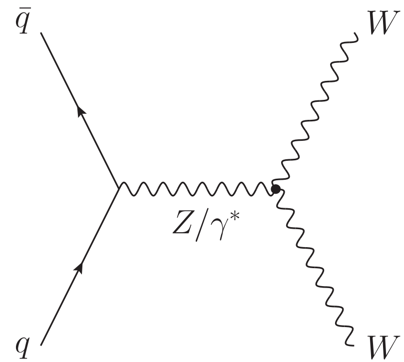
Measurements of $W^+W^- + \geq 1$ jet cross sections at $\sqrt{s}$ = 13 TeV with the ATLAS detector

Jack C. MacDonald
on behalf of the ATLAS WW analysis team

LHC EW WG MB, 28th May 2021

Introduction

- **WW measurements provide precision tests of Standard Model (SM)**
  - Sensitive to properties of gauge boson self-interactions
  - Test of perturbative quantum chromodynamics (pQCD) and electroweak (EW) theory



- **Important background for $H \to WW$ measurements and BSM searches**
- **Previous WW measurements at the LHC:**
  - $\sqrt{s}$ = 7 TeV arXiv:1210.2979 arXiv:1306.1126
    $\sqrt{s}$ = 8 TeV arXiv:1603.01702 arXiv:1507.03268 arXiv:1608.03086
    $\sqrt{s}$ = 13 TeV arXiv:1702.04519 arXiv:1905.04242 arXiv:2009.00119
    (2013 → 2020)
  - All limit number of hadronic jets to reduce backgrounds

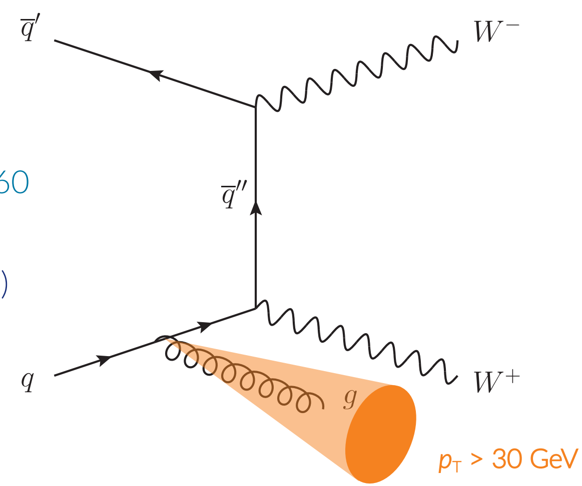
Most recent ATLAS measurement inclusive over jets → focus of this talk 2021

→ shown at Moriond in March 2021

Pre-print: arXiv:2103.10319

Motivation

- **Jet-inclusive differential measurements made for first time at LHC**
- **Improved precision in fully inclusive measurement** (when combined with jet veto measurement)
- **Improved sensitivity to BSM physics**
  - Effective field theory (EFT) interference term less helicity suppressed than in jet veto case → arXiv:1707.08060



Analysis strategy

- **Count $pp \to e\nu\mu\nu$ (+ jets) events** in SR (+ high $p_T^{\text{lead.lep/jet}}$ regions)
  - Data binned in 12 observables*
- **Estimate backgrounds**
  - Dominant contribution from $t\bar{t}$ events → estimate with data-driven method
  - Fake leptons (data-driven), Z+jets, diboson, $V\gamma$
- **Unfolded result = detector^{-1} (data – backgrounds)**

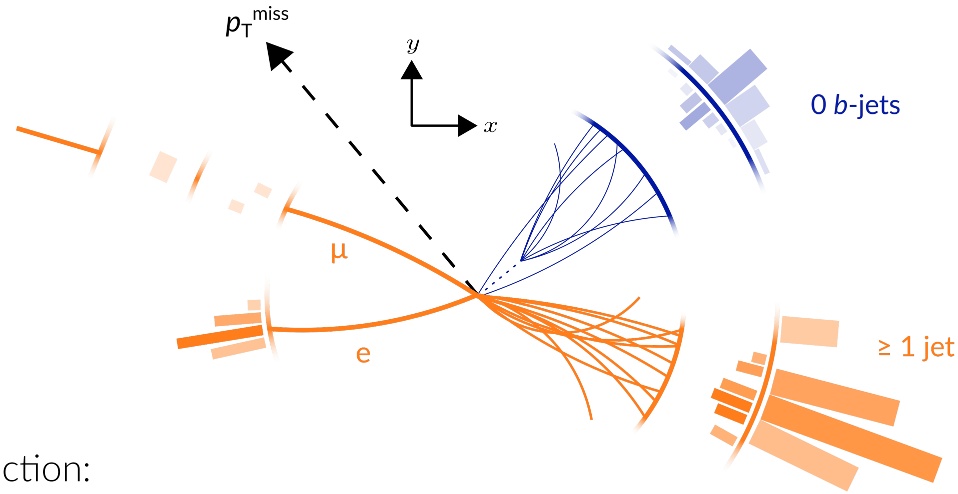
* $p_T^{\text{lead.lep.}}$, $p_T^{\text{sublead.lep.}}$, $m_{e\mu}$, $p_{T,e\mu}$, $y_{e\mu}$, $\Delta\varphi_{e\mu}$, $\cos\theta^*$
$p_T^{\text{lead.jet.}}$, $m_{T,e\mu}$, H_T, S_T, N_{jets},

Event selection

- Lepton and jet cuts:

p_T^e, p_T^μ > 27 GeV

≥ 1 jet
p_T > 30 GeV, $|\eta|$ < 4.5



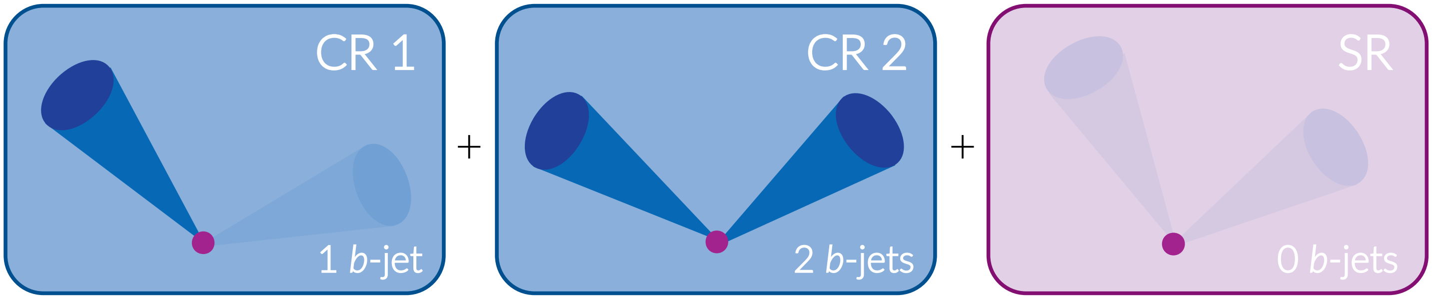
- Selections for background rejection:

different flavour leptons	*Reduces:* $Z/\gamma^* \to ee/\mu\mu$
$m_{e\mu}$ > 85 GeV	*Reduces:* $Z/\gamma^* \to \tau\bar{\tau}$ where $\tau \to e/\mu\ (+\nu\nu)$ resonant $gg \to H \to WW$ contributions
b-jet veto p_T > 20 GeV, $\lvert\eta\rvert$ < 4.5 DL1r 85% efficiency	*Reduces:* $t\bar{t}$ and Wt

Top estimate ($t\bar{t}$ + Wt)

- **Top events account for ~60% of events in signal region (SR)** → dominated by $t\bar{t}$ events
  - Use data-driven 'b-tag counting' method inspired by $t\bar{t}$ cross-section measurement → arXiv:1910.08819
  - Two control regions (CRs) with different numbers of *tagged* b-jets

$N_{t\bar{t}}$ =



$$N_{t\bar{t}} = N_{t\bar{t}}^{1b} + N_{t\bar{t}}^{2b} + N_{t\bar{t}}^{\mathrm{SR}}$$

$$\mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu} = \mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu}\cdot 2\epsilon_b\left(1 - C_b\epsilon_b\right) + \mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu}\cdot C_b\epsilon_b^2 + \mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu}\cdot\left(1 - 2\epsilon_b + C_b\epsilon_b^2\right)$$

$\mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu}$: $t\bar{t}$ count (up to b-jet selection)

$N_{t\bar{t}}^{1b}$ + $N_{t\bar{t}}^{2b}$: $t\bar{t}$ enhanced CRs → count events in data (subtract small backgrounds from MC)

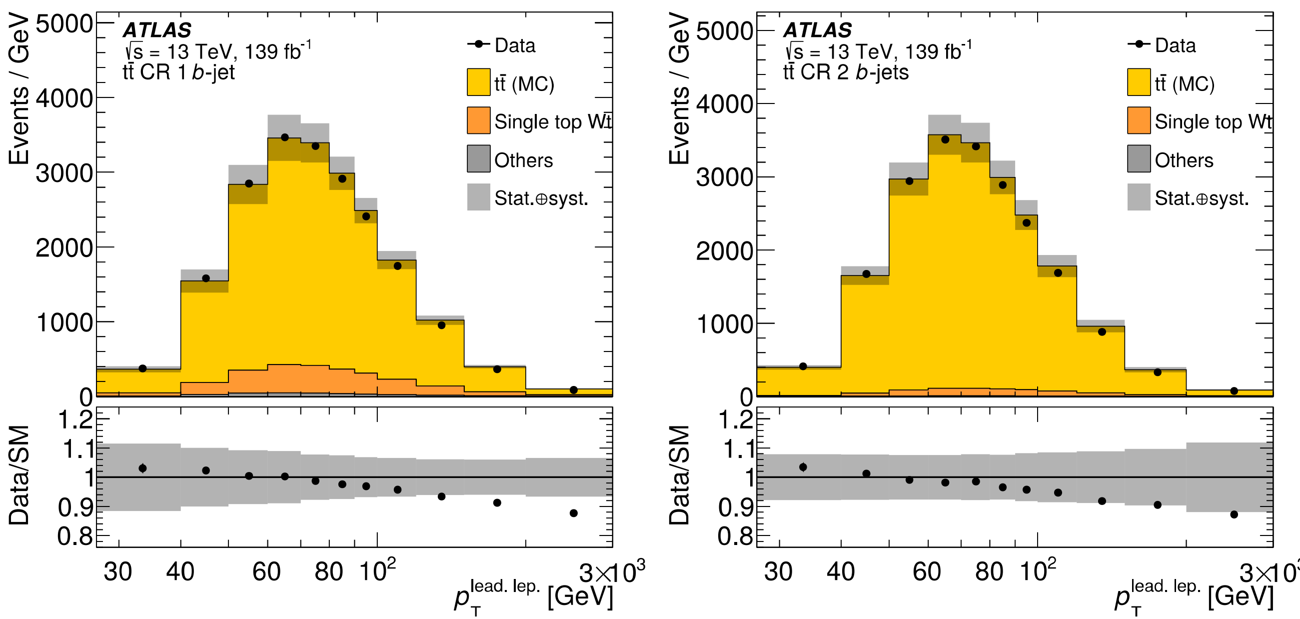
$N_{t\bar{t}}^{\mathrm{SR}}$: what we want (solve simultaneous equations)

- **$t\bar{t}$ modelling only enters in C_b**
- **Repeated in each bin for differential measurements**
  (b-jet p_T > 20 GeV so CR 2 also defined for 1 jet bin)

ϵ_b = b-jet selection efficiency
C_b = correlation factor (MC)

Top estimate ($t\bar{t}$ + Wt)

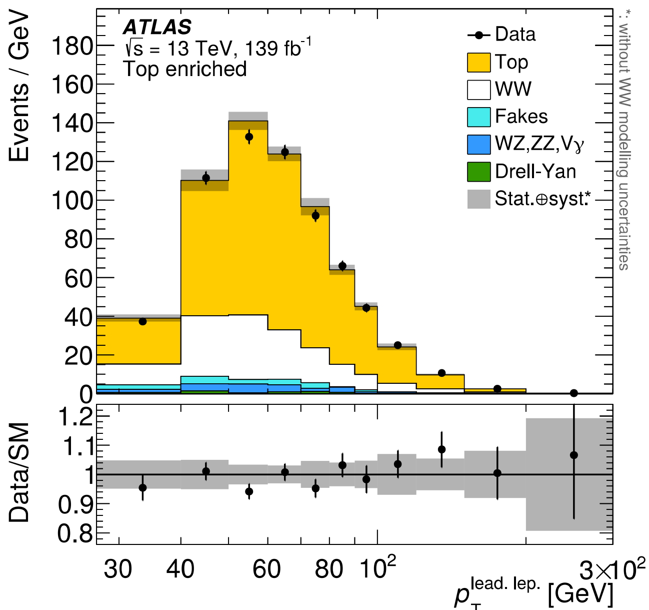
- **Detector-level distributions in two CRs**
  - Excess of events predicted at high $p_T^{\text{lead.lep.}}$ corrected for by data-driven $t\bar{t}$ estimate
- **Single top (Wt) contribution from MC**

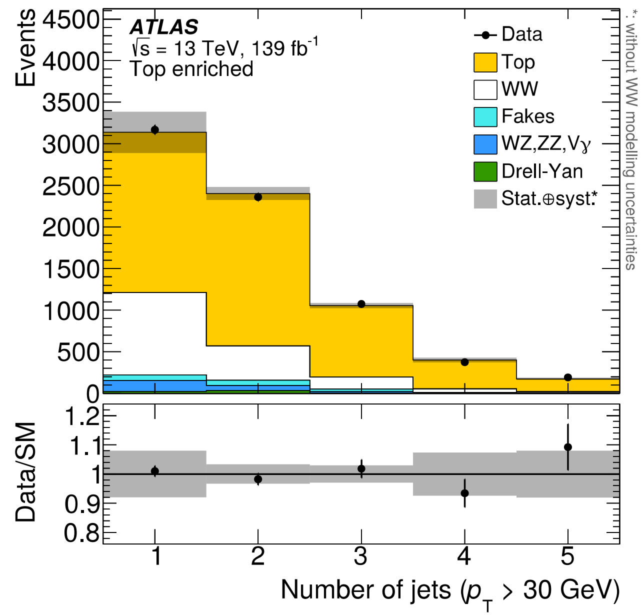


arXiv:2103.10319

Top estimate ($t\bar{t}$ + Wt) validation

- **Estimate strongly reduces systematic uncertainties:** 15% (pure MC) → 2.8%
  - Anti-correlation between some $t\bar{t}$ and Wt systematics reduces total uncertainty
- **Extensive closure tests performed**
- **Check estimate in top enriched validation region (VR):** m_{lj} < 140 GeV, $\Delta\varphi_{e\mu} < \pi/2$ (+ SR)





arXiv:2103.10319

Fake lepton estimate

- **Fake = jet misidentified as lepton / lepton from heavy flavour (HF) decay (3%)**, mainly *W*+jets
- **Poorly modelled** → estimate contribution with data-driven fake factor (FF) method
  - Use two auxiliary regions: *dijet* and *ID+anti-ID*

e/μ selection in SR = ID
orthogonal selection = anti-ID

Dijet region (lepton candidate balanced by jet)

Extract (p_T, η, flavour-dependent) **fake factors** here

lepton candidate selection

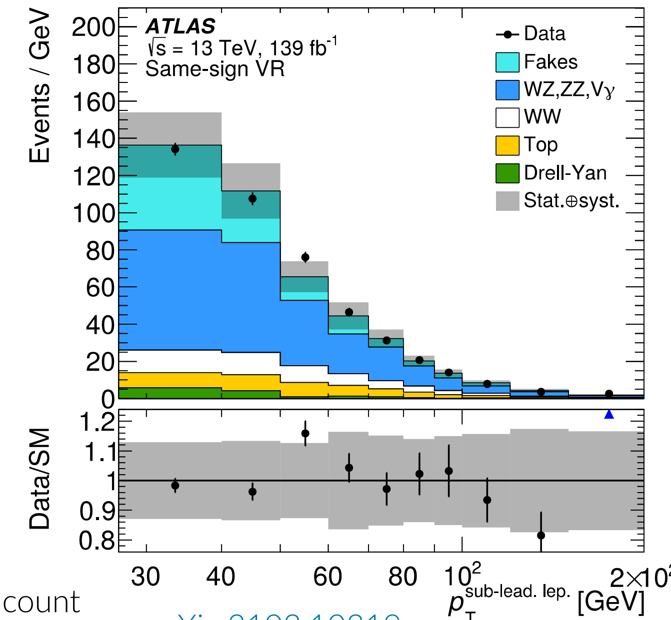
$$\mathrm{FF} = \frac{N_{\mathrm{ID}} - N^{\mathrm{prompt}}_{\mathrm{ID,MC}}}{N_{\mathrm{anti\text{-}ID}} - N^{\mathrm{prompt}}_{\mathrm{anti\text{-}ID,MC}}} = \frac{\mathrm{ID}}{\mathrm{anti\text{-}ID}}$$

ID+anti-ID region (SR selection with one ID → anti-ID)

Apply fake factors here

$$N^{\mathrm{non\text{-}prompt}}_{\mathrm{ID+ID}} = \mathrm{FF} \times \left(N_{\mathrm{ID+anti\text{-}ID}} - N^{\mathrm{prompt}}_{\mathrm{ID+anti\text{-}ID,MC}}\right)$$
$$= \frac{\mathrm{ID}}{\mathrm{anti\text{-}ID}} \times (\mathrm{ID+anti\text{-}ID}) = (\mathrm{ID+ID})$$

fake count in SR



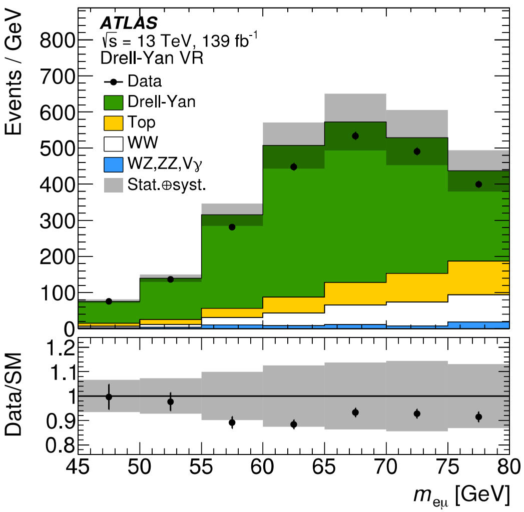
arXiv:2103.10319

- **FF estimate validated in same sign region**
- **Total uncertainty of 40%**

Other backgrounds

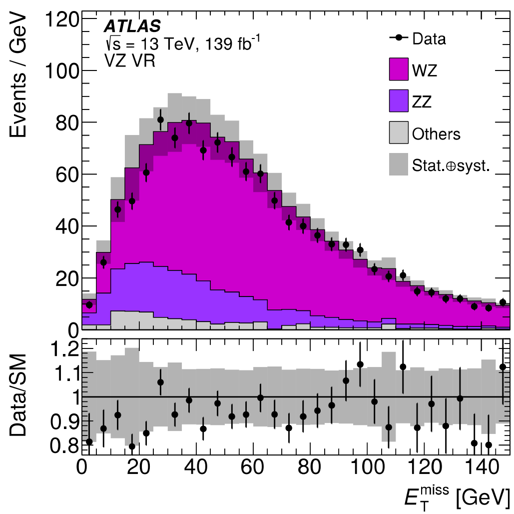
- **All remaining backgrounds estimated from simulation and validated in dedicated VRs**
  - Account for ~3% of events in SR → Sherpa 2.2.X
  - Z+jets (Drell-Yan), VZ, $V\gamma$ → triboson negligible (< 0.1% of selected events)
  - Systematic uncertainties from alternative signal models

$Z \to \tau\tau$



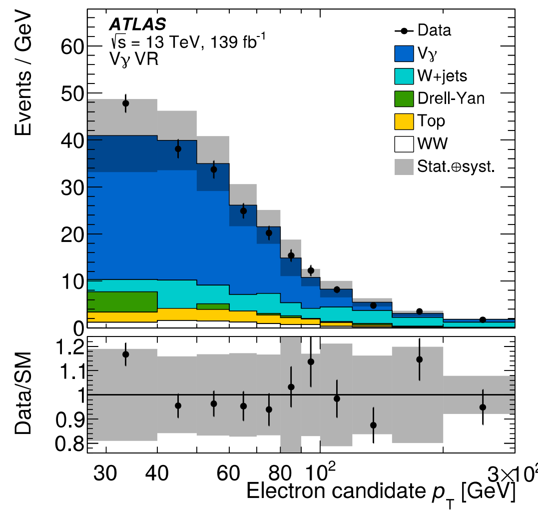
45 < $m_{e\mu}$ < 80 GeV
E_T^{miss} < 20 or $p_{T,e\mu}$ < 30 GeV

missed lepton



3 leptons
80 < m_{ll}^{SF} < 100 GeV

γ misidentified as e



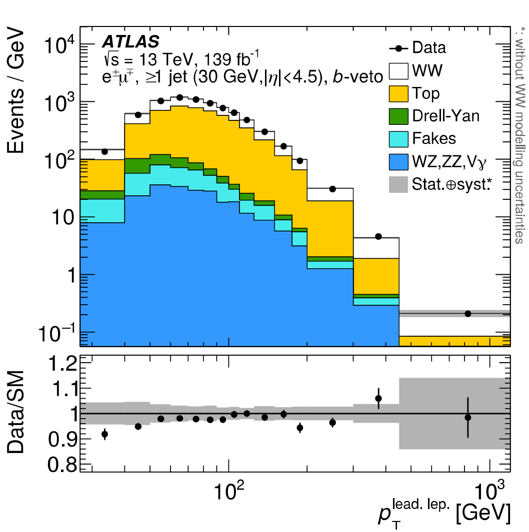
relaxed lepton criteria
same + opposite charge $e\mu$

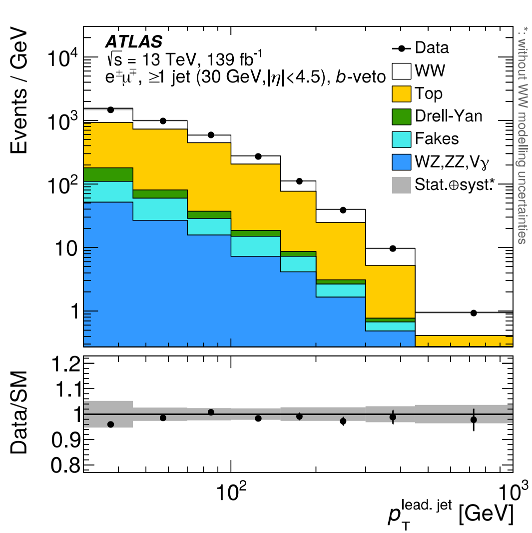
arXiv:2103.10319

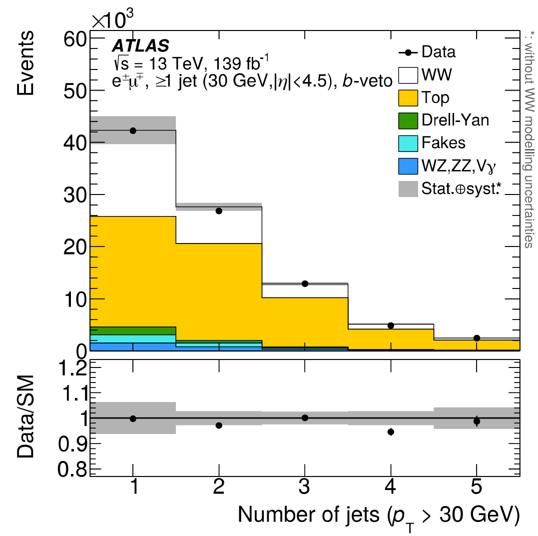
Results — detector-level distributions

- **Comparison of data with detector-level predictions**
  - Very good agreement seen
  - Slight over-prediction at low $p_T^{\text{lead.lep.}}$ with nominal Sherpa prediction (covered by signal modelling uncertainties)

	Signal region		$p_T^{\text{lead. jet}} > 200$ GeV	
Data	89 239		5825	
Total SM	91 600 ± 2500		5980 ± 150	
WW	28 100 ± 1200	31%	2480 ± 60	42%
Total bkg.	63 500 ± 1800	69%	3500 ± 140	58%
Top	55 800 ± 1500	61%	3030 ± 110	51%
Drell–Yan	2200 ± 700	2%	66 ± 9	1%
Fake leptons	2700 ± 1100	3%	140 ± 70	2%
$WZ, ZZ, V\gamma$	2800 ± 500	3%	270 ± 70	4%







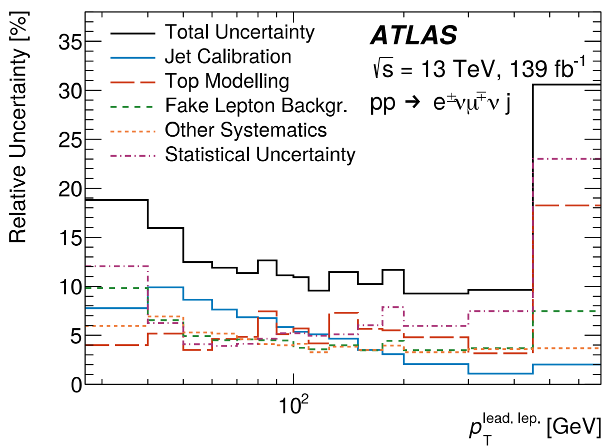
arXiv:2103.10319

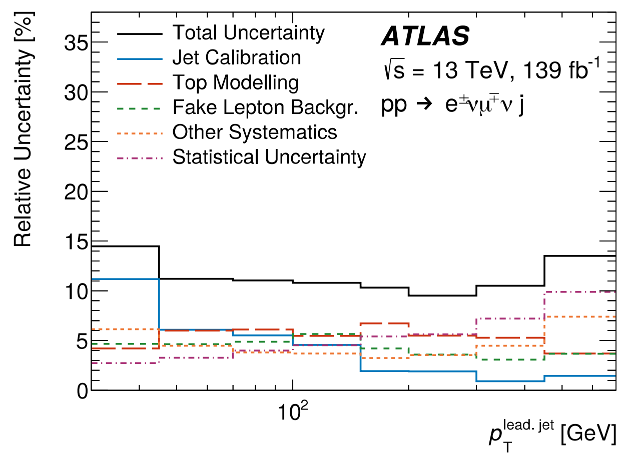
Results — fiducial-level cross sections

- **Fiducial phase space chosen as close to measurement phase space as possible**
  - No $\tau \to l\nu\nu$ on fiducial level (correction of ~9%)
  - Nominal fiducial region inclusive over jet flavour, but additional *b*-veto selection available on HEPData

Fiducial selection requirements		
p_T^{ℓ}	$>$	27 GeV
$\lvert\eta^{\ell}\rvert$	$<$	2.5
$m_{e\mu}$	$>$	85 GeV
p_T^{j}	$>$	30 GeV
$\lvert y^{j}\rvert$	$<$	4.5

- **Correct for detector effects** using iterative Bayesian unfolding method
  - Systematic uncertainties from varying unfolding inputs, statistical from toys → total uncertainty ~10% dominated by jet calibration, top and fake contributions





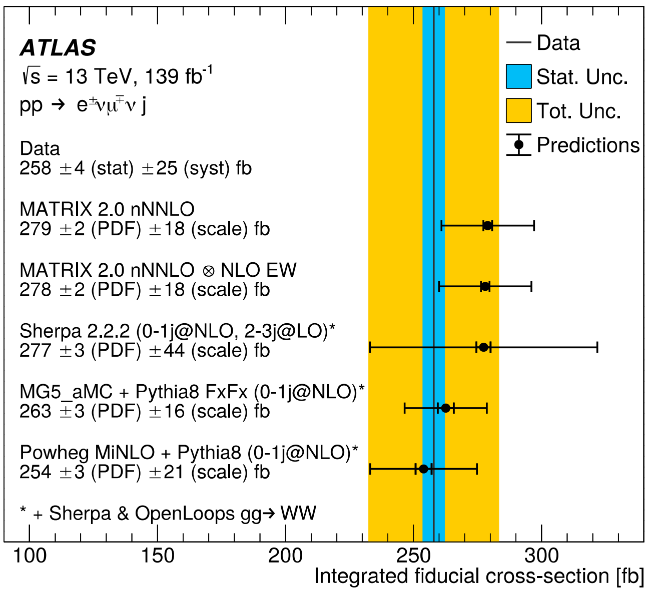
arXiv:2103.10319

fiducial-level cross sections

- Unfolded results compared to variety of fixed order and NLO+PS predictions

Process	Generator	Parton shower	PDF	Matrix element $\mathcal{O}(\alpha_S)$
$q\bar{q} \to WW$	MATRIX 2.0	–	NNPDF3.1	NNLO
$gg \to WW$	MATRIX 2.0	–	NNPDF3.1	NLO
$q\bar{q} \to WW$	Sherpa 2.2.2	Sherpa	NNPDF3.0	NLO (0–1 jet), LO (2–3 jets)
$q\bar{q} \to WW$	Powheg MiNLO	Pythia 8	NNPDF3.0	NLO (0–1 jet)
$q\bar{q} \to WW$	MadGraph 2.3.3	Pythia 8	NNPDF3.0	NLO (0–1 jet)
$gg \to WW$	Sherpa 2.2.2 + OpenLoops	Sherpa	NNPDF3.0	LO (0–1 jet)

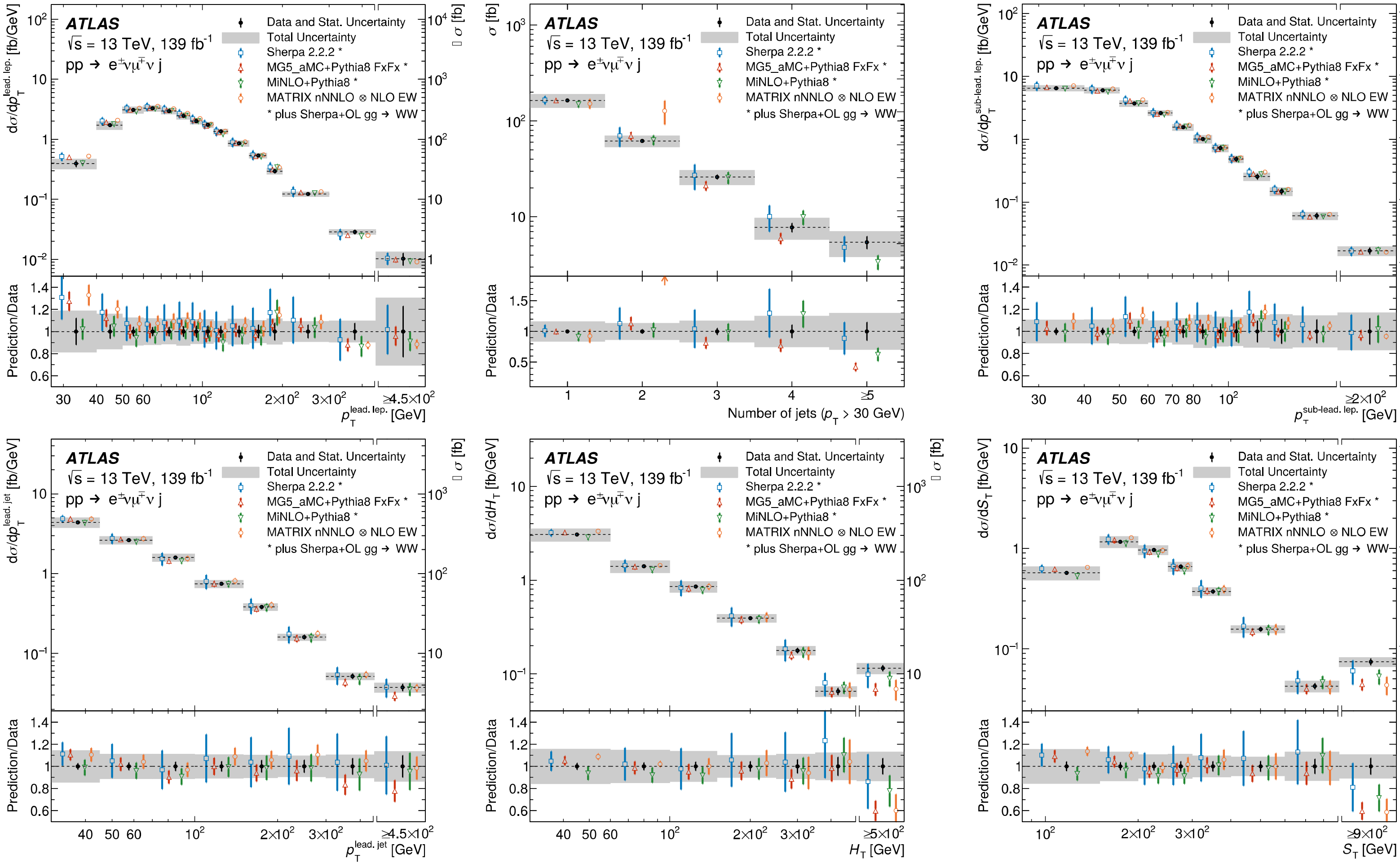
+ EW corrections for $qq \to WWj$ (MATRIX 2.0 rows)



- **Integrated fiducial cross section shows excellent agreement with theoretical predictions**
  - Relatively large scale uncertainty for Sherpa due to LO matrix elements with up to 3 jets
- **Differential distributions** (following slides) in general **show very good agreement**
  - χ^2/n.d.f. values for nominal Sherpa 2.2.2 prediction all <1 (excluding $m_{e\mu}$ in high jet p_T region = 1.4) → similar for other predictions

arXiv:2103.10319

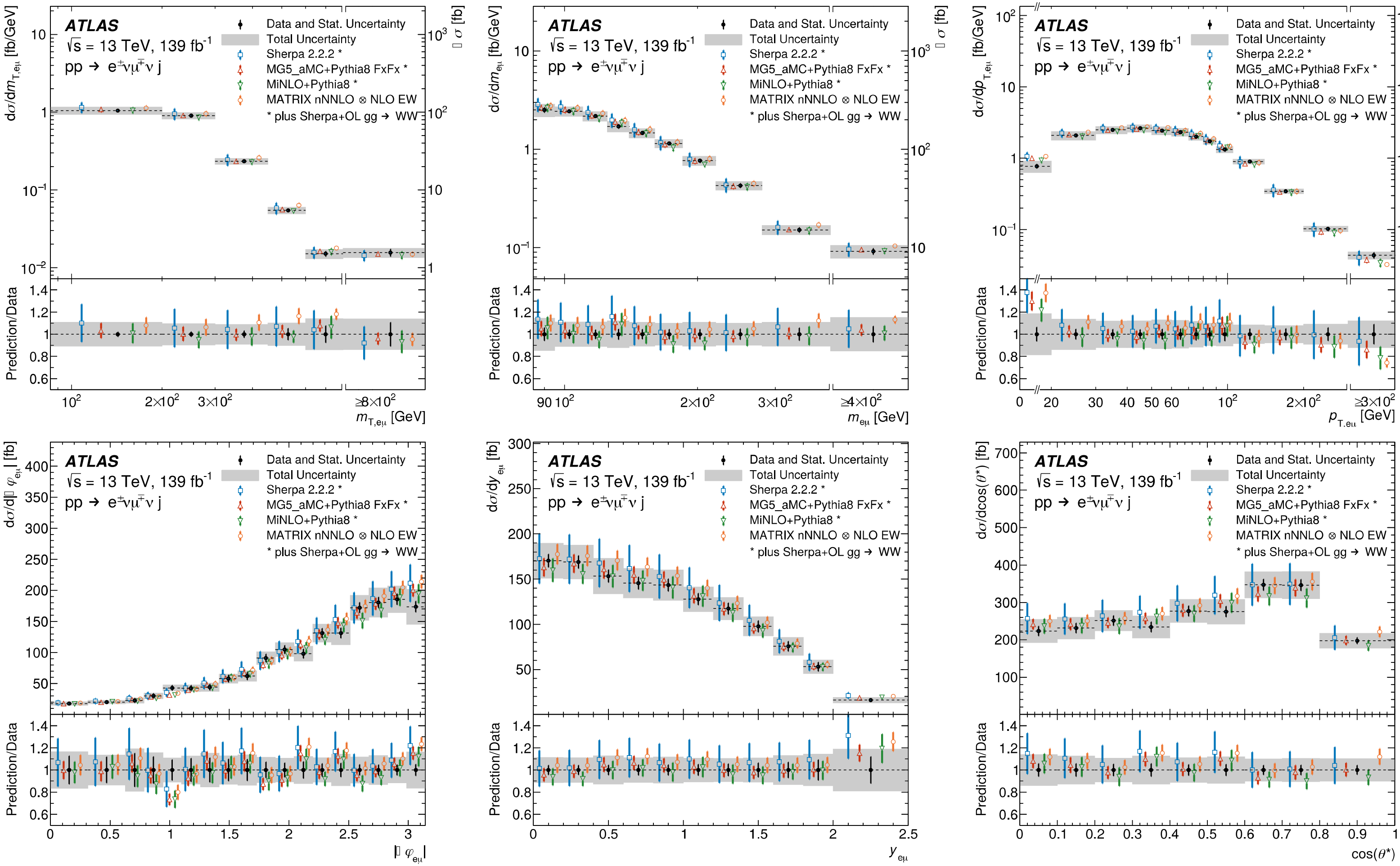
Results fiducial-level cross sections



arXiv:2103.10319

Results

fiducial-level cross sections



EFT interpretation

- **SM can be considered as EFT** with additional dim. > 4 operators suppressed by some high energy scale Λ

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_{i,d>4} \frac{c_i^{(d)}}{\Lambda^{d-4}} Q_i^{(d)}$$

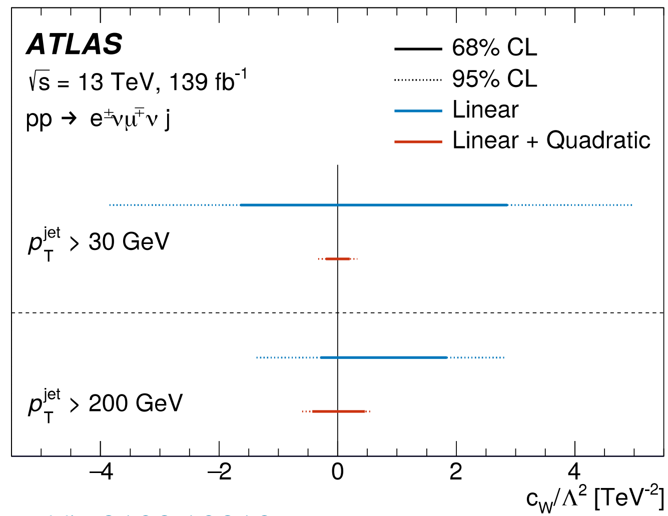
- **Small scale EFT study**: focus on one dim. 6 operator
  - Analysis sensitive to Q_W affecting gauge boson self-couplings ⟶ arXiv:1008.4884
  - Importance of SM+BSM interference term in cross section expected to increase with jet p_T ⟶ arXiv:1707.08060

$$\sigma = \sigma_{\rm SM} + \frac{c_W}{\Lambda^2}\sigma_{\rm int} + \frac{c_W^2}{\Lambda^4}\sigma_{\rm BSM}$$

purely SM | SM+BSM | purely BSM

Fitting and results

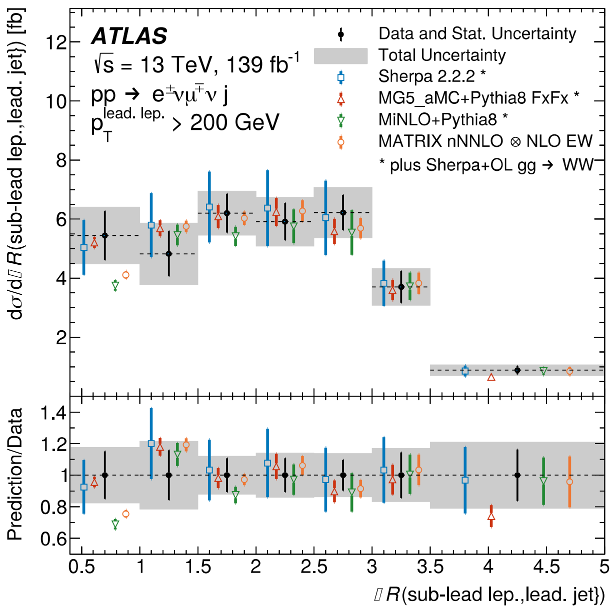
- **Perform likelihood fits in dedicated $p_T^{\rm lead.jet}$ > 200 GeV region** using (unfolded) $m_{e\mu}$ distribution
- **Impact of quadratic term seen to reduce** compared to $p_T^{\rm lead.jet}$ > 30 GeV (nominal SR) fit

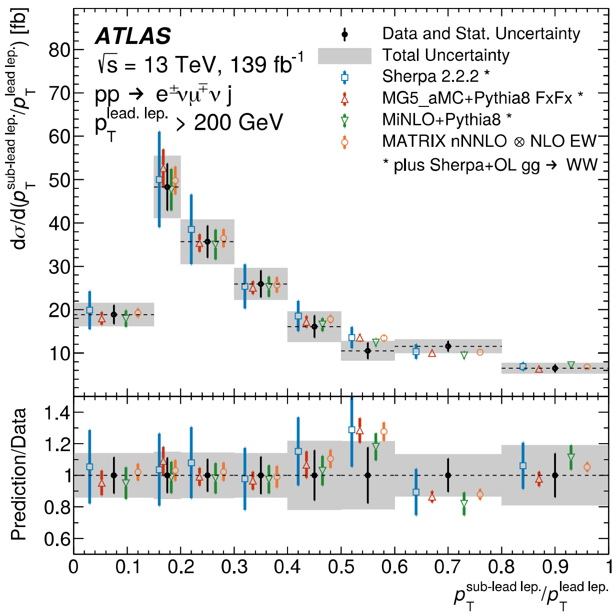


arXiv:2103.10319

High lepton p_T region

- **Additional differential cross sections in $p_T^{\text{lead.lep.}} > 200$ GeV region**
  - Targets *W*+jets event topologies with a soft *W* emission from a jet
  - Affected by ‘giant *K*-factors’ corresponding to large higher order EW and QCD corrections
- **Unfolded distributions show good agreement with theoretical predictions**





	$p_T^{\text{lead. lep.}} > 200$ GeV	
Data	3873	
Total SM	3960 ± 120	
WW	1740 ± 50	44%
Total bkg.	2210 ± 110	56%
Top	1920 ± 90	49%
Drell–Yan	42 ± 6	1%
Fake leptons	70 ± 40	2%
$WZ, ZZ, V\gamma$	180 ± 40	4%

arXiv:2103.10319

Summary

- **First *WW* jet-inclusive differential measurements performed at LHC**
- **Reduction of uncertainties in dominant top background** using powerful data-driven $t\bar{t}$ estimate
- **Fiducial and differential cross sections agree with theoretical predictions** up to highest measured p_{T} and for up to 5 jets
- **Improved sensitivity to EFT interference term** in high jet p_{T} region

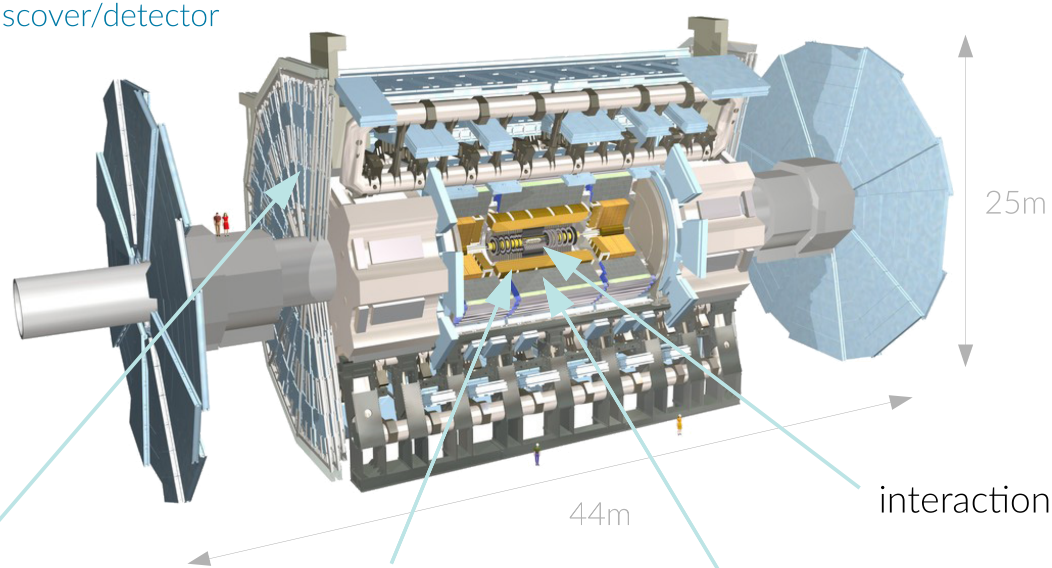
Analysis team proceeding to look at *WW* + 0 jets

In near future combine results for high precision fully inclusive measurement

Backup

ATLAS detector and reconstruction

https://atlas.cern/discover/detector



Muon spectrometer **Tracker (ID)** **EM** and **hadronic calorimeters**

Muons = tracks in tracker and muon spectrometer
$|\eta| < 2.5$

Electrons = calorimeter deposits and tracks
$|\eta| < 2.47$ (excl. $1.37 < |\eta| < 1.52$)

Jets = (anti-k_T) clustered calorimeter deposits and tracks
$|\eta| < 4.5$

Secondary vertices from *b*-hadron decays allow for *b*-jet identification (DL1r)

Detailed event selection

Selection	Criteria
Lepton p_T	> 27 GeV
Lepton η	$\lvert\eta\rvert < 2.47$ and not $1.37 < \lvert\eta\rvert < 1.52$ (electron)
	$\lvert\eta\rvert < 2.5$ (muon)
Lepton identification	`TightLH` (electron), `Medium` (muon)
Lepton isolation	`Gradient` (electron), `Tight_FixedRad` (muon)
Lepton impact parameter	$\lvert d_0/\sigma_{d_0}\rvert < 5, 3$ (electron, muon)
	$\lvert z_0 \cdot \sin\theta\rvert < 0.5$ mm
Jet selection	$p_T > 30$ GeV, $\lvert\eta\rvert < 4.5$
b-jet selection	$p_T > 20$ GeV, $\lvert\eta\rvert < 2.5$, DL1r (85% eff. WP)
Lepton selection	1 electron and 1 muon of opposite charge,
	no additional lepton with $p_T > 10$ GeV, `Loose` isolation,
	and `LooseLH` (electron) / `Loose` (muon) identification
Number of jets	≥ 1
Number of b-jets	0
$m_{e\mu}$	> 85 GeV
High $p_T^{\text{lead. jet}}$ selection	$p_T^{\text{lead. jet}} > 200$ GeV

arXiv:2103.10319

MC samples

Process	Generator	Parton shower	Matrix element $\mathcal{O}(\alpha_S)$	Normalization
$q\bar{q} \to WW$	SHERPA 2.2.2	SHERPA	NLO (0–1 jet), LO (2–3 jets)	Generator†
$gg \to WW$	SHERPA 2.2.2	SHERPA	LO (0–1 jet)	Generator
$t\bar{t}$	POWHEG-BOX v2	PYTHIA 8	NLO	NNLO+NNLL
Wt	POWHEG-BOX v2	PYTHIA 8	NLO	NLO+NNLL
Z+jets	SHERPA 2.2.1	SHERPA	NLO (0–2 jets), LO (3–4 jets)	NNLO
WZ, ZZ	SHERPA 2.2.2	SHERPA	NLO (0–1 jet), LO (2–3 jets)	Generator†
$W\gamma$, $Z\gamma$	SHERPA 2.2.8	SHERPA	NLO (0–1 jet), LO (2–3 jets)	Generator†
VVV	SHERPA 2.2.2	SHERPA	NLO (0–1 jet), LO (2–3 jets)	Generator†

†: The cross-section calculated by SHERPA is found to be in good agreement with the NNLO result .

arXiv:2103.10319

Top estimate ($t\bar{t}$ + Wt) calculation details

- **Number of $t\bar{t}$ events passing $e\mu$ selection:**

$$N_{t\bar{t}} = \mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu}$$

$e\mu$ selection efficiency

$t\bar{t}$ (+jets) cross-section

- **Number of $t\bar{t}$ events in CRs obtained from data** (backgrounds estimated with MC)

$$N_{t\bar{t}}^{2b} = \mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu} \cdot C_b\epsilon_b^2$$
$$N_{t\bar{t}}^{1b} = \mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu} \cdot 2\epsilon_b\left(1 - C_b\epsilon_b\right)$$

ϵ_b = efficiency to find and tag a b-jet
ϵ_{bb} = efficiency to find and tag two b-jets
$C_b = \epsilon_{bb}/\epsilon_b^2$ = correlation factor

- **Solve for** $\epsilon_b = \dfrac{2N_{t\bar{t}}^{2b}}{C_b\left(N_{t\bar{t}}^{1b} + 2N_{t\bar{t}}^{2b}\right)}$ $\quad \mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu} = \dfrac{C_b}{4}\dfrac{\left(N_{t\bar{t}}^{1b} + 2N_{t\bar{t}}^{2b}\right)^2}{N_{t\bar{t}}^{2b}}$

- **Obtain estimate in SR**

$$N_{t\bar{t}}^{SR} = N_{t\bar{t}} - N_{t\bar{t}}^{1b} - N_{t\bar{t}}^{2b}$$
$$= \mathcal{L}\sigma_{t\bar{t}}\epsilon_{e\mu} \cdot \left(1 - 2\epsilon_b + C_b\epsilon_b^2\right)$$

$$N_{t\bar{t}}^{SR} = \frac{C_b}{4}\frac{\left(N_{t\bar{t}}^{1b} + 2N_{t\bar{t}}^{2b}\right)^2}{N_{t\bar{t}}^{2b}} - N_{t\bar{t}}^{1b} - N_{t\bar{t}}^{2b}$$

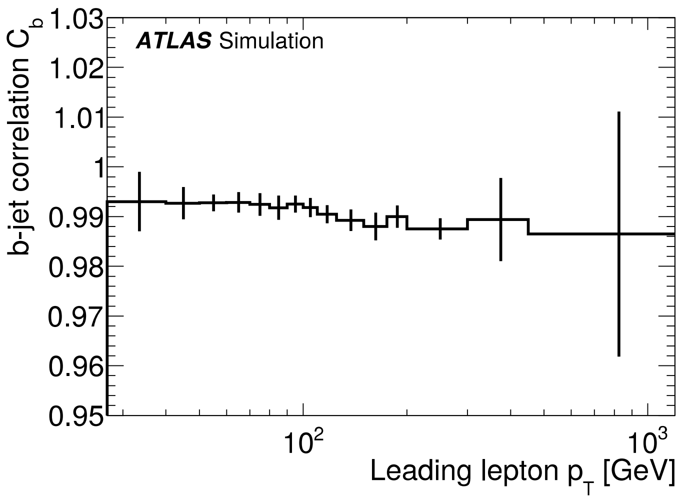
$$N_{t\bar{t}}^{ib} = N_{\text{data}}^{ib} - N_{\text{bkg,MC}}^{ib}$$

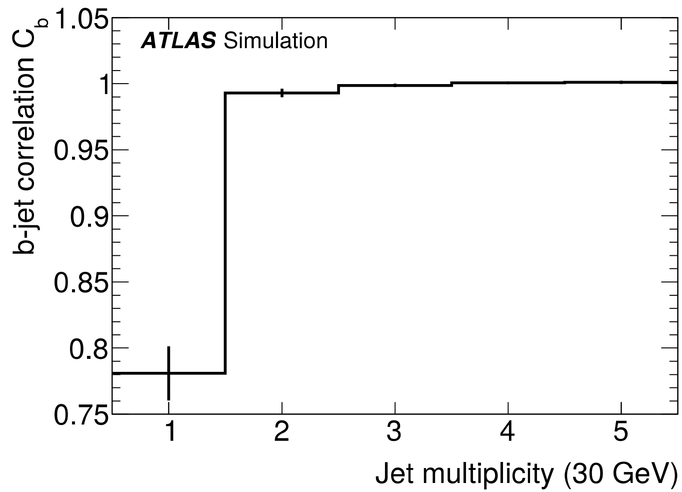
$$C_b = \frac{4N_{t\bar{t},\text{MC}}^{0+1+2b}N_{t\bar{t},\text{MC}}^{2b}}{\left(N_{t\bar{t},\text{MC}}^{1b} + 2N_{t\bar{t},\text{MC}}^{2b}\right)^2}$$

only use of $t\bar{t}$ modelling

Top estimate ($t\bar{t}$ + Wt) — b-jet correlation factor, C_b

- **Inclusive value:** $C_b = 0.991 \pm 0.002$
- **Example differential distributions:**





arXiv:2103.10319

Background estimates

Region	Observed	Predicted ± Error	Purity
$t\bar{t}$ CR 1b	260 971	268 000 ± 19 000	87%
$t\bar{t}$ CR 2b	257 777	267 000 ± 21 000	96%
Top enriched	7167	7000 ± 1000	72%
Same-sign VR	5095	5000 ± 600	25%
Drell–Yan VR	11 824	13 000 ± 1600	74%
VZ VR	14 770	14 000 ± 1900	94%
$V\gamma$ VR (OS)	2720	2670 ± 240	63%
$V\gamma$ VR (SS)	2401	2250 ± 240	76%

arXiv:2103.10319

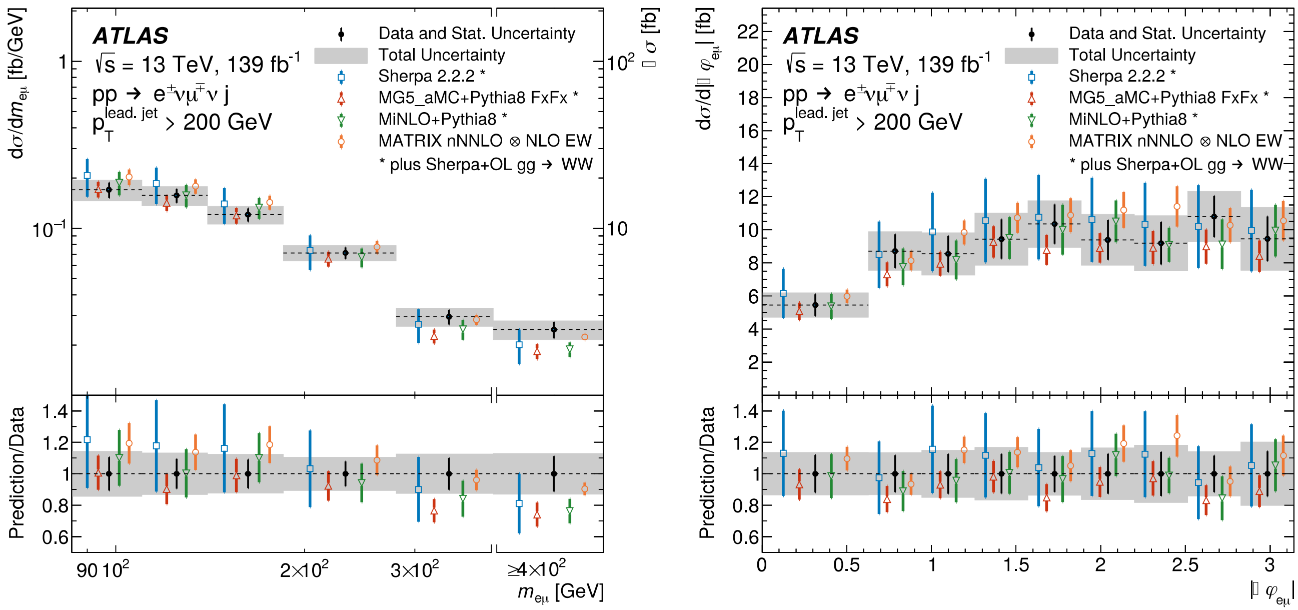
Results — uncertainty breakdown

Uncertainty source	Relative effect
Total uncertainty	10%
Signal region statistical uncertainty	1.1%
Data-driven background and MC statistics	1.2%
Jet calibration	6.3%
Top modelling	4.5%
Fake-lepton background	4.3%
Signal modelling	2.7%
Other background	2.3%
Flavour tagging	2.3%
Luminosity	1.9%
Other systematic uncertainties	0.6%

arXiv:2103.10319

Results

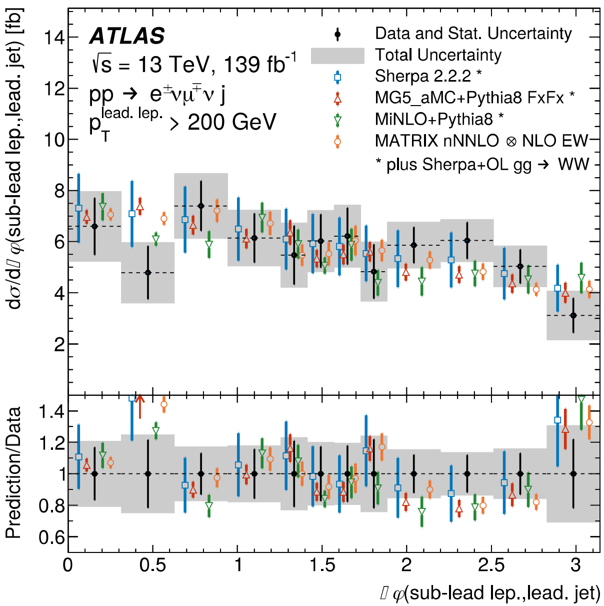
high jet p_T region

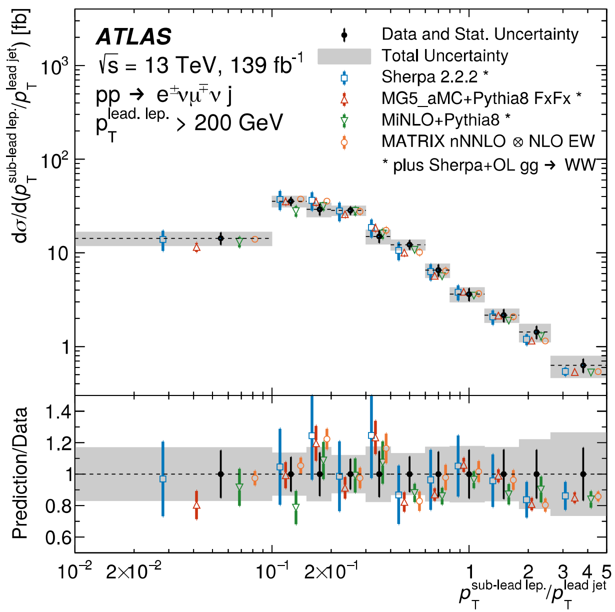


arXiv:2103.10319

Results

high lepton p_T region





arXiv:2103.10319