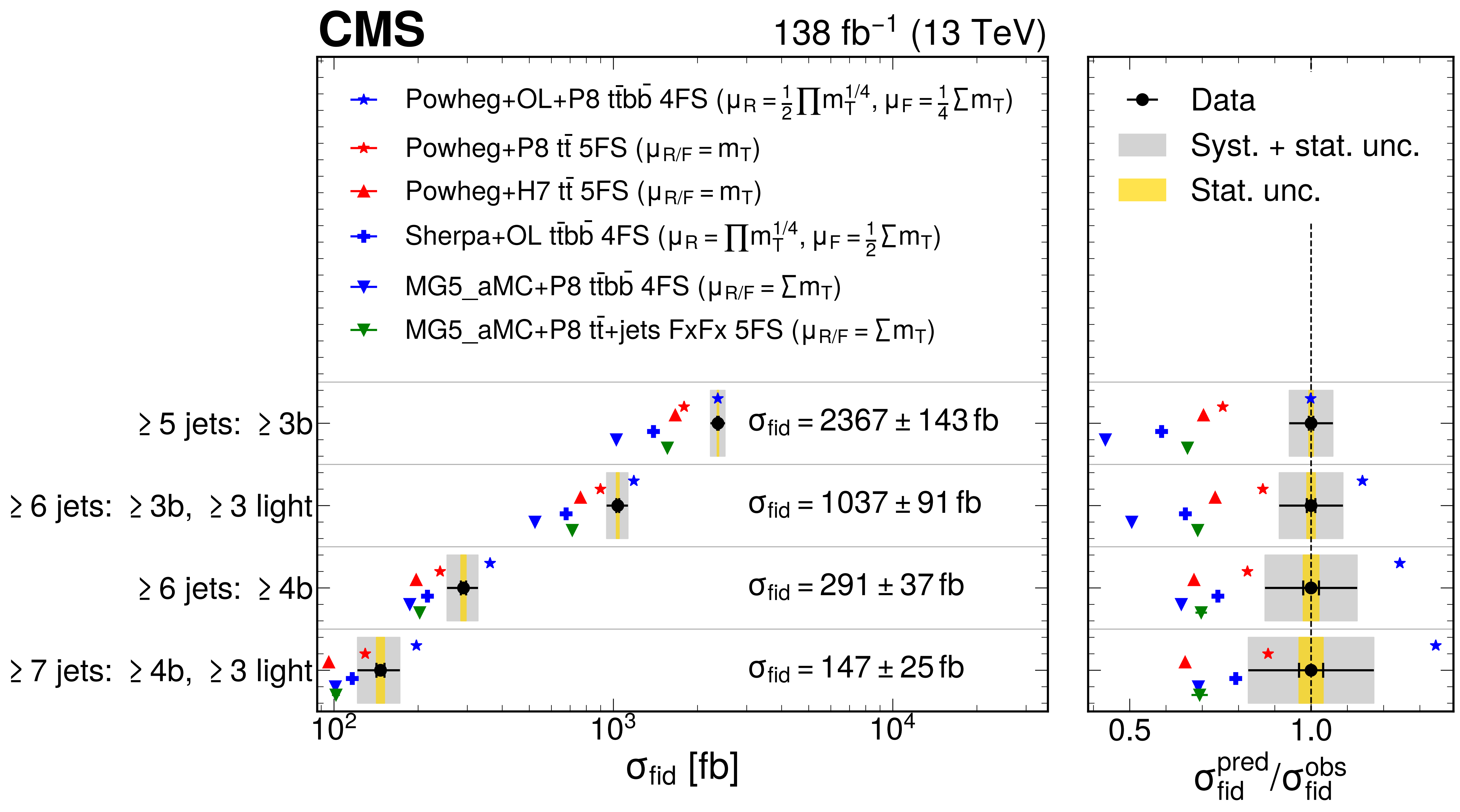
Signal model: Powheg+OL+P8 tt̄bb̄ 4FS

arXiv:2309.14442

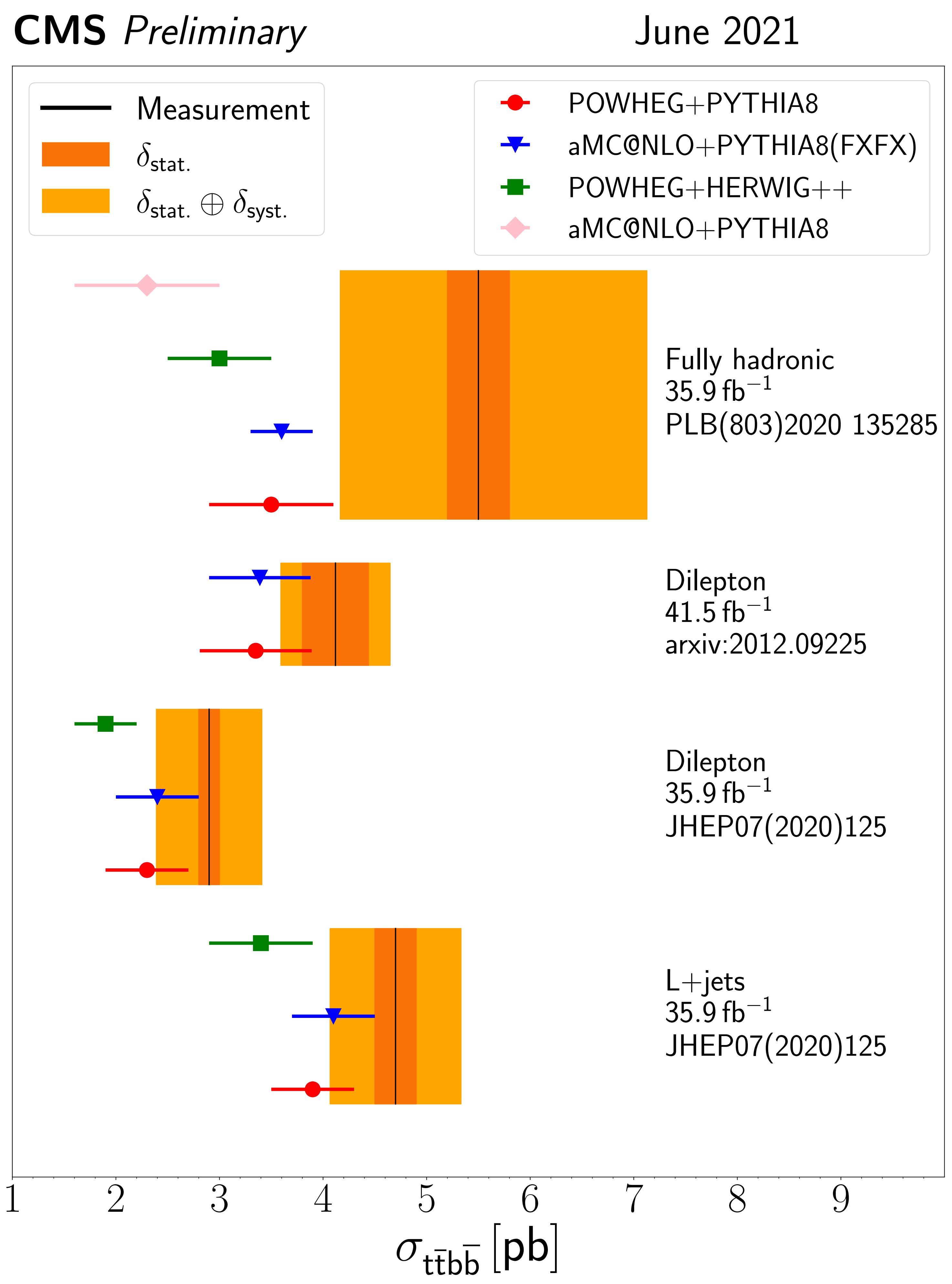
Inclusive cross section measurements

- 4 different phase spaces
- Observed the same trend as before - measured value is higher than the prediction

Visible Phase Space



CMS Top Quark Summary Figures
Full Phase Space

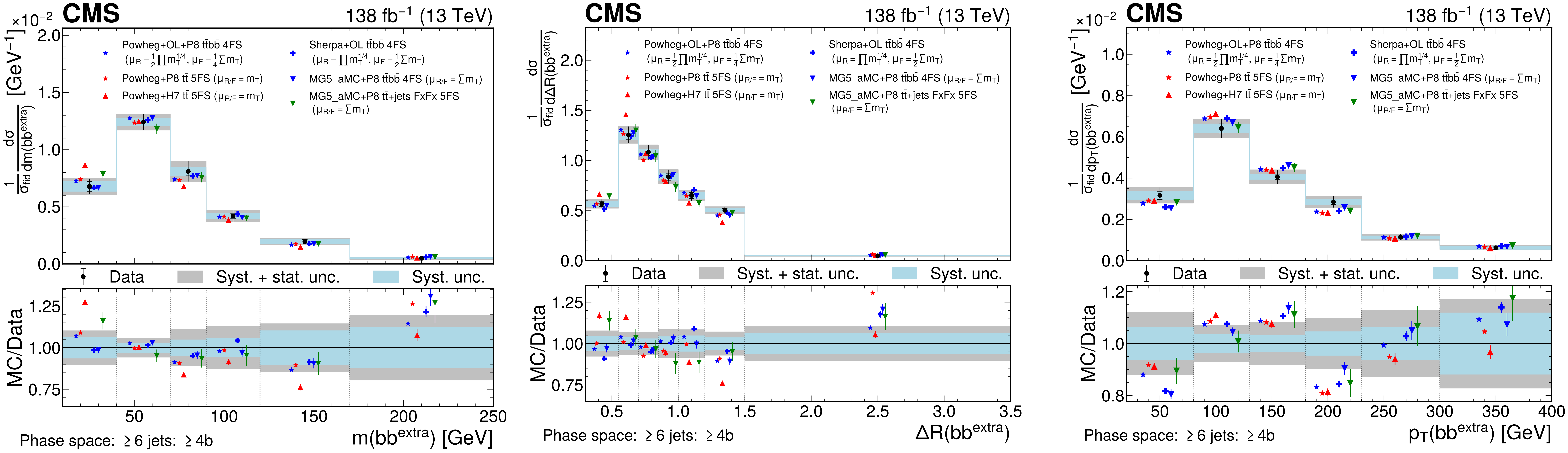


arXiv:2309.14442

Differential cross section

Extra b jets using the closest b jets

- ≥6j , ≥4b phase space: HERWIG tends to produce two additional b jets with smaller angle than data

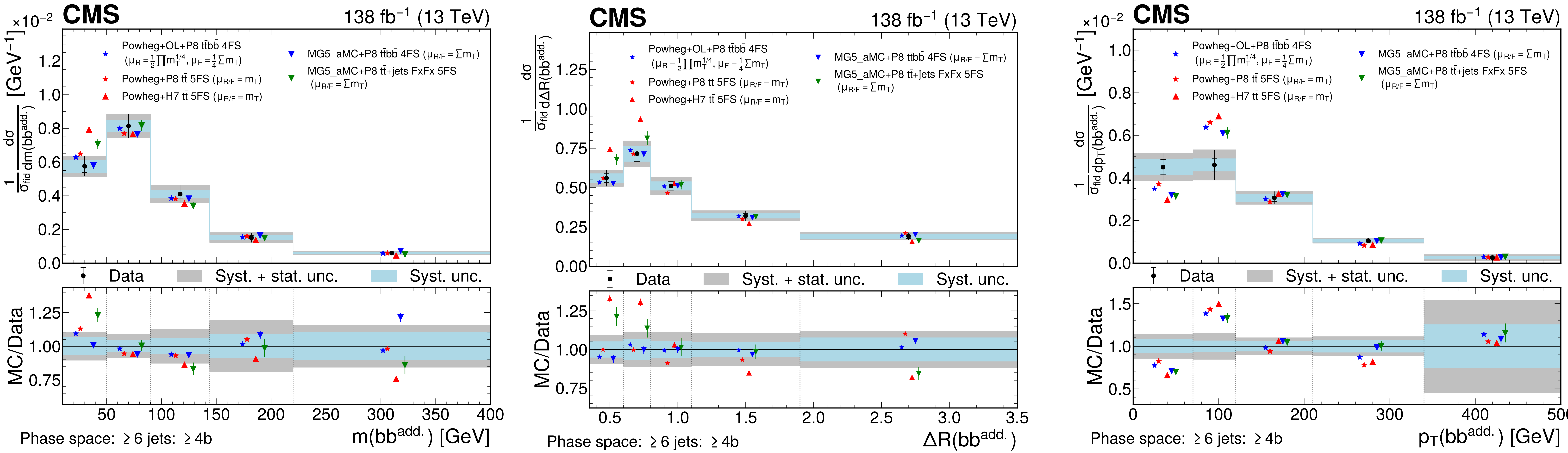


arXiv:2309.14442

Differential cross section

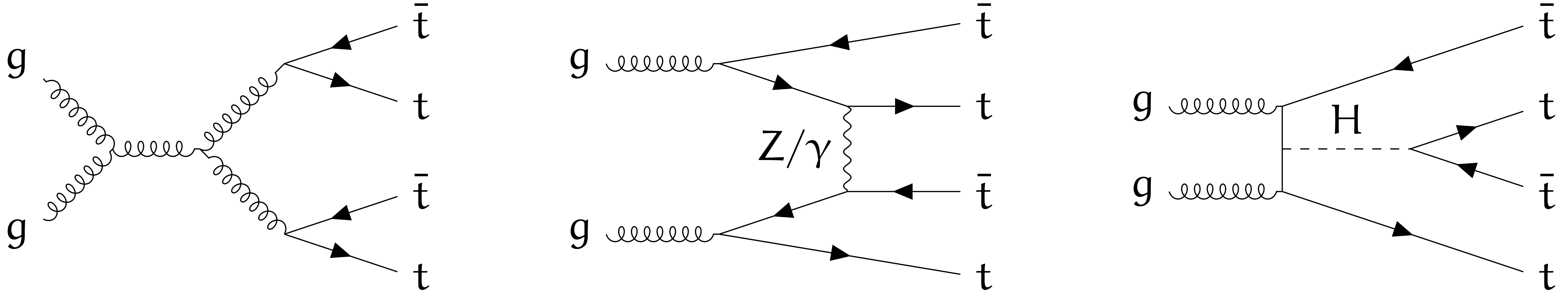
Additional b jets not from top quark

- ≥6j , ≥4b phase space: HERWIG tends to produce two additional b jets with smaller angle than data

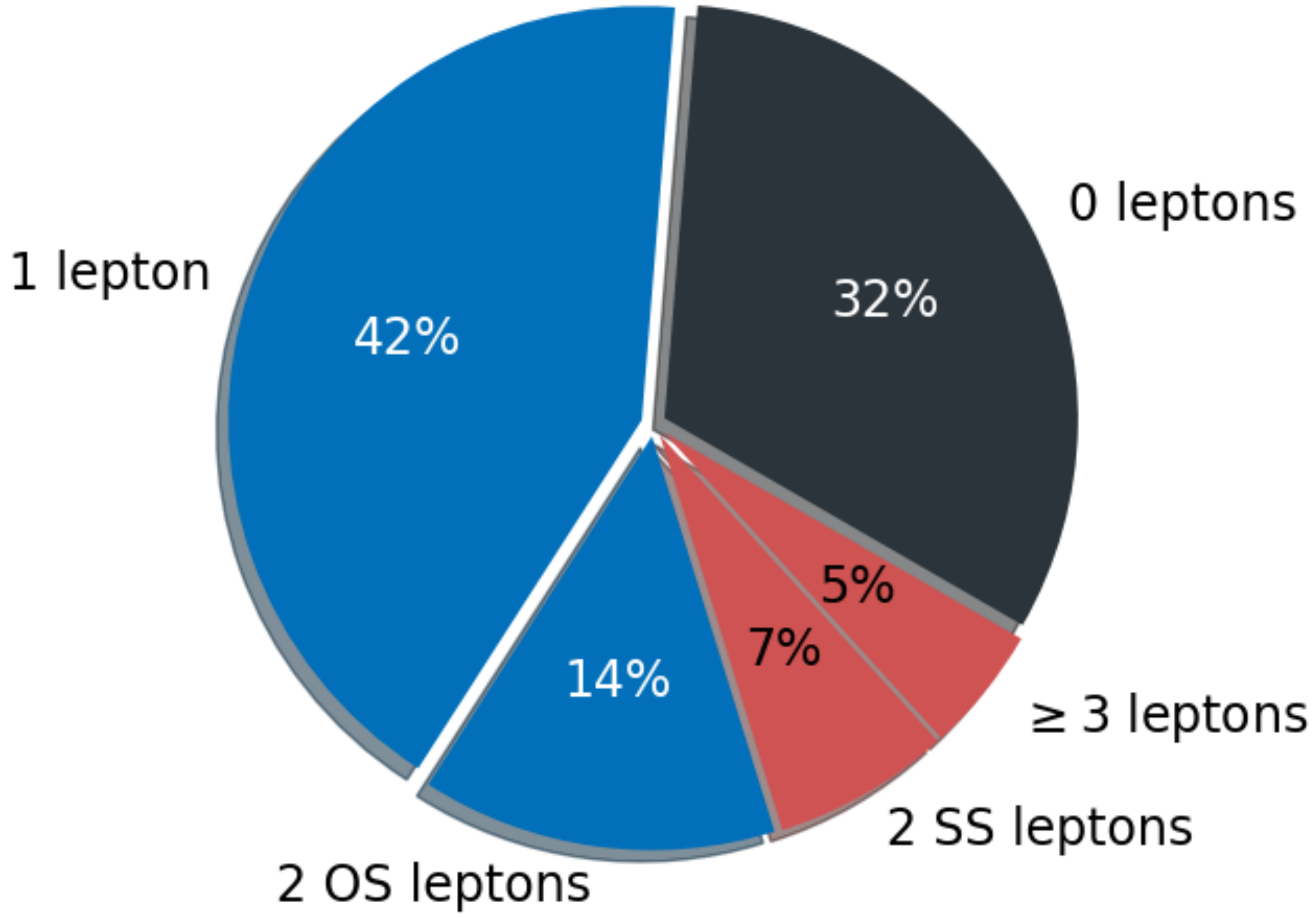


Four top searches

- Four top production is a very rare SM process
  - 12.0 ± 2.4 fb (NLO) - JHEP 02 (2018) 031
  - $13.4^{1.0}_{-1.8}$ fb (NLO+NLL') - arXiv:2212.03259 (see the talk by Anna Kulesza)
- Probe of top-Higgs Yukawa coupling
- This process is the heaviest final state observed at the LHC
- Sensitive to new physics and effective field theory operator



Four top final states



- Four top quarks have large object multiplicity
  - 4 b quarks (jets) and the decay products of 4 W bosons
- Three different channels
  - All hadronic (0L)
  - Single lepton and opposite sign dilepton (1L, OSDL) : Large branching ratio and $t\bar{t}$ background
  - Same-sign dilepton and multilepton (SSDL, ML) : smaller branching ratio and clean signature
- Heavy use of machine learning techniques to maximize signal to background ratio
  - Boosted Decision Trees (BDT)
  - Graph Neural Networks (GNNs)

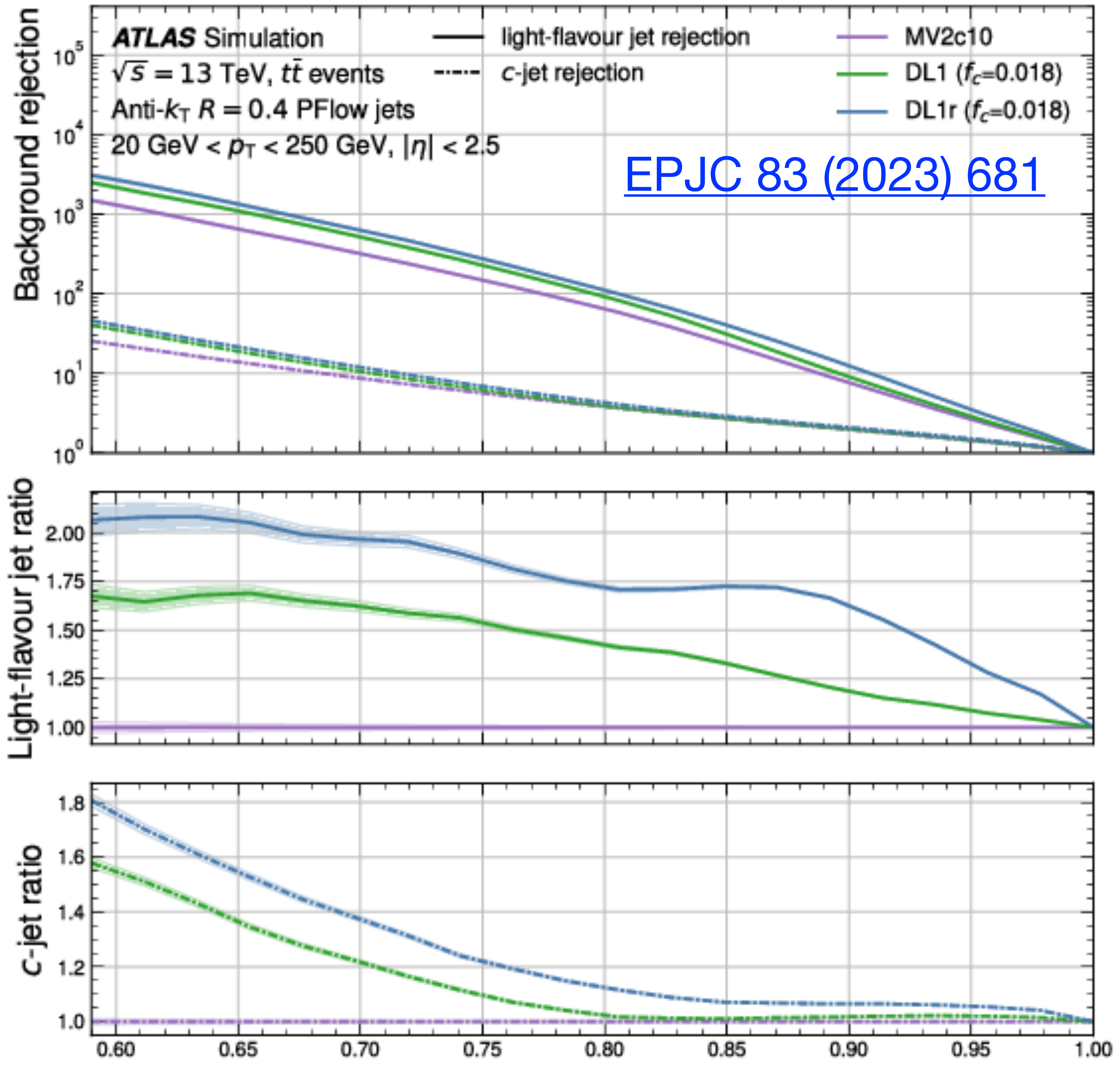
[ATLAS : EPJC 83 (2023) 496](#)

[CMS : PLB 847 (2023) 138290](#)

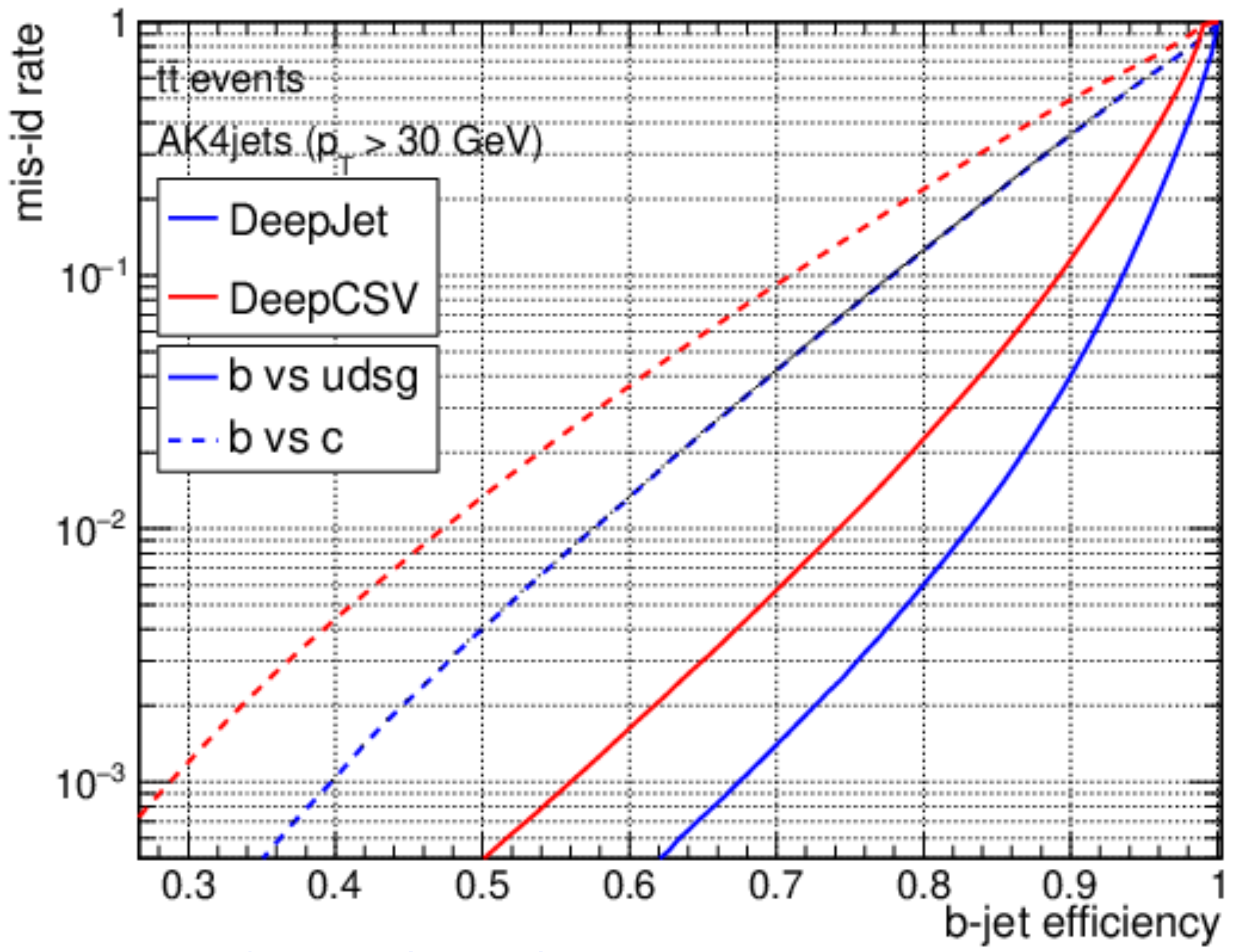
Observation

First observation from ATLAS and CMS

- Re-analysis of Run 2 data for SSDL and ML channels
  - better lepton identification method
  - improved b-tagging
- Major improvements
  - GNN (ATLAS) and multiclass BDTs (CMS)
  - Better estimation of $t\bar{t}X$ backgrounds
  - Better handling the uncertainty on 3 tops



EPJC 83 (2023) 681



JINST 15 (2020) P12012

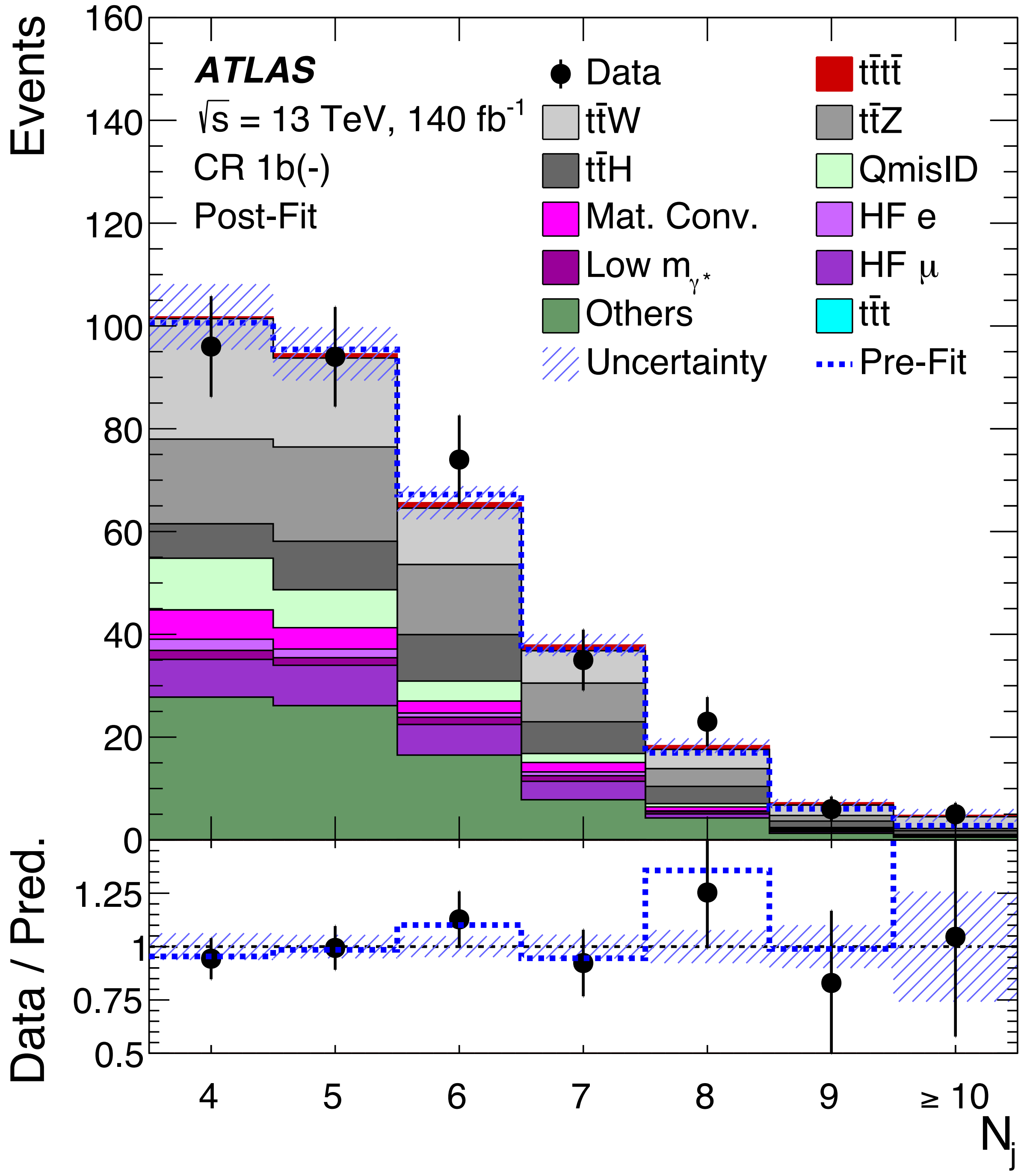
ATLAS : EPJC 83 (2023) 496

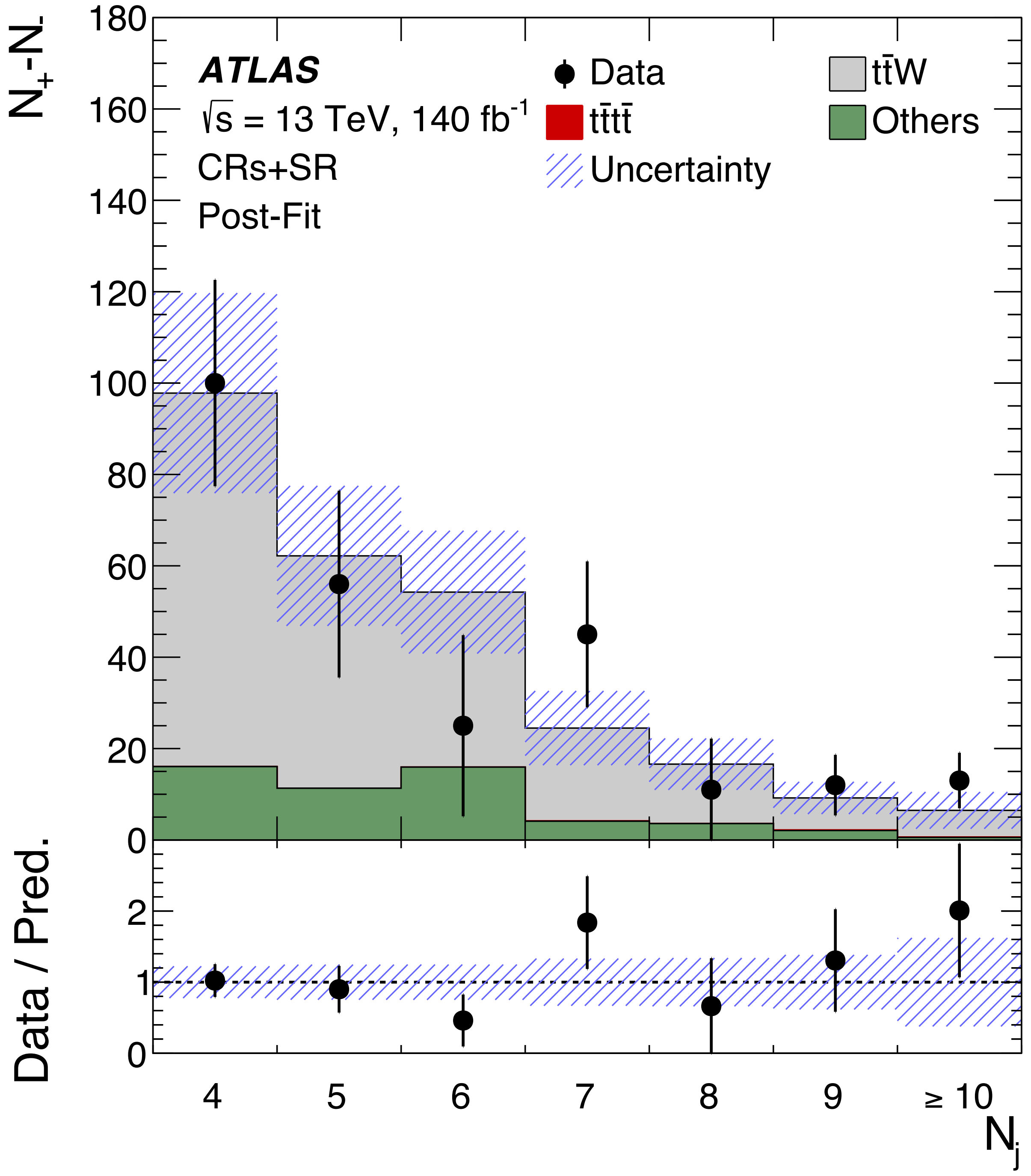
Background modeling (ATLAS)

$t\bar{t}W$ modeling

- N_{jets} distributions are corrected using data (ATLAS)
  - $R(j) = N(j+1)/N(j)$, j is the jet multiplicity
  - Staircase scaling: $R(j) = a_0$ for high jet multiplicities
  - Poisson scaling: $R(j) = a_1/(1+n)$, n is num. of additional jets
- 4 dedicated control regions to determine a_0, a_1, $2NF_{4jets}$
- Use the difference $N_+ - N_-$ to validate $t\bar{t}W$

$t\bar{t}W$ background	a_0	a_1	$\text{NF}_{t\bar{t}W^+(4\text{jet})}$	$\text{NF}_{t\bar{t}W^-(4\text{jet})}$
Value	0.51 ± 0.10	$0.22^{+0.25}_{-0.22}$	$1.27^{+0.25}_{-0.22}$	$1.11^{+0.31}_{-0.28}$



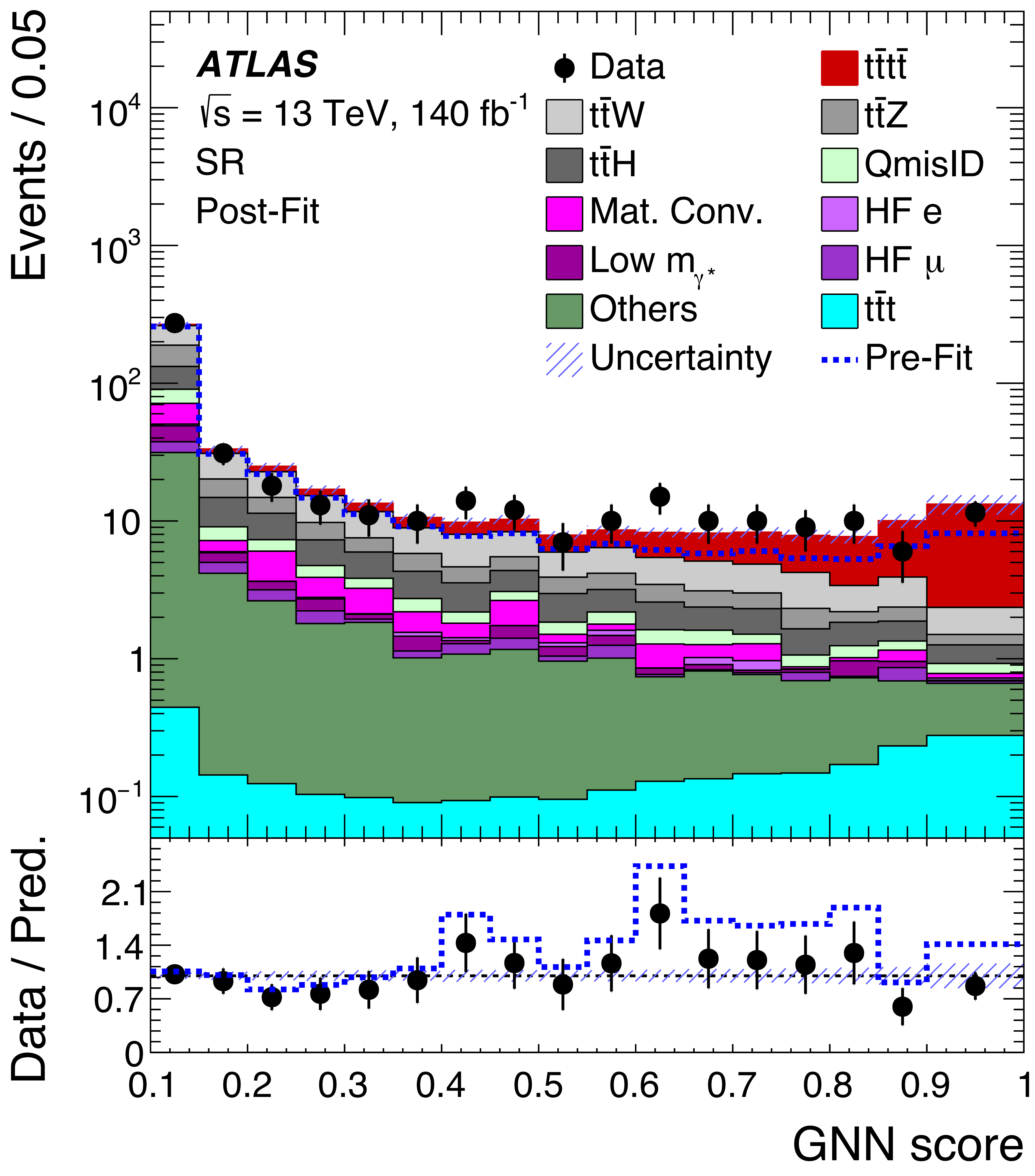


[ATLAS : EPJC 83 (2023) 496]

Signal extraction (ATLAS)

- Signal extraction
  - 4 control regions for $t\bar{t}W$ and 4 control regions for non-prompt and conversions background
  - Combines SSDL and ML events for signal region
  - $\geq$ 6 jets, $\geq$2 b jets, HT (Σjet and lepton) > 500 GeV
  - Graph Neural Network (GNN) to separate signal from background
- Sensitivity: 6.1σ observed (4.7σ expected)
- Measured cross section: $22.5^{+4.7}_{-4.3}(\text{stat})^{+4.6}_{-3.4}(\text{syst})$ fb

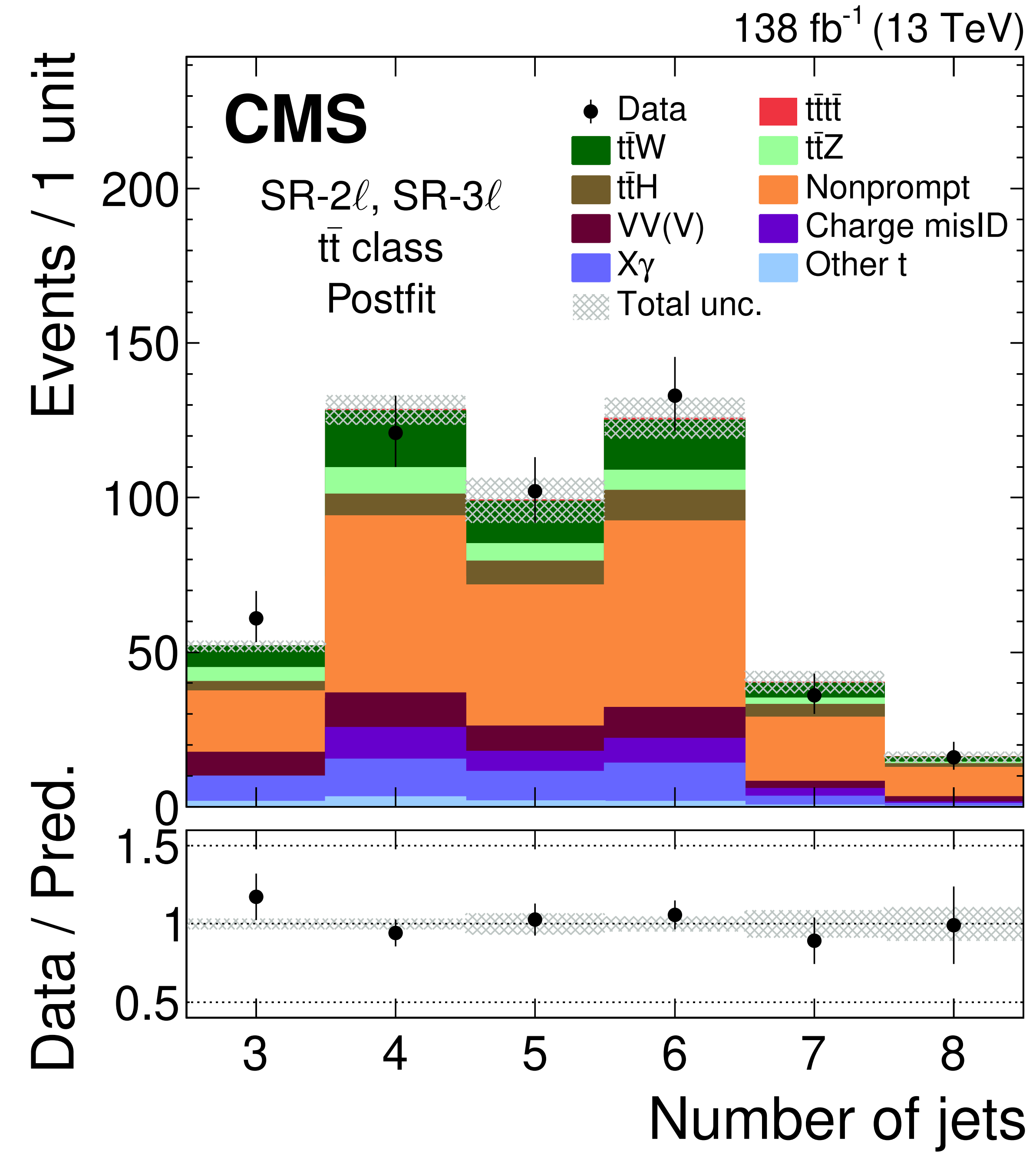
~ 27%

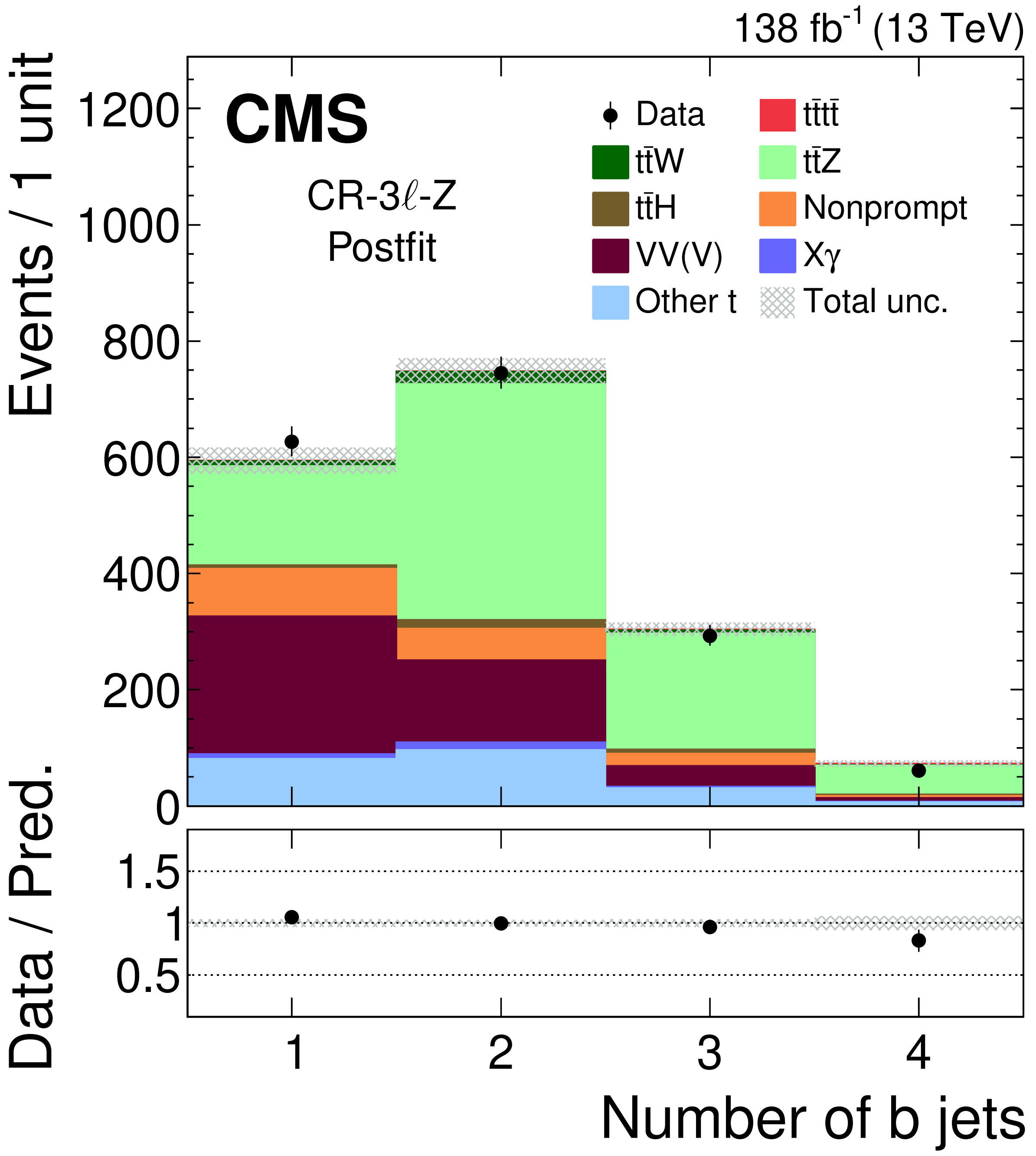


CMS : PLB 847 (2023) 138290

Background estimation (CMS)

- $t\bar{t}W$ modeling: NLO QCD MC
  - Additional large uncertainty on $t\bar{t}W$ + jets (mis-modeling additional (b) jets)
  - Free-floating normalization in fit
    - Postfit normalization: 990 ± 98 fb (compatible with CMS measurement) JHEP 07 (2023) 219
    - Constrained by 2 control regions and multi-class BDT
- Z control regions (3 and 4 lepton channels)
  - $|m_{ll} - m_Z| < 15$ GeV
  - Allow for free-floating $t\bar{t}Z$ normalization in fit
    - Postfit normalization: 945 ± 81 fb (compatible with CMS measurement) EPJC 80 (2020) 428
  - Control over WZ & ZZ with additional (b) jets

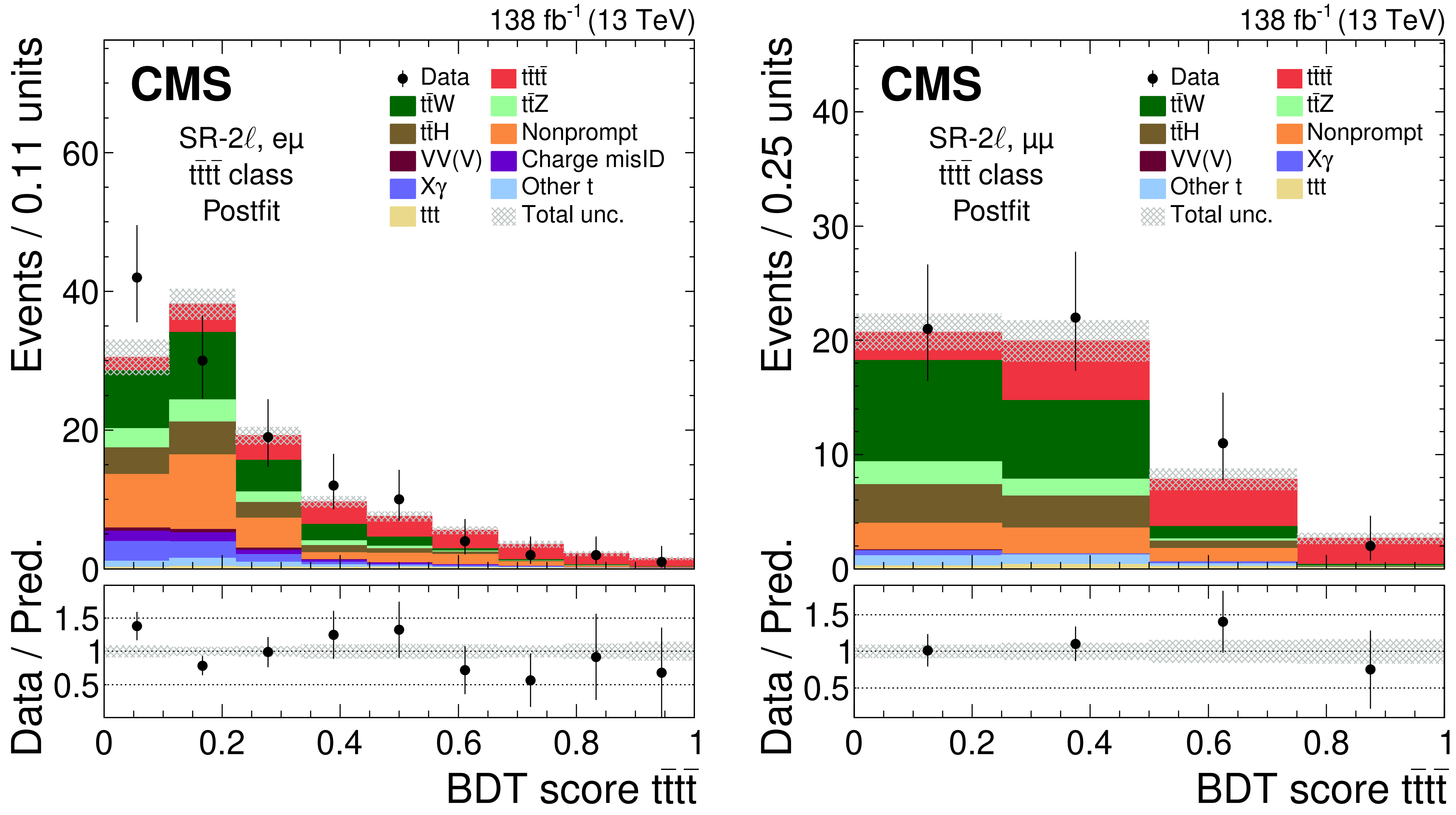


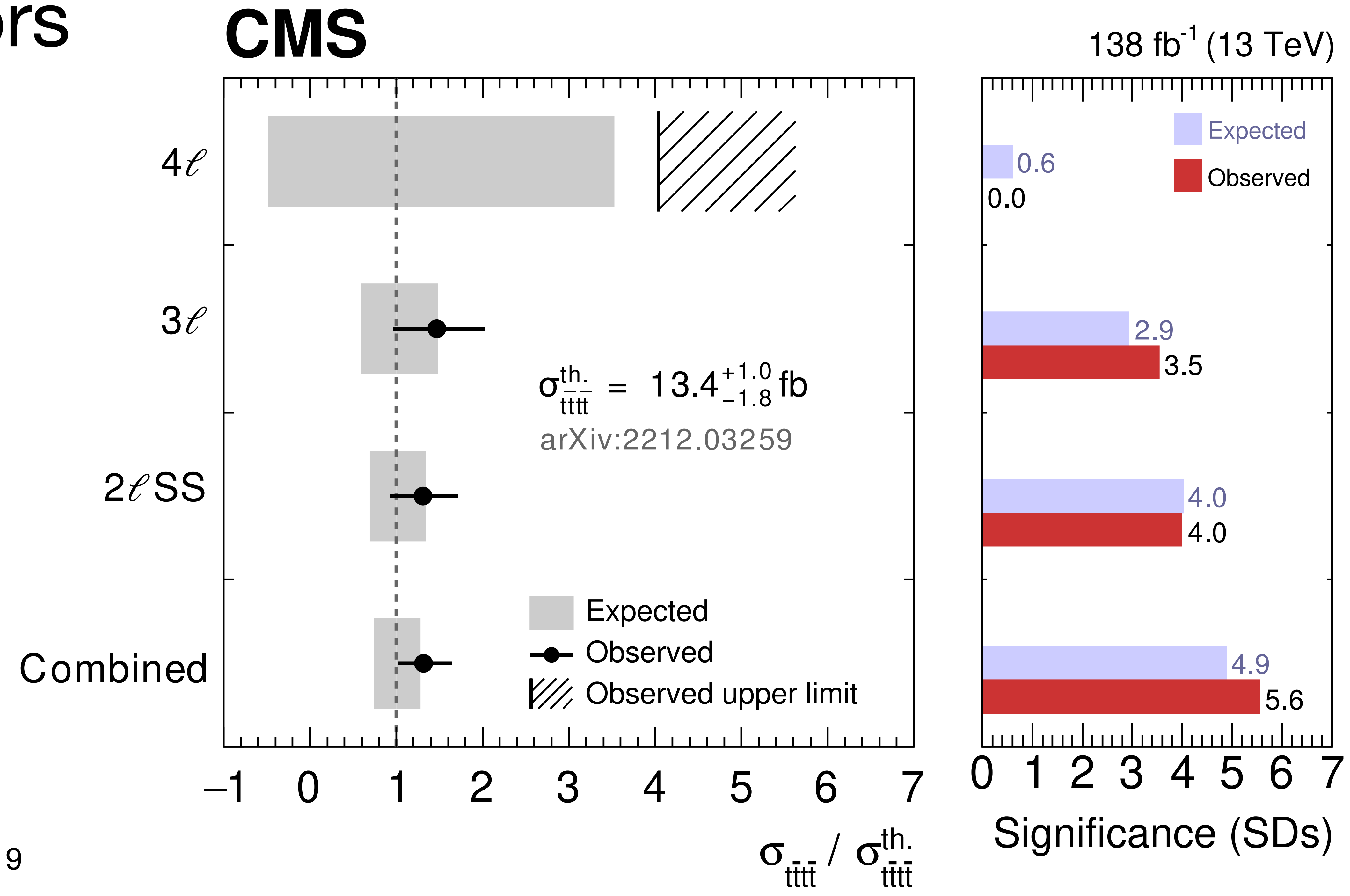


CMS : PLB 847 (2023) 138290

Signal extraction (CMS)

- Signal extraction
  - 3 signal regions - SSDL, 3L and 4 L
  - BDT multi-classification : $t\bar{t}t\bar{t}$, $t\bar{t}X$, $t\bar{t}$
  - SSDL : split into 3 different lepton flavors
- Sensitivity
  - 5.6σ observed (4.9σ expected)
- Measured cross section
  - $17.7^{+3.7}_{-3.5}(\text{stat})^{+2.3}_{-1.9}(\text{syst})$ fb ~ 24%



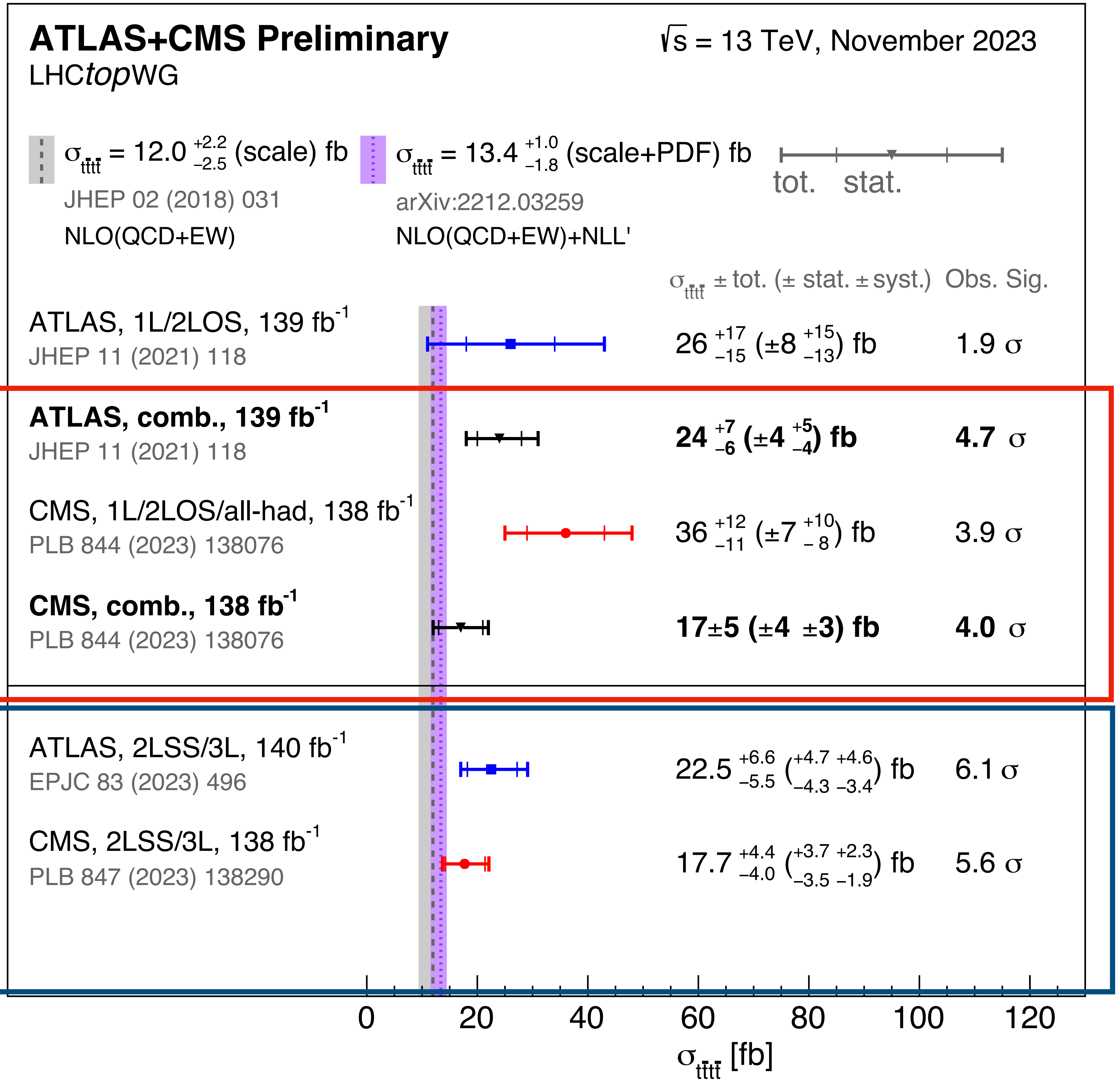


Four top summary

Evidence

Observation

In this talk

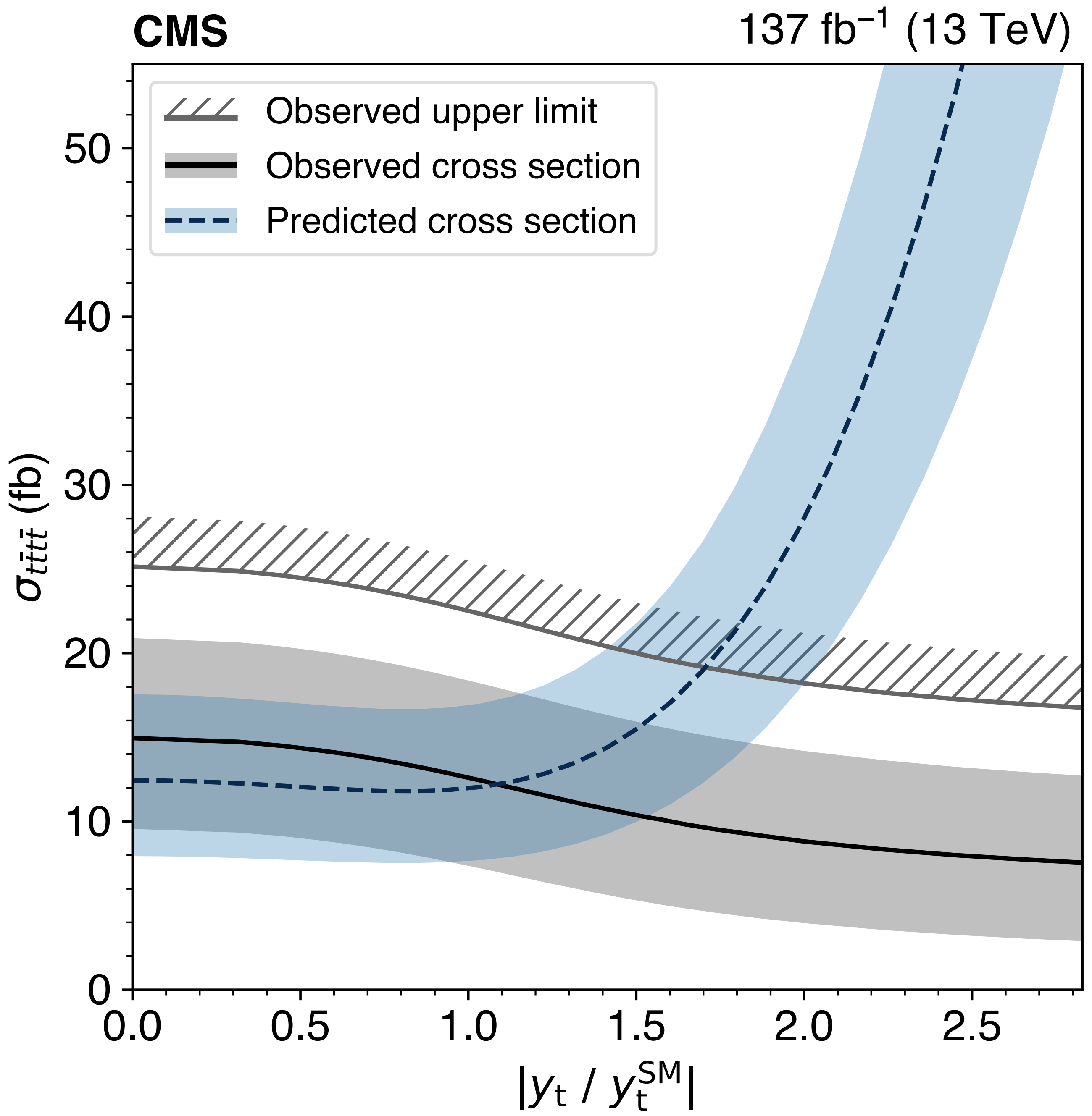


[ATLAS : EPJC 83 (2023) 496](#)

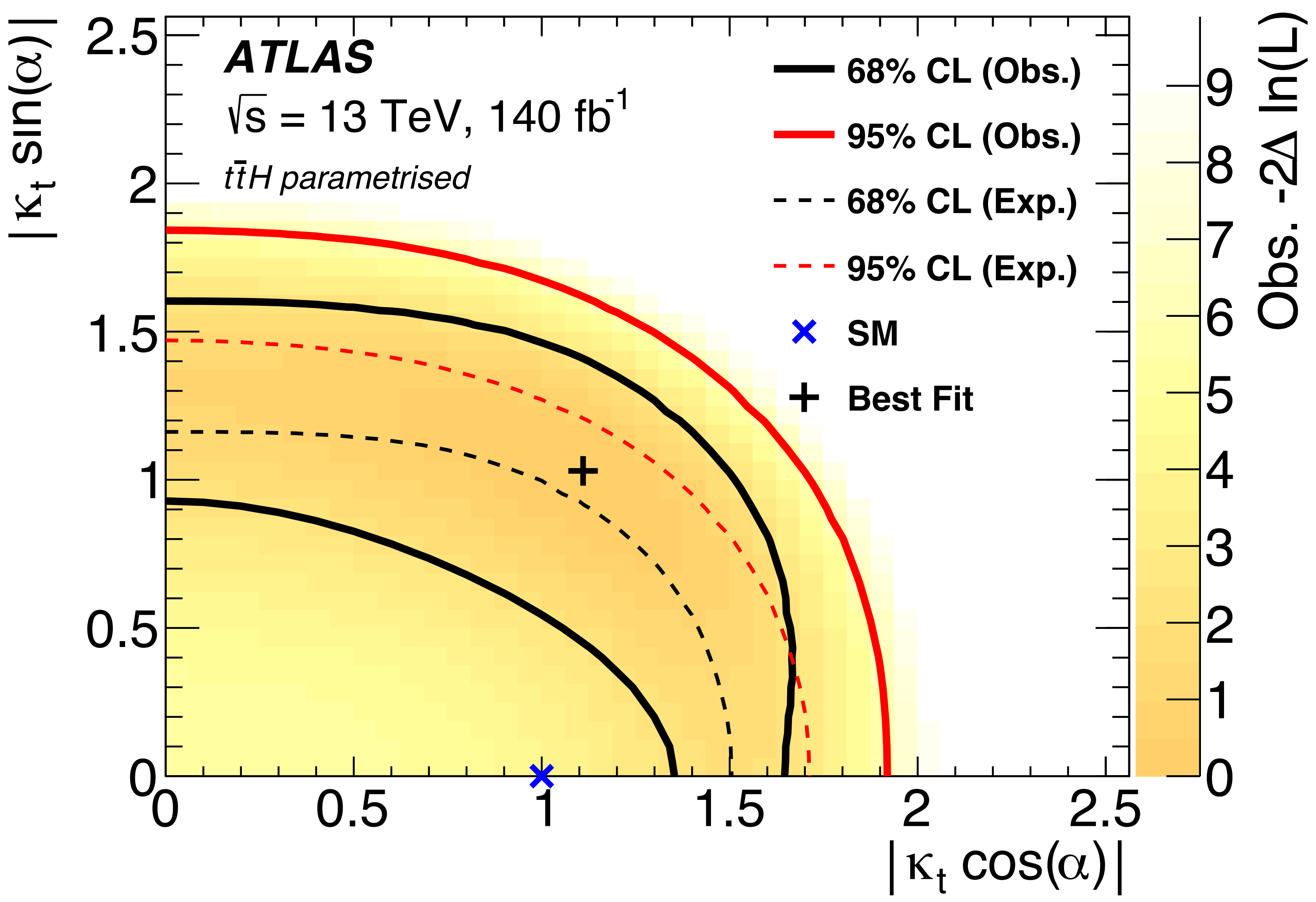
[CMS:EPJC 80 (2020) 75](#)

Interpretation

- Limits on top-quark Yukawa coupling



$k_t < 1.7$



$k_t < 1.9$

k_t = $t\bar{t}H$ signal strength

α = mixing angle between CP even and CP odd components

- Limits on EFT operators (ATLAS)
  - $t\bar{t}t\bar{t}$ is sensitive to four heavy fermion operators
  - probe the BSM models

$$\sigma_{t\bar{t}t\bar{t}} = \sigma^{SM}_{t\bar{t}t\bar{t}} + \frac{1}{\Lambda^2}\sum_i C_i \sigma_i^{(1)} + \frac{1}{\Lambda^4}\sum_{i\leq j} C_i C_j \sigma_{i,j}^{(2)}$$

Operators	Expected C_i/Λ^2 [TeV $^{-2}$]	Observed C_i/Λ^2 [TeV $^{-2}$]
O^1_{QQ}	[-2.5, 3.2]	[-4.0, 4.5]
O^1_{Qt}	[-2.6, 2.1]	[-3.8, 3.4]
O^1_{tt}	[-1.2, 1.4]	[-1.9, 2.1]
O^8_{Qt}	[-4.3, 5.1]	[-6.9, 7.6]

Limits on four heavy flavor interpretation

Conclusion

- For $t\bar{t}b\bar{b}$ cross section measurements, generally predictions under-estimate cross sections
  - mis-modeling due to only NLO in QCD or new physics
- Twice more data in Run 3 and more data-driven techniques can give us more information
- Should make use of EFT approach for possible new physics for the $t\bar{t}b\bar{b}$ and $t\bar{t}t\bar{t}$
  - Differential measurement may be crucial in this approach