Opposite-sign WW production with Run3 data in ATLAS

ALBERTO PLEBANI

IOP Joint APP and HEPP Annual Conference 2026

Edinburgh – 8th April 2026

Supervisor: Dr. Sarah Louise Williams

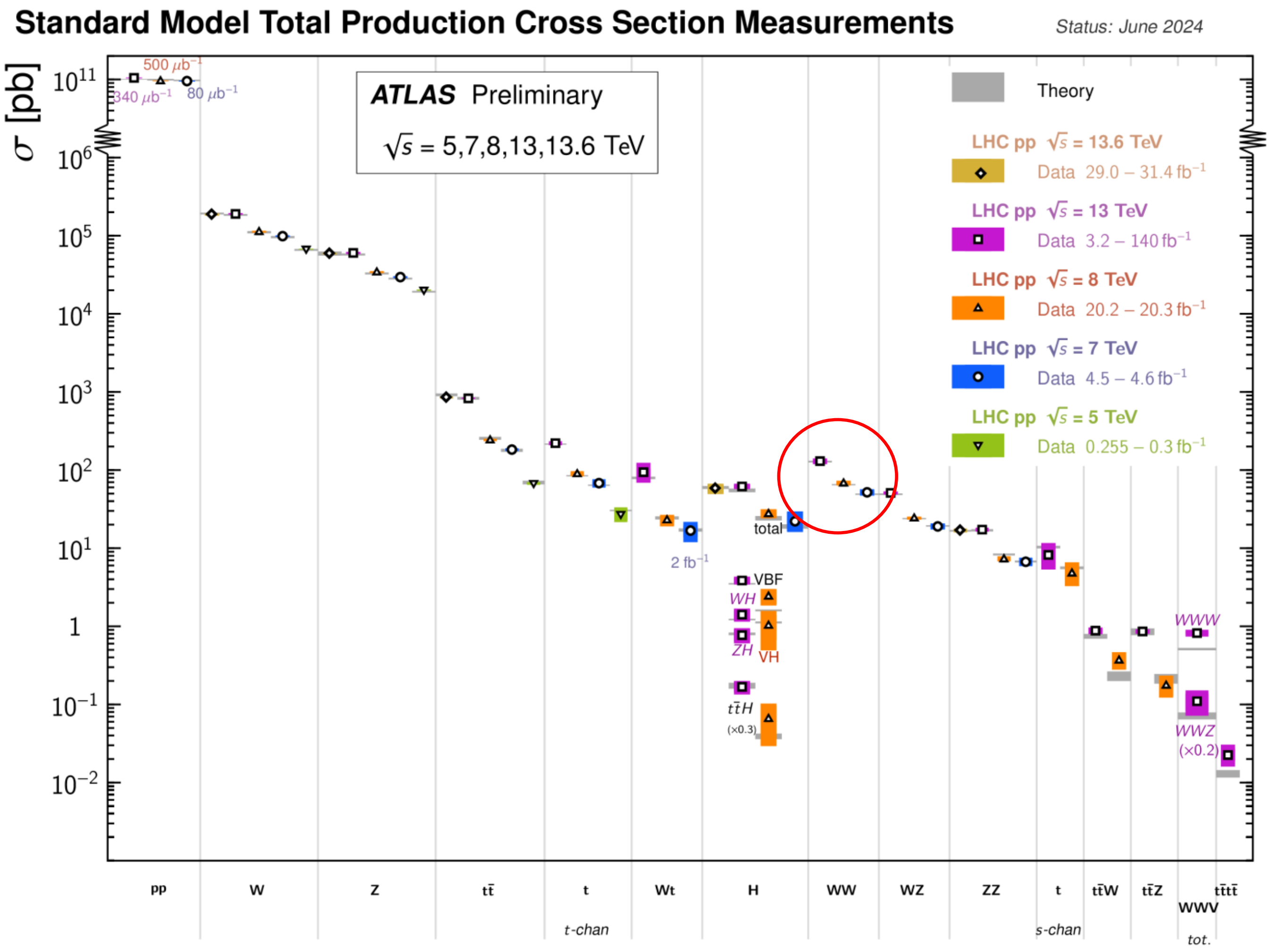
University of Cambridge, St Edmund's College

HARDING Distinguished Postgraduate SCHOLARS PROGRAMME

ATLAS EXPERIMENT

Standard Model Physics at the LHC

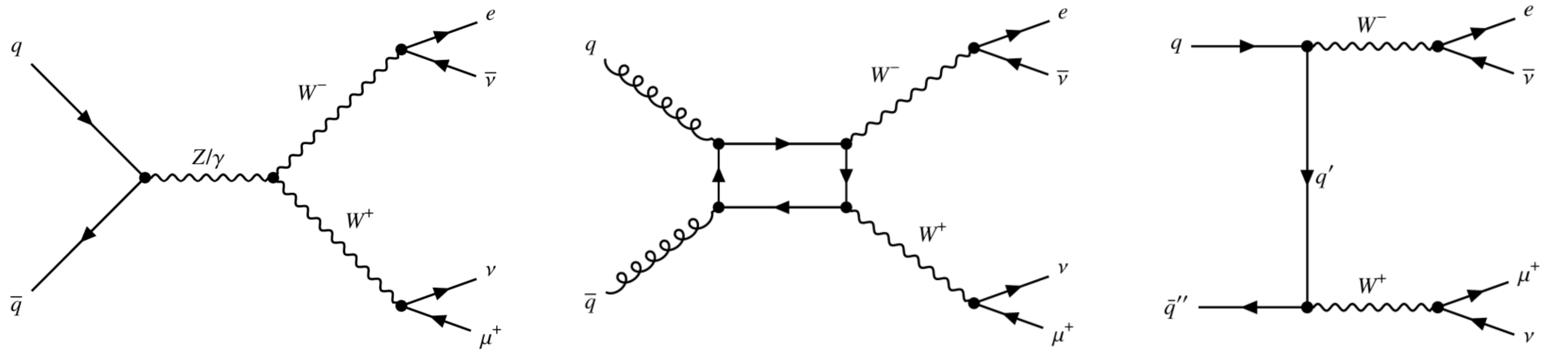
- Standard Model tested across 7 orders of magnitude with good agreement with predictions
- WW process:
  - Evidence in Run1 at $\sqrt{s} = 8$ TeV
  - Measurement in 2015 at $\sqrt{s} = 8$ TeV
  - Differential measurement at $\sqrt{s}$ =13 TeV with partial Run2 (36.1 fb^{-1})
  - Differential, EFT and asymmetry at $\sqrt{s} = 13$ TeV for full Run2 (140 fb^{-1})
- This analysis (team from Cambridge and Manchester)
  - First partial Run3 WW measurement at $\sqrt{s} = 13.6$ TeV (163.9 fb^{-1})
  - Differential measurement with focus on CP-Violation (CPV) and WW polarisation observables
  - Will also perform EFT interpretation



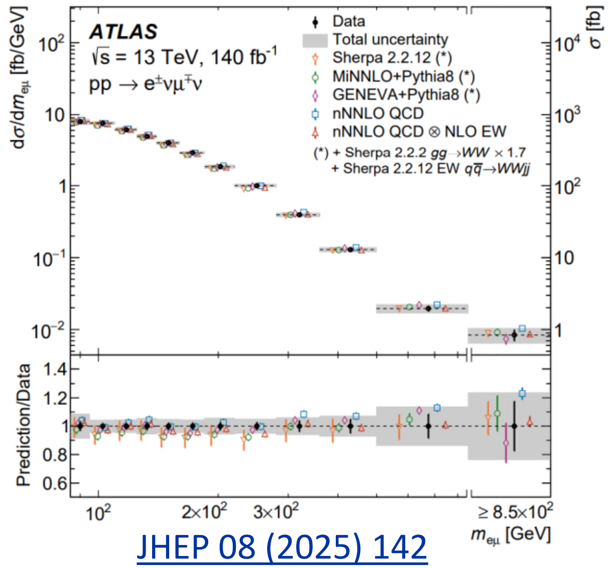
ATL-PHYS-PUB-2024-011

WW production

- Opposite-sign opposite-flavour WW final state

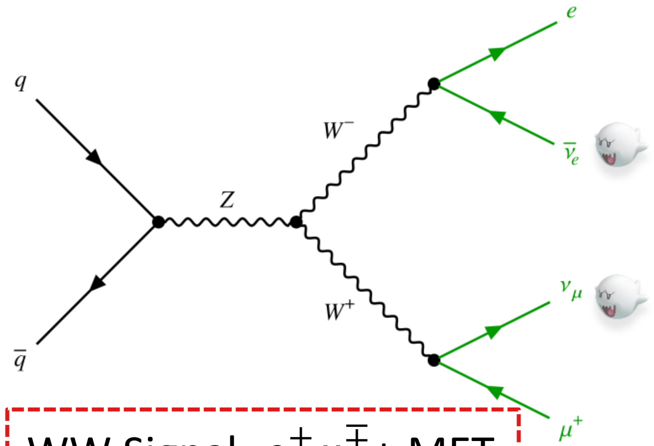


- Predicted cross section (Matrix 2.1 NNLO x NLO EW): $\sigma(W^+W^-) = 123 \pm 1$ (PDF) ± 2 (scale) fb
- Full Run2 analysis ($\sqrt{s}$ =13 TeV, 140 fb^{-1}):
  - Inclusive cross-section $\sigma\left(W^+W^- \to e^\pm\nu\,\mu^\mp\nu\right) = 127 \pm 1\,(stat.) \pm 4\,(syst.)$ fb
  - Differential cross section measured as a function of kinematic variables
- Partial Run3 analysis ($\sqrt{s}$ =13.6 TeV, 163.9 fb^{-1} :
  - Goal is to unfold variables sensitive to CPV and= WW polarisation
  - First look at CP-odd observables (signed $\Delta\phi(\ell\ell)$)
  - Enhance sensitivity to linear EFT over previous measurements
    - In particular Triple Gauge Coupling

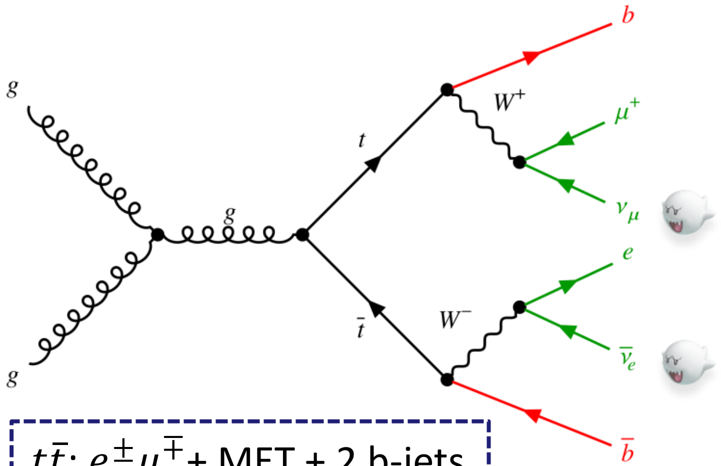


JHEP 08 (2025) 142

Event selection

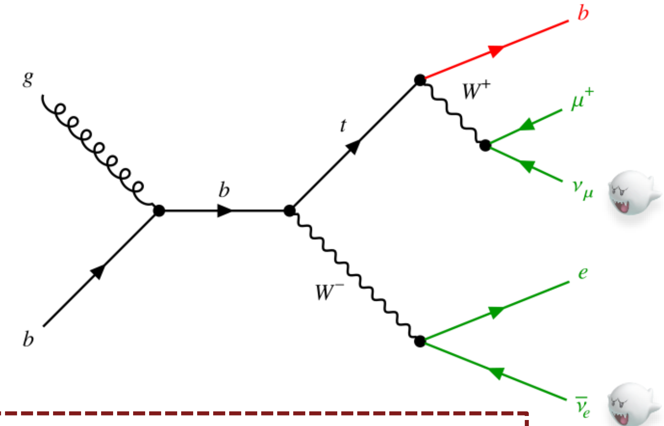


WW Signal: $e^{\pm}\mu^{\mp}$+ MET

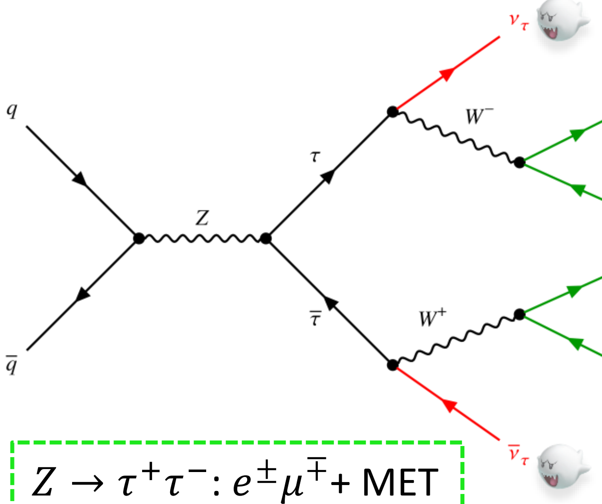


$t\bar{t}$: $e^{\pm}\mu^{\mp}$+ MET + 2 b-jets
=> b-jet veto

Sample	Yields ± Stat.	Percentage (%)
WW	57 485.7 ± 77.2	34.0
$Z + jets$	3636.5 ± 324.4	2.2
WZ	2366.5 ± 9.2	1.4
VV	130.6 ± 1.1	0.1
$t\bar{t}$	77 451.1 ± 762.8	45.8
$V\gamma$	174.1 ± 25.5	0.1
$W + jets$	100.7 ± 64.6	0.1
VVV	201.1 ± 0.5	0.1
VV EWK	768.0 ± 9.0	0.5
Single-top	19 748.4 ± 28.8	11.7
Fakes	6877.6 ± 72.4	4.1
Total mc23ade	168 940.4 ± 839.2	1
Run 3 Data	170 773.0 ± 413.2	-
Data/MC	1.01	-

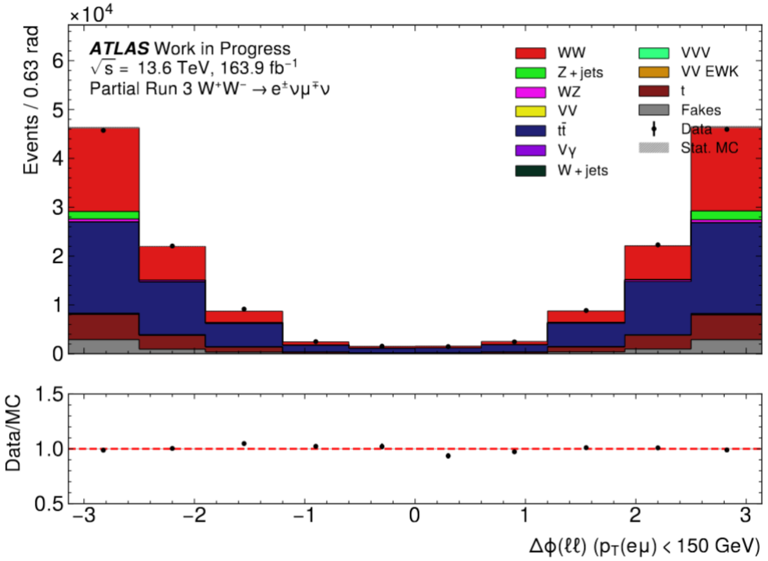


Single top: $e^{\pm}\mu^{\mp}$+ MET + 1 b-jet
=> b-jet veto



$Z \rightarrow \tau^{+}\tau^{-}$: $e^{\pm}\mu^{\mp}$+ MET
-> mll > 100 GeV

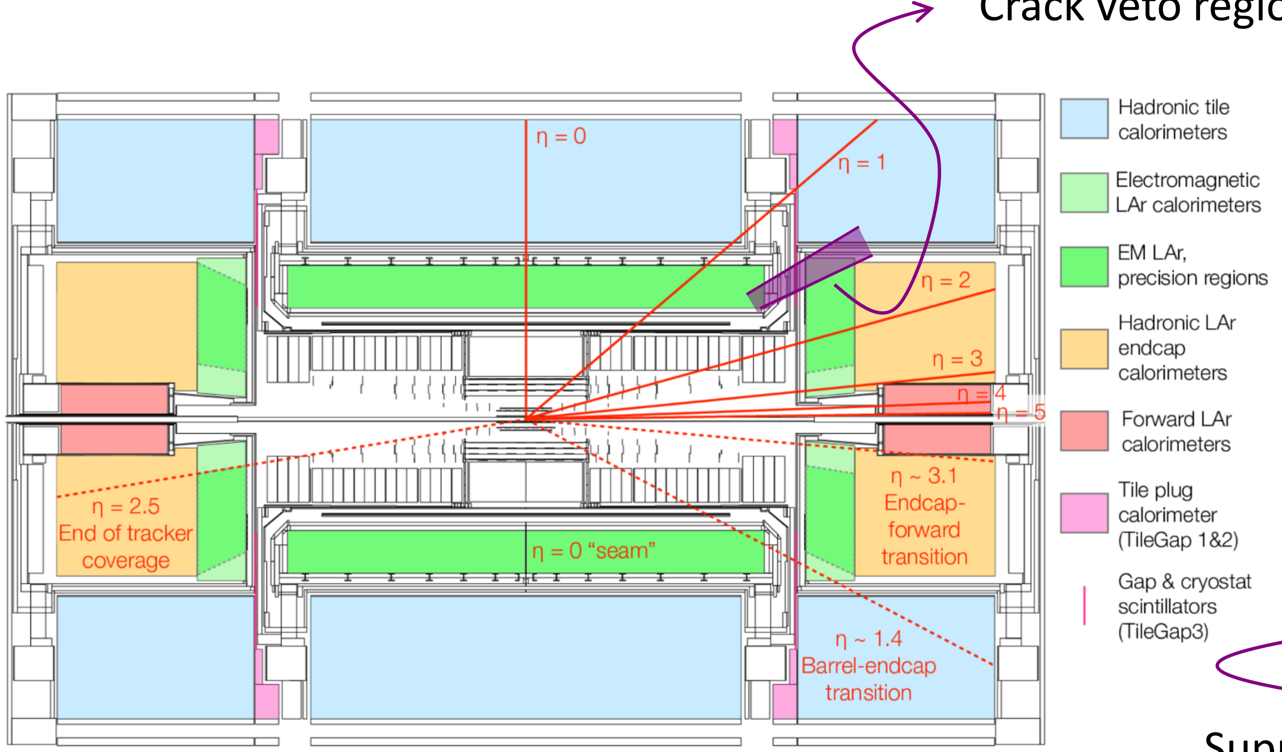
Good data/MC agreement

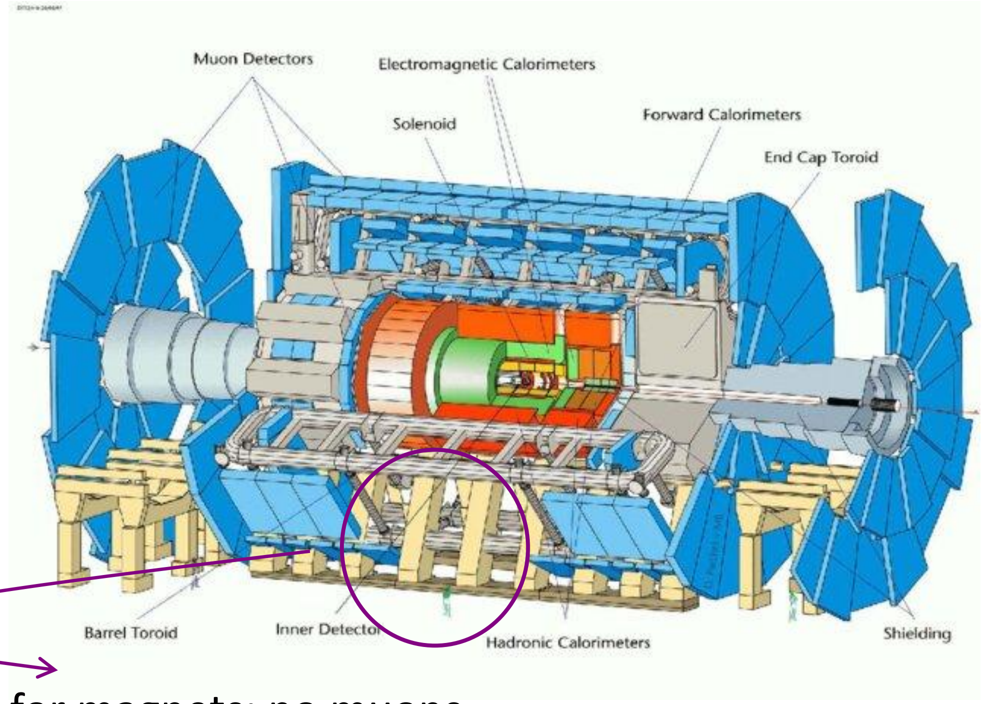


ATLAS detector effects

- ATLAS has a finite detector acceptance and resolution
- These effects lead to mismeasurements of the final state objects, which make it harder to compare with the truth-level prediction

Crack veto region: no electrons

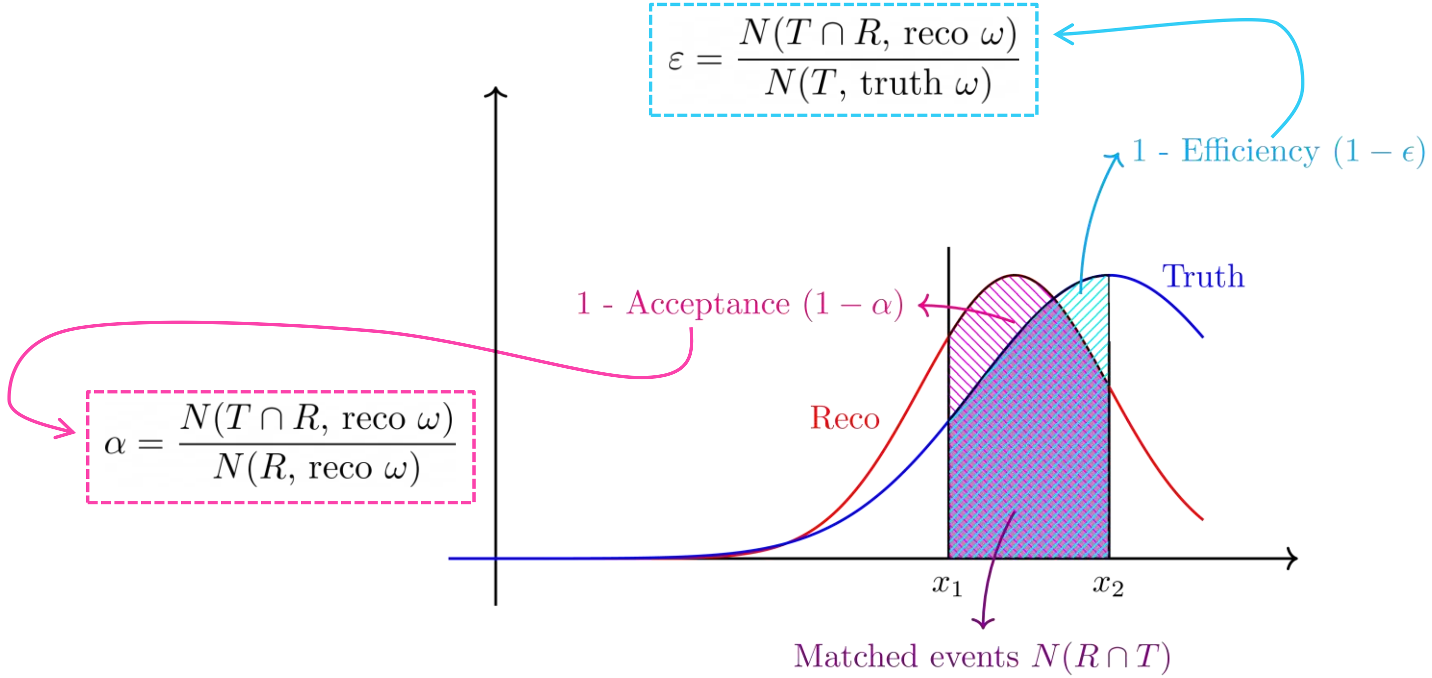


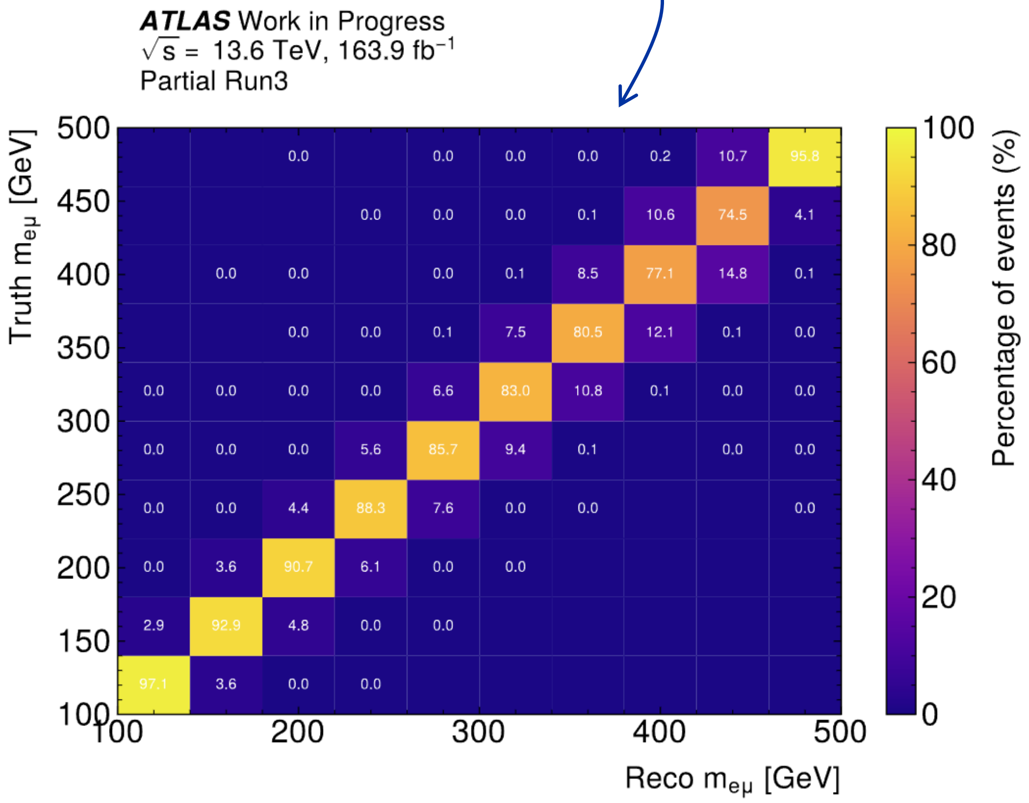


Support for magnets: no muons

Unfolding

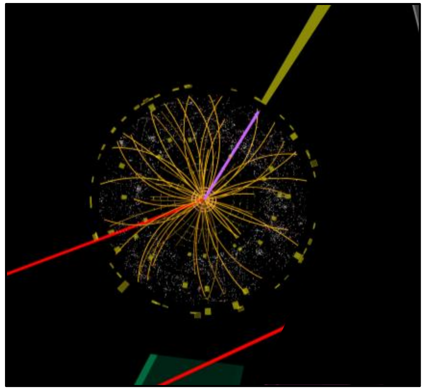
- Unfolding: correcting the measured distribution to account for the detector's effects
- The unfolding procedure can be parametrised as follows:
  - Response matrix: probability of measuring an event in reco given the event is measured in truth
  - Migration matrix: probability of observing both reco and truth
  - Reconstruction efficiency: probability of measuring reco
  - Fiducial acceptance : probability of measuring truth



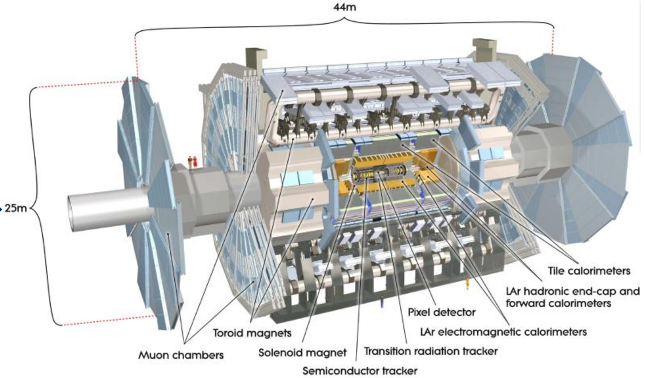


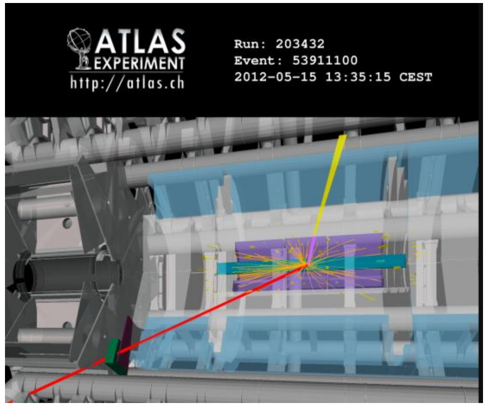
Reco-level data analysis workflow

Detector interaction

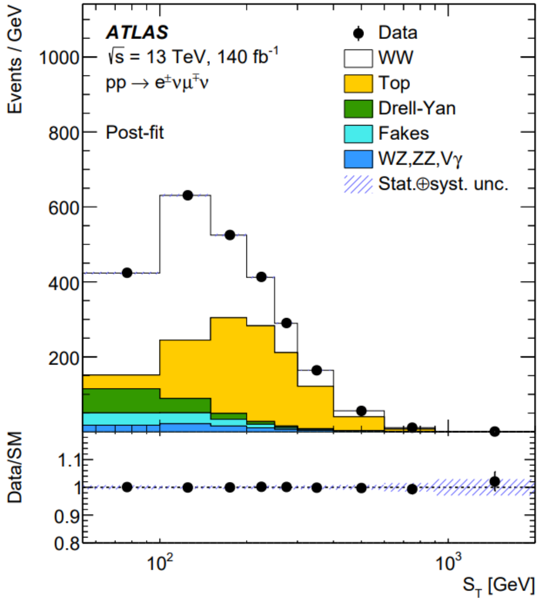


Collision

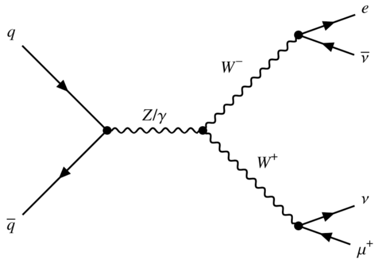




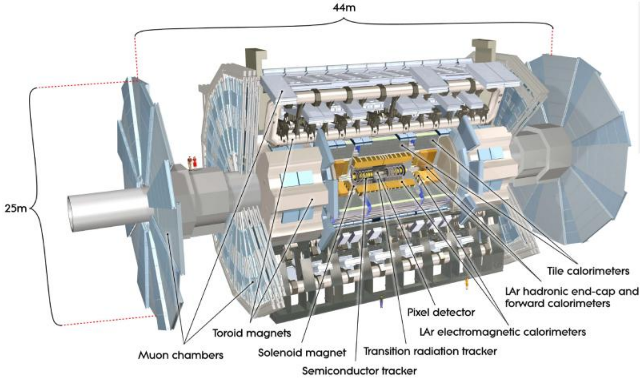
Detector reconstruction



Analysis

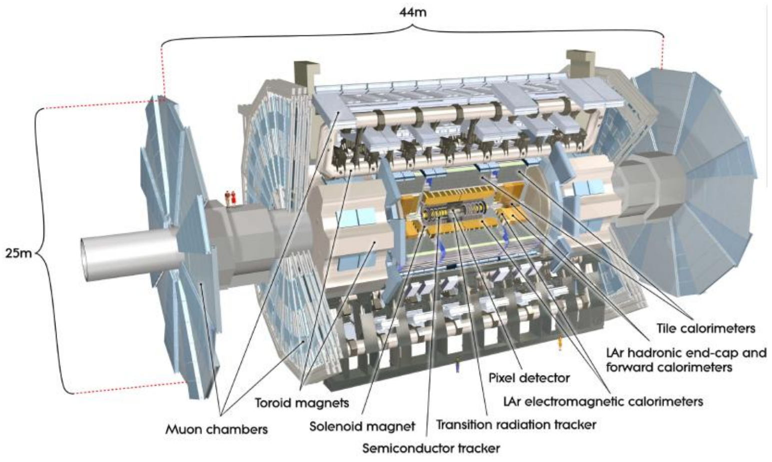


Physics model

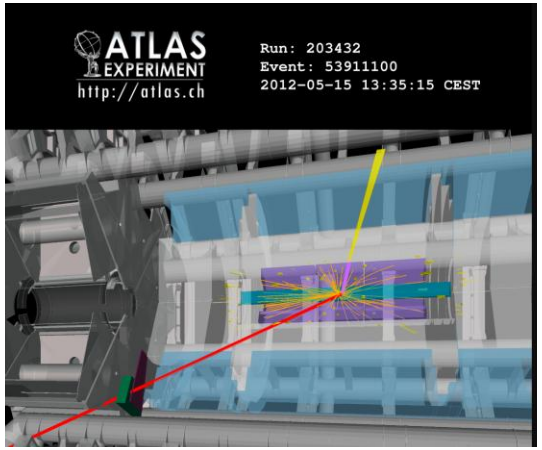


Detector simulation

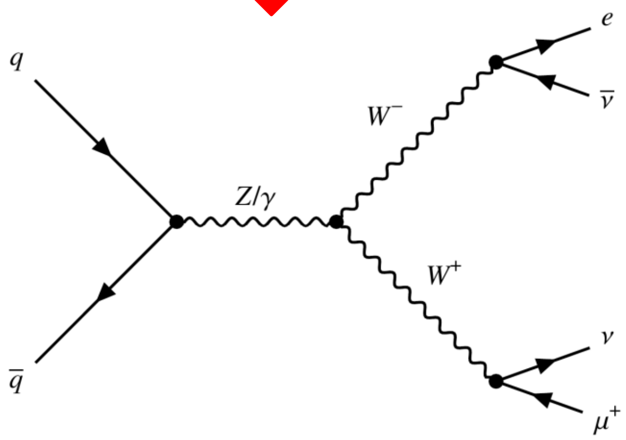
Unfolding analysis workflow



Detector simulation

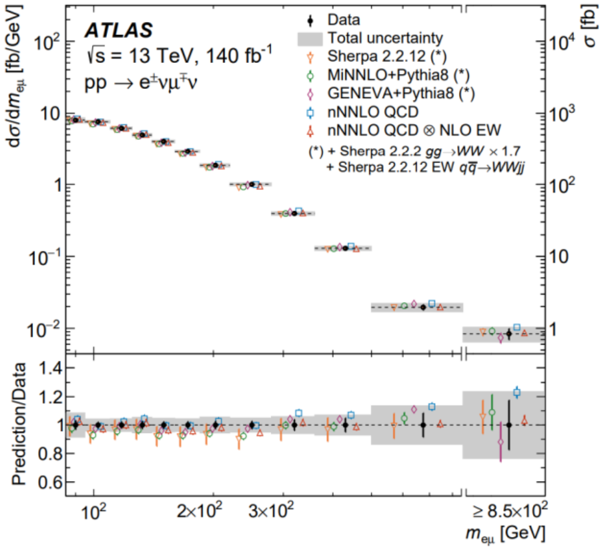


Detector reconstruction



Physics model

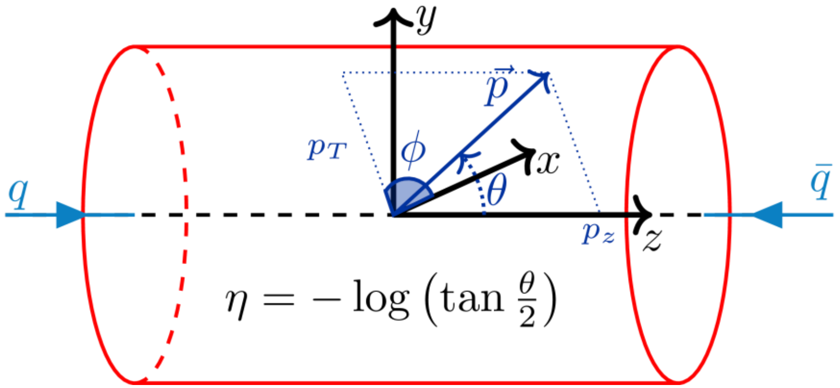
Unfolding



Variables to unfold

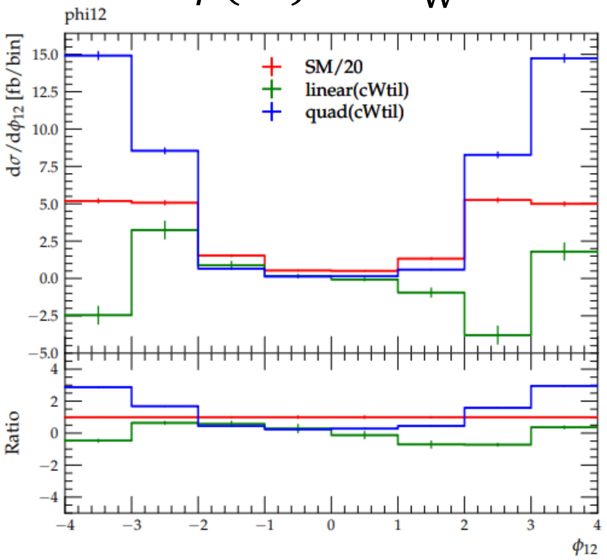
- Unfolding variables which are sensitive to polarisation and potential BSM effects, particularly EFT

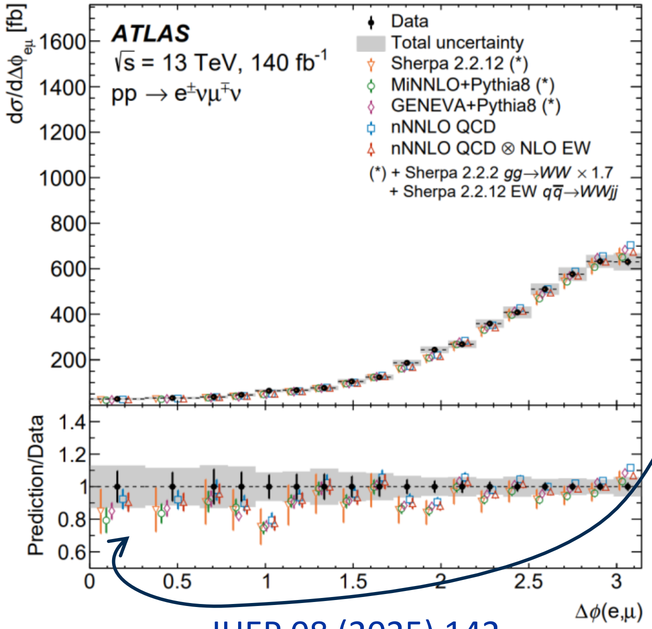
Variable	Description	Motivation
$p_T(\ell_1)$	p_T of the leading lepton	Interesting for single-boson polarisation studies
E_T^{miss}	Missing transverse energy	Unfolded in previous WW analyses
$m_{e\mu}$	Invariant mass of the electron-muon pair	Provides sensitivity to EFT and polarisation of the WW system
$p_T(e,\mu)$	Transverse momentum of the electron-muon pair	Provides sensitivity to PDFs, EFT and polarisation of the WW system
$\Delta\eta(\ell\ell)$	Pseudorapidity difference between rapidity-ordered leptons	Sensitive to CP-odd operators
$\Delta\phi(\ell\ell)$	Azimuthal difference between rapidity-ordered leptons	Sensitive to CP-odd operators, showed a mild tension with SM
$p_T(e,\mu)$ vs $\Delta\phi(\ell\ell)$	Double-differential cross section	Study $\Delta\phi(\ell\ell)$ only in the high-p_T region, where EFT effects rise



ATLAS coordinate system

$\Delta\phi(\ell\ell)$ for $O_{\widetilde{W}}$

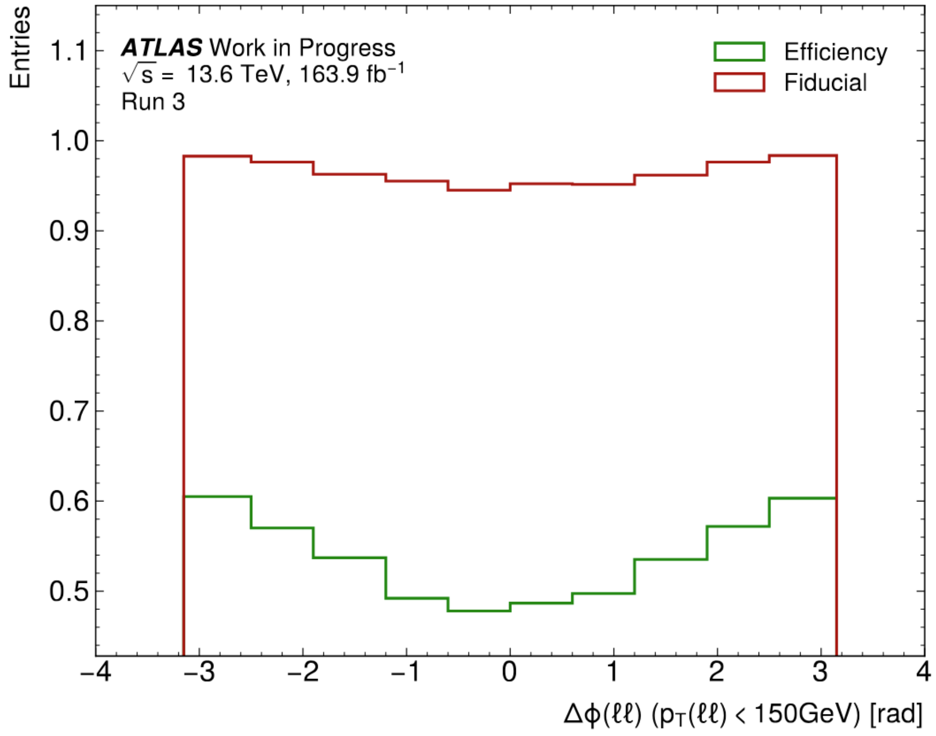


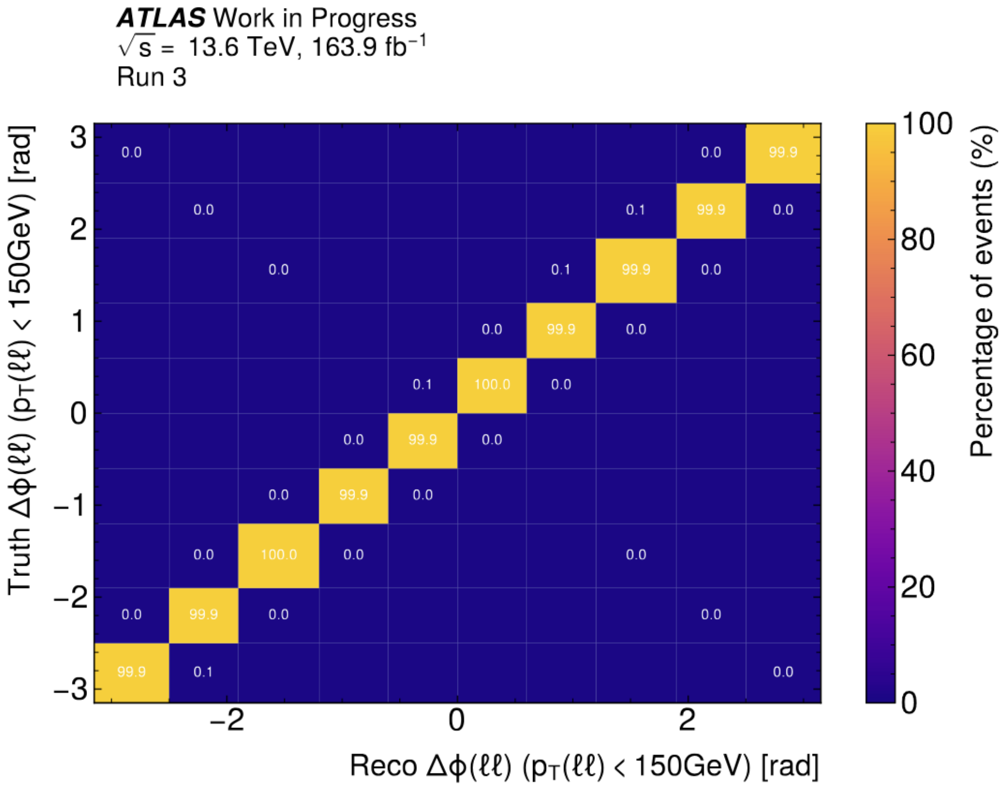


JHEP 08 (2025) 142

Unfolding $\Delta\phi(\ell\ell)$ vs $p_T(\ell\ell)$

- Low-$p_T(\ell\ell)$ bin: $0 \leq p_T(\ell\ell) \leq 150$ GeV
- Bin optimisation:
  - $\geq 70\%$ diagonal in Migration Matrix (limit migration) and ≥ 400 signal events per bin
  - Stack plot (left), fiducial and efficiency corrections (middle) and MM (right)



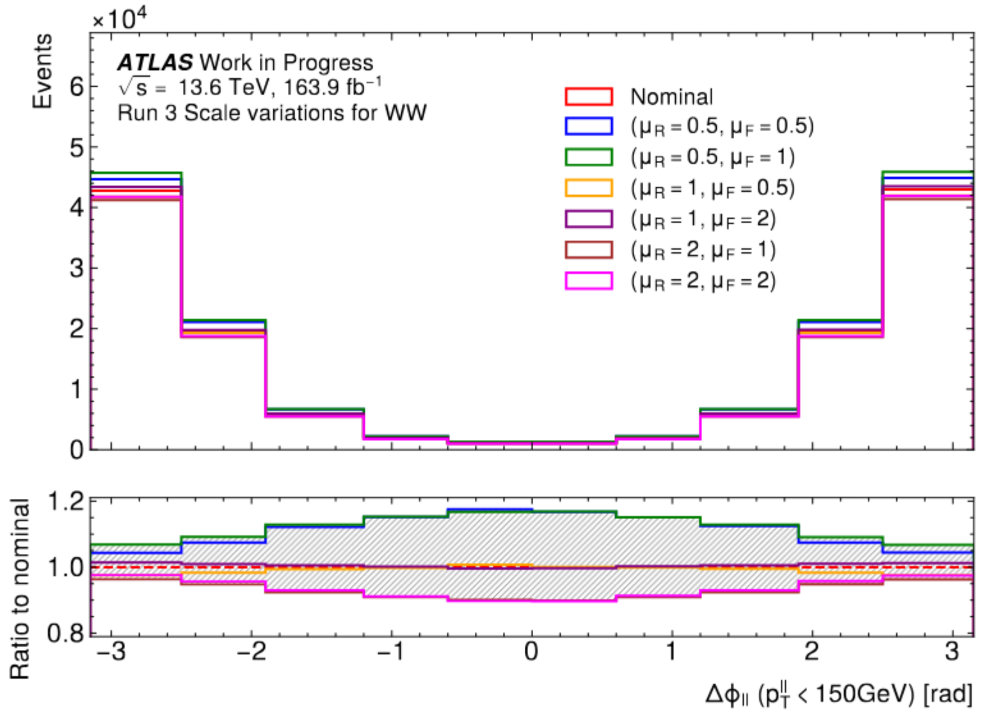


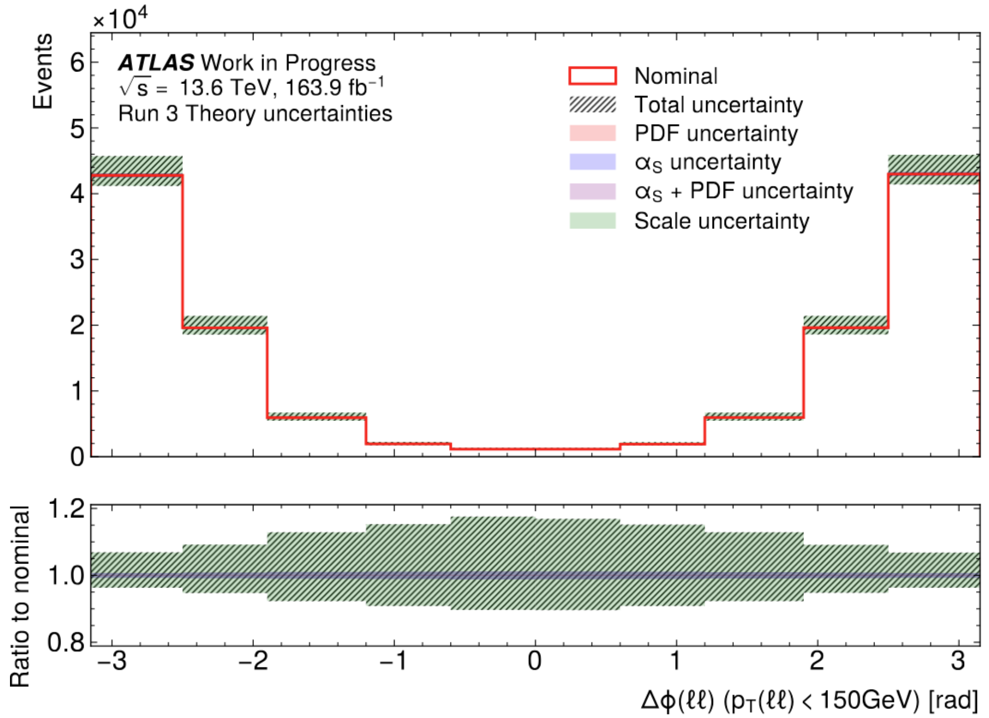
Signal truth: variations

- Renormalisation and factorisation scale:
  - μ_R affects α_S directly $\alpha_S(\mu_R^2) = \dfrac{\alpha_S(\mu_0^2)}{1 + \beta_0 \alpha_S(\mu_0)^2 \ln \frac{\mu_R^2}{\mu_0^2}}$
  - μ_F affects PDFs
- Evaluated as 7-point envelope

$$\{(\mu_R,\ \mu_F)\} = \{(0.5, 0.5); (0.5, 1); (1, 0.5); (1, 1); (2, 1); (1, 2); (2, 2)\}$$

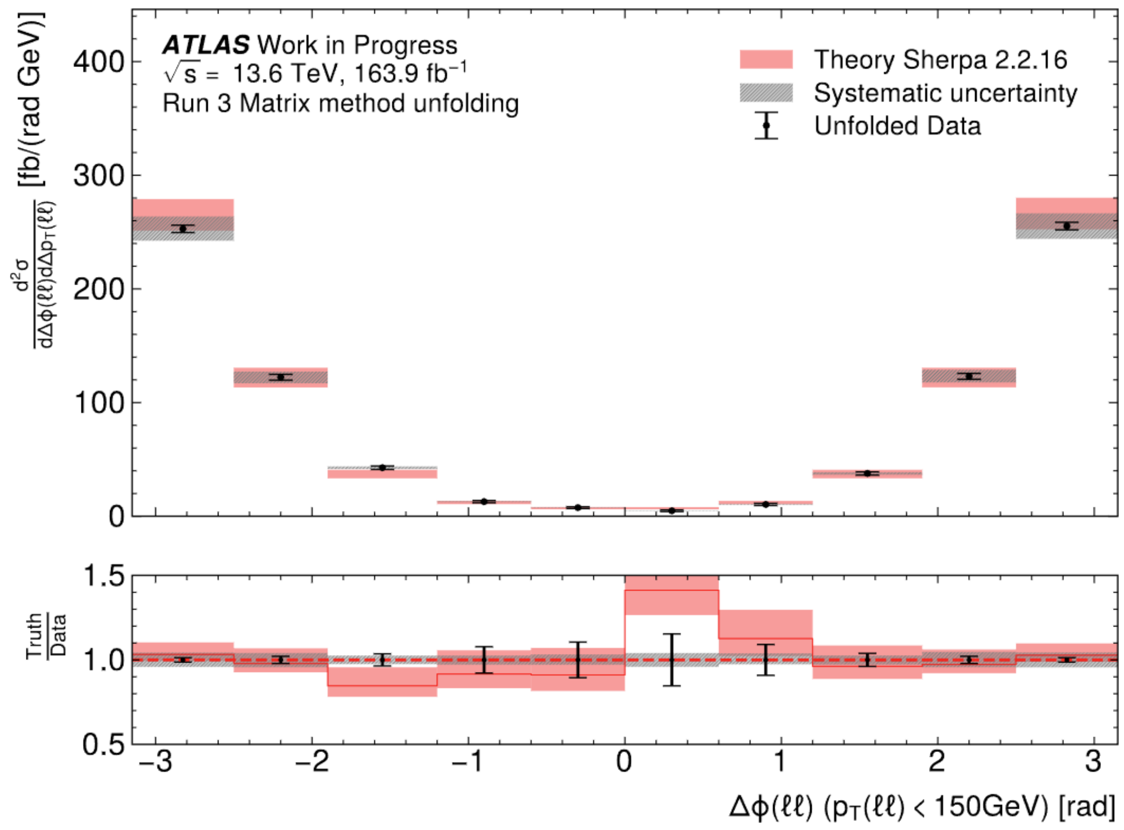
- PDF and α_S:
  - Using PDF4LHC21 recommendations
    - PDF uncertainty evaluated through 100 variations
    - α_S varied between 0.118 (nominal) $\pm$ 0.001
    - Two uncertainties combined and symmetrised
- Scale uncertainty dominates



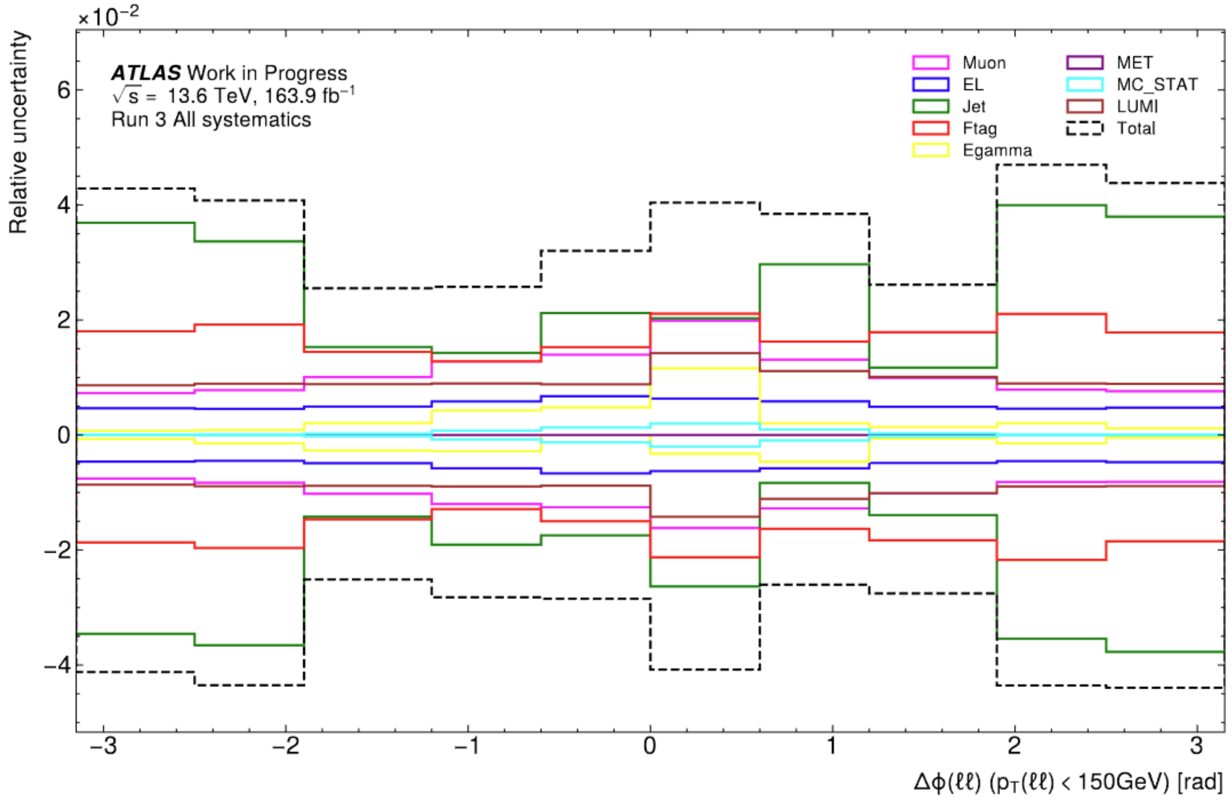


Data unfolding w/ systematics

Good agreement with Sherpa theory prediction



Experimental systematic uncertainty less than 5% in every bin



Summary and conclusions

- WW process is sensitive to potential EFT and CP-violating effects
- Presented first look into OSOF WW differential measurement with ATLAS Run3 data
  - $163\ \mathrm{fb}^{-1}$ of data -> highest recorded to date
  - Focus on enhancing sensitivity to linear EFT over previous measurements, particularly on triple gauge coupling
  - First measurement of CP-odd sensitive variables (signed $\Delta\phi(\ell\ell)$)
  - Will extend to neural network observables to probe CP-violating effects
- Good agreement between theory and prediction for CP-odd sensitive variables
- Analysis nearing ATLAS internal review, stay tuned for publication this Summer!

Thank you for your attention

Alberto Plebani

HARDING
Distinguished Postgraduate
SCHOLARS PROGRAMME

ATLAS
EXPERIMENT

Event selection and yields

- Event selection mimics Run2 selection as close as possible, except:
  - GN2 @90% eff (before was DL1r @85%)
  - Updated isolation and identification WPs

Event selection		
Exactly 2 opposite-charged leptons ($1e$ and 1μ), $p_T > 27$GeV $\lvert\eta(e)\rvert \leq 2.47$ with crack veto $1.37 \leq \lvert\eta(e)\rvert \leq 1.52$ $\lvert\eta(\mu)\rvert \leq 2.5$		
Electrons: ID: TightLH, Isol: Tight_VarRad Muons: ID: Medium, Isol: PFlowTight_VarRad		
$\lvert d_0/\sigma_{d_0}\rvert < 5(e),\ 3(\mu),\ \lvert z_0 \cdot \sin\theta\rvert < 0.5$ mm		
No additional electron with: $p_T > 10$ GeV, $\lvert\eta\rvert \leq 2.47$, ID: LooseBLayerLH, Isol: Loose_VarRad		
No additional muon with: $p_T > 10$ GeV, $\lvert\eta\rvert \leq 2.7$, ID: Loose, Isol: NonIso		
Inclusive in jet multiplicity for $p_T(j) > 30$ GeV and $\lvert\eta(j)\rvert \leq 4.5$ JVT and fJVT applied N_{b-jets} @90% eff. (GN2 PCBT > 1) $= 0$		
$m_{e\mu} >$ 100 GeV Veto on non-prompt leptons		
Triggers:		
mc23a	mc23d	mc23e
HLT_e26_lhtight_ivarloose_L1EM22VHI HLT_e60_lhmedium_L1EM22VHI HLT_e140_lhloose_L1eEM26M HLT_mu24_ivarmedium_L1MU14FCH HLT_mu50_L1MU14FCH	HLT_e26_lhtight_ivarloose_L1eEM26M HLT_e60_lhmedium_L1eEM26M HLT_e140_lhloose_L1EM22VHI HLT_mu24_ivarmedium_L1MU14FCH HLT_mu50_L1MU14FCH	HLT_e26_lhtight_ivarloose_L1eEM26M HLT_e60_lhmedium_L1eEM26M HLT_e140_lhloose_L1EM22VHI HLT_mu24_ivarmedium_L1MU14FCH HLT_mu50_L1MU14FCH

Sample	Yields ± Stat.	Percentage (%)
WW	57 485.7 ± 77.2	34.0
$Z + jets$	3636.5 ± 324.4	2.2
WZ	2366.5 ± 9.2	1.4
VV	130.6 ± 1.1	0.1
$t\bar{t}$	77 451.1 ± 762.8	45.8
$V\gamma$	174.1 ± 25.5	0.1
$W + jets$	100.7 ± 64.6	0.1
VVV	201.1 ± 0.5	0.1
VV EWK	768.0 ± 9.0	0.5
Single-top	19 748.4 ± 28.8	11.7
Fakes	6877.6 ± 72.4	4.1
Total mc23ade	168 940.4 ± 839.2	1
Run 3 Data	170 773.0 ± 413.2	-
Data/MC	1.01	-

Fake lepton background

- Small but not insignificant contribution of fake leptons to the signal region (<10%), largely from W+jets.
- Data-driven methods were used to calculate the fake lepton background.
- The nominal fake lepton background was estimated using the asymptotic matrix method.
- Fake factor method also used to provide a systematic variation.
- Fake lepton efficiencies were calculated in a Z+jets CR
- Fake efficiencies also calculated in a ttbar CR to provide a systematic variation for fake leptons originating in heavy flavour decays.

Object	Requirment	Criteria
Lepton	p_T	>27 GeV (Leading electron), >10 GeV (Sub-leading electron) OR >25 GeV (Leading muon), >10 GeV (Sub-leading muon)
	$\|\eta\|$	<2.47 (electron), <2.5 (muon)
	Identification	LooseBLayerLH (electron), Loose, (muon)
	Isolation	Loose_VarRad (electron), NonIso (muon)
	Track Parameters	$\|d_0$ significance$\| < 5$ (electron), $\|z_0 \sin\theta\| < 0.5$ (electron)
	Overlap Removal	passes $e-\mu$ OR, $e-jet$ OR, passes $\mu-jet$ OR
Jet	p_T	>35 GeV
	$\|\eta\|$	<2.5
	b-tagging	GN2 PCBT >1 (equivalent to 90% eff. WP), $\|\eta\|<2.5$
	Vertex Tagging	JVT FixedEffPt WP
Event	Trigger	HLT_e26_lhtight_ivarloose_L1EM22VHI OR HLT_e60_lhmedium_L1EM22VHI OR HLT_e140_lhloose_L1EM22VHI OR HLT_e140_lhloose_L1EM22VHI OR HLT_mu24_ivarmedium_L1MU14FCH OR HLT_mu50_L1MU14FCH
	Trigger-matching	Tag-lepton required to be matched to at least one trigger decision
	Leptons	2 leptons, same flavour, opposite sign pair, no additional electrons of $p_T > 10$ GeV, LooseBLayerLH Identification and Loose_VarRad Isolation, no additional muons of $p_T > 10$ GeV, Loose Identification and NonIso Isolation
	Number of b-jets	0
	m_{ll} mass window	$\|m_Z - m_{ll}\| < 10$ GeV

Z+jets CR definition

Fake leptons VR

- Fakes VR:
  - Orthogonal to SR due to same-sign requirement on $\ell\ell$ pair
  - Good agreement in both normalisation and shape for mc23a

Sample	Yields		Percentage (%)
WW	1.6 ±		0.1
ttbar	282.2 ±		15.1
Fakes	827.6 ±		44.3
VV	596.3 ±		31.9
Single top	4.0 ±		0.2
Z+jets	57.8 ±		3.1
VVV	17.4 ±		0.9
VV+jets	83.3 ±		4.5
Vgamma	0.1 ±		0.0
Total	1870.2 ±	nan	1
Data	1870.0 ±	43.2	-
Data/MC	1.00		-

Object	Requirment	Criteria
Lepton	p_T	$>$ 27 GeV (electron), $>$ 27 GeV (muon)
	$\|\eta\|$	$<$ 2.47 (electron), $<$ 2.5 (muon)
	Identification	TightLH (electron), Medium, (muon)
	Isolation	Tight_VarRad (electron), PflowTight_VarRad (muon)
	Track Parameters	$\|d_0$ significance$\| < 5$ (electron), $\|d_0$ significance$\| < 3$ (muon), $\|z_0 \sin\theta\| < 0.5$
	Overlap Removal	passes $e-\mu$ OR, $e-jet$ OR, passes $\mu-jet$ OR
Jet	p_T	$>$ 30 GeV
	$\|\eta\|$	$<$ 4.5
	b-tagging	GN2 PCBT $>$ 1 (equivalent to 90% eff. WP), $\|\eta\| < 2.5$
	Vertex Tagging	JVT FixedEffPt WP
Event	Trigger	HLT_e26_lhtight_ivarloose_L1EM22VHI OR HLT_e60_lhmedium_L1EM22VHI OR HLT_e140_lhloose_L1EM22VHI OR HLT_e140_lhloose_L1EM22VHI OR HLT_mu24_ivarmedium_L1MU14FCH OR HLT_mu50_L1MU14FCH
	Leptons	1 electron and 1 muon of same sign, no additional electrons of $p_T >$ 10 GeV, LooseBLayerLH Identification and Loose_VarRad Isolation, no additional muons of $p_T >$ 10 GeV, Loose Identification and NonIso Isolation
	Number of b-jets	0
	$m_{e\mu}$	$>$ 85 GeV

ttbar background

- Ttbar constitutes the largest background in the signal region
- A data-driven approach known as b-tag counting was used to estimate it.
- This approach was adopted in the Run 2 WW analysis.
- Ttbar yields are derived from data in 2b-jet and 1b-jet CRs
- The yield in the 0b-jet signal region is estimated by exploiting correlation effects between the 1b and 2b CR yields

$$N_{1b}^{t\bar{t}} = N_{1b} - N_{1b}^{\text{others}} = N_{\geq 0b}^{t\bar{t}} \cdot 2\varepsilon_b \left(1 - C_b \varepsilon_b\right) \,,$$
$$N_{2b}^{t\bar{t}} = N_{2b} - N_{2b}^{\text{others}} = N_{\geq 0b}^{t\bar{t}} \cdot C_b \varepsilon_b^2 \,,$$
$$N_{0b}^{t\bar{t}} = N_{\geq 0b}^{t\bar{t}} \cdot \left(1 - 2\varepsilon_b + C_b \varepsilon_b^2\right) \,,$$

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

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

ttbar VR

- ttbar VR:
  - Extra cuts at $m_{lj} < 140\ GeV$ and $\Delta\phi(e\mu) < \pi/2$
- Good agreement in shape and normalisation

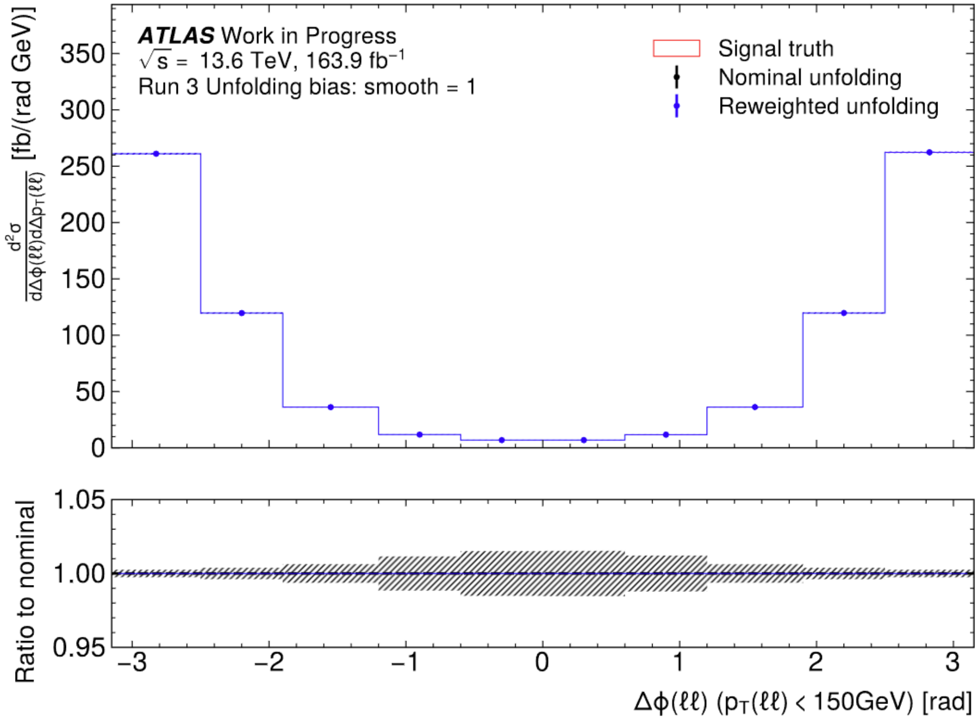
Sample	Yields		Percentage (%)
WW	223.1 ±		17.6
ttbar	810.0 ±		64.0
Fakes	51.8 ±		4.1
VV	25.5 ±		2.0
Single top	146.1 ±		11.5
Z+jets	5.5 ±		0.4
VVV	1.2 ±		0.1
VV+jets	2.9 ±		0.2
Vgamma	0.1 ±		0.0
Total	1266.3 ±	95.2	1
Data	1357.0 ±	36.8	-
Data/MC	1.07		-

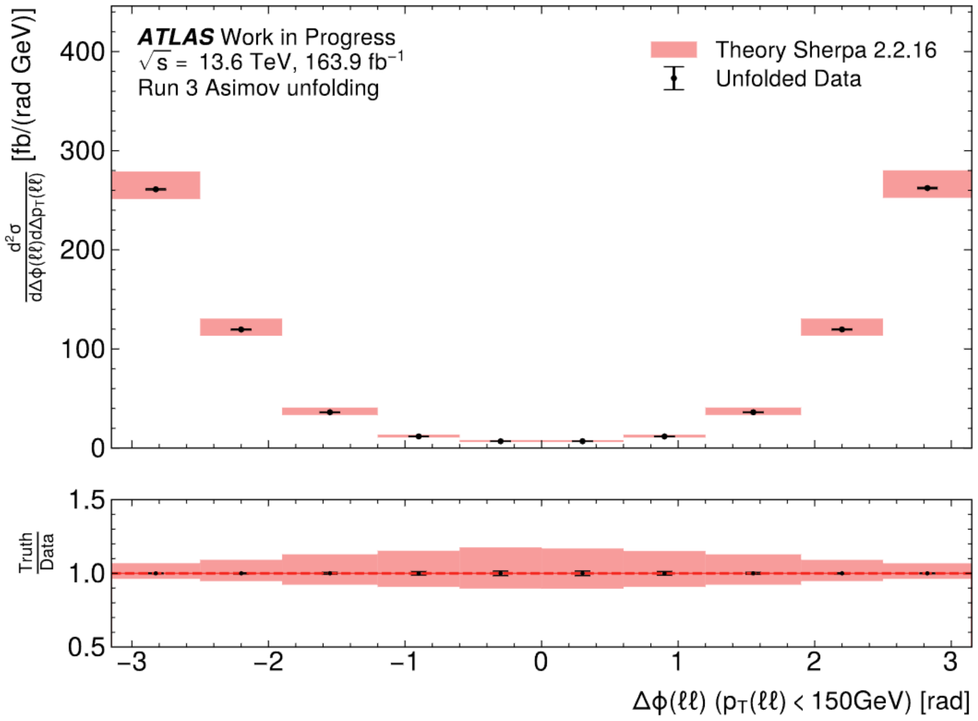
Object	Requirment	Criteria
Lepton	p_T	> 27 GeV (electron), > 27 GeV (muon)
	$\|\eta\|$	< 2.47 (electron), < 2.5 (muon)
	Identification	TightLH (electron), Medium, (muon)
	Isolation	Tight_VarRad (electron), PflowTight_VarRad (muon)
	Track Parameters	$\|d_0$ significance$\| < 5$ (electron), $\|d_0$ significance$\| < 3$ (muon), $\|z_0 \sin\theta\| < 0.5$
	Overlap Removal	passes $e-\mu$ OR, $e-jet$ OR, passes $\mu-jet$ OR
Jet	p_T	> 30 GeV
	$\|\eta\|$	< 4.5
	b-tagging	GN2 PCBT > 1 (equivalent to 90% eff. WP), $\|\eta\| < 2.5$
	Vertex Tagging	JVT FixedEffPt WP
Event	Trigger	HLT_e26_lhtight_ivarloose_L1EM22VHI OR HLT_e60_lhmedium_L1EM22VHI OR HLT_e140_lhloose_L1EM22VHI OR HLT_e140_lhloose_L1EM22VHI OR HLT_mu24_ivarmedium_L1MU14FCH OR HLT_mu50_L1MU14FCH
	Leptons	1 electron and 1 muon of opposite charge, no additional electrons of $p_T > 10$ GeV, TightLH Identification and Tight_VarRad Isolation, no additional muons of $p_T > 10$ GeV, Medium Identification and PflowTight_VarRad Isolation
	Number of b-jets	0
	$m_{e\mu}$	> 85 GeV
	$m_{\ell j}$	< 140 GeV
	$\Delta\phi_{e\mu}$	$< \pi/2$

Unfolding: bias uncertainty and closure

- Data-driven bias uncertainty
  - Obtained by reweighting MC to reflect data

- Closure test:
  - Unfold MC-to-MC
  - Perfect closure observed





Unfolding setup

- Unfolding performed using custom-made IBU script with RooUnfold
- Bin optimised so that each bin has at least 500 entries **and** at least 70% in diagonal of the migration matrix (details in backup)
- Performed unfolding both MC-to-MC (closure test) and using real Data
  - For fakes, using both MC and data-driven fakes estimation
  - Plan to include data-driven estimation also for ttbar
- Uncertainties:
  - All instrumental systematics added
  - Currently working on theoretical (PDF and scale)
- Added theoretical uncertainties on the truth distribution

Theory: scale uncertainties calculation

- Scale variations:
  - 7-point envelope:
    Nominal
    - $\{(\mu_R,\ \mu_F)\} = \{(0.5, 0.5); (0.5, 1); (1, 0.5); (1, 1); (2, 1); (1, 2); (2, 2)\}$
- Ex. for signal $W^+W^- \rightarrow e^{\pm}\nu\mu^{\mp}\nu$ (701050, *Sh_llvv_os*):
  - Take syst-variated reco and truth for each pair of μ_R and μ_F and calculate MM, efficiency, and acceptance
  - Take NOSYS for backgrounds to subtract from data
  - Unfold the syst-variated truth
  - Get the uncertainty
- Repeat for all processes (WZ, ttbar, Zee, and so on) independently
  - Still not fully implemented

Theory: PDF uncertainties calculation

- Working with the 100 NNPDF variations, combining them according to PDF4LHC21 recs:
  - (N_eig = 100)

$$\delta^{\mathrm{pdf}}\mathcal{F} = \sqrt{\sum_{k=1}^{N_{\mathrm{eig}}} \left(\mathcal{F}^{(k)} - \mathcal{F}^{(0)}\right)^2}$$

  - Two extra variations for $\alpha_S = 0.117$ and $\alpha_S = 0.119$ (nominal is $\alpha_S = 0.118$)

$$\delta^{\alpha_s}\sigma = \frac{\sigma(\alpha_s = 0.119) - \sigma(\alpha_s = 0.117)}{2},$$

  - Combining the two:
    - $$\delta^{\mathrm{PDF}+\alpha_s}\sigma = \sqrt{(\delta^{\mathrm{pdf}}\sigma)^2 + (\delta^{\alpha_s}\sigma)^2}\,.$$

- Consider together all processes with the same PDF set
  - Ex. take all Sherpa processes which use NNPDF3.0 NNLO Hessian PDF set, change reco and truth (only if signal) and calculate MM, efficiency, and acceptance. Use the syst-variated estimation of the backgrounds (if affected, if not consider NOSYS) and unfold
    - Obtain one uncertainty for every group of PDF (NNPDF3.0 NLO Hessian, NNPDF3.0 NLO 4FS, NNPDF3.0 NLO, NNPDF3.0 NNLO Hessian, NNPDF3.0 NNLO 4FS)
  - Currently implemented only for signal, working on adding others

Theory uncertainties

DSID	Process	Generator	Nominal Hessian PDF	PDF set	PDF for scale	PDF set for scale	Scale variations names
522024	$t\bar{t}Z(\to e\bar{e})$	MadGraph	93300 (40 + 2 for α_S)	PDF4LHC21_40	303000	NNPDF NLO Hessian	MUR10_MUF20
522028	$t\bar{t}Z(\to \mu\bar{\mu})$	MadGraph	93300 (40 + 2 for α_S)	PDF4LHC21_40	303000	NNPDF NLO Hessian	MUR10_MUF20
522032	$t\bar{t}Z(\to \tau\bar{\tau})$	MadGraph	93300 (40 + 2 for α_S)	PDF4LHC21_40	303000	NNPDF NLO Hessian	MUR10_MUF20
525955	tWZ	MadGraph	93300 (40 + 2 for α_S)	PDF4LHC21_40	303000	NNPDF NLO Hessian	MUR10_MUF20
545027	tZq	MadGraph	93700 (40 + 2 for α_S)	PDF4LHC21_40_4FS	260400	NNPDF NLO 4FS	MUR10_MUF20
545028	$\bar{t}Zq$	MadGraph	93700 (40 + 2 for α_S)	PDF4LHC21_40_4FS	260400	NNPDF NLO 4FS	MUR10_MUF20
601230	$t\bar{t}$	PowhegPythia	93300 (40 + 2 for α_S)	PDF4LHC21_40	260000	NNPDF NLO	MUR1_MUF2
601354	tW	PowhegPythia	93300 (40 + 2 for α_S)	PDF4LHC21_40	260000	NNPDF NLO	MUR1_MUF2
700772	$\tau\tau\gamma$	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700792	$Z \to \tau\bar{\tau}$ (+B)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700793	$Z \to \tau\bar{\tau}$ (+C)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700794	$Z \to \tau\bar{\tau}$ (+ LF)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700901	$Z \to \tau\bar{\tau}$ low_Mll (+B)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700902	$Z \to \tau\bar{\tau}$ low_Mll (+C)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700903	$Z \to \tau\bar{\tau}$ low_Mll (+LF)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700997	$t\bar{t}W$	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701000	$\ell\ell\ell\ell jj$ (EW)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701005	$\ell\ell\ell\nu jj$ (EW)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701010	$\ell\ell\nu\nu jj$ (OS) (EW)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701020	$\ell\ell\ell\ell jj$ (Int. EW)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701025	$\ell\ell\ell\nu jj$ (Int. EW)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701030	$\ell\ell\nu\nu jj$ (OS) (Int. EW)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701040	$\ell\ell\ell\ell$	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701045	$\ell\ell\ell\nu$	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701050	$\ell\nu\ell\nu$ (OS)	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701085	$Z(\to qq)Z(\to \ell\ell)$	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701090	$Z(\to bb)Z(\to \ell\ell)$	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701105	$W(\to qq)Z(\to \ell\ell)$	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
701165	$\ell\ell\gamma jj$	Sherpa	93300 (40 + 2 for α_S)	PDF4LHC21_40	303200	NNPDF NNLO Hessian	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700866	$WWW \to \ell\ell\ell\nu\nu\nu$	Sherpa	93700 (40 + 2 for α_S)	PDF4LHC21_40_4FS	261400	NNPDF NNLO 4FS	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2
700867	$WZZ \to \ell\ell\ell\nu\nu\nu$	Sherpa	93700 (40 + 2 for α_S)	PDF4LHC21_40_4FS	261400	NNPDF NNLO 4FS	MUR1_MUF2_PDF303200_PSMUR1_PSMUF2

Iterative Bayesian Unfolding

- Assume an initial hypothesis for the probability density of the truth distribution f(t)
  - $f(t) = \int dr\ g(r)\ P(t|r)$, with $g(r)$ taken from data and $P(t|r)$ taken from MC simulations
  - $P(t|r)$ is then updated iteratively as
  - The n^{th} step is obtained as

$$P(t|r) = \frac{P(r|t)f(t)}{\int f(t)P(r|t)dt} = \frac{P(r|t)f(t)}{g(r)}$$

$$f^{n+1}(t) = \int f^{r}(t)\frac{g(r)_{\mathrm{data}}}{g^{n}(r)}P(r|t)dt$$

  - Procedure repeated until difference between $f^{n+1}(t)$ and $f(t)$ is less than a certain threshold

Bootstrap

- Implemented bootstrap technique in unfolding calculation to estimate statistical uncertainty
- Error comparison:
  - Stat_uncertainty is the error coming from the "nominal" unfolding
  - Bootstrap_error is the one coming from the bootstrap
  - Comparable, with bootstrap being usually smaller

```
(Pdb) print(np.array(stat_uncertainty))
[39.66092158 35.32825551 46.28077764 59.5925267  71.00373422 80.09907369
 86.42675813 84.24197714 88.01504912 89.05609265 87.59173586 83.66074768
 81.72525735 77.31381775 77.15241043 73.69661204 65.25220743 60.70280759
 54.10084857 48.68811855 46.06392141 40.08269998 36.10017056 28.05210092
 27.99039308 25.7299486  30.99368253]
(Pdb) print(np.array(bootstrap_error))
[38.44243091 35.37961289 45.79619605 59.46766351 72.612459   78.36038555
 87.00428931 82.43005963 85.84026613 87.6244471  85.96676988 84.16738236
 82.73307236 76.96109972 73.74238683 76.0298312  65.82961615 61.14980692
 55.00073527 47.50449302 47.10482908 39.70725019 35.10139281 28.12479942
 27.75629444 26.1913554  32.08084817]
(Pdb) difference_error = bootstrap_error - stat_uncertainty
(Pdb) print(np.array(difference_error / stat_uncertainty))
[-0.0307227   0.00145372 -0.01047047 -0.00209528  0.0226569  -0.02170672
  0.00668232 -0.02150849 -0.02470922 -0.01607577 -0.01855159  0.00605582
  0.01233174 -0.00456216 -0.04419854  0.03165979  0.00884888  0.00736373
  0.0166335  -0.02431036  0.02259703 -0.00936688 -0.02766684  0.00259155
 -0.00836354  0.01793267  0.03507701]
(Pdb)
```

7.1.3 Bootstrap techniques for statistical uncertainty

The statistical uncertainty on the unfolded distribution is estimated using the bootstrap method. For this, the `BootstrapGenerator` package [22] was used. This method consists of generating a set of replicas of the measured spectrum, all derived by fluctuating each event in the data distribution by a Poisson distribution centred around the value of the measured data. In practice, this is obtained by performing the following steps:

- Generate a response by using the background-subtracted-data as per Eq. (7.15)
- For every bin in the data distribution, get the total number of expected events N_{meas} as the number of measured data events
- Sample N_{meas} times a weight $\omega_i \sim \text{Poisson}(1)$ from a Poisson distribution with mean 1, and fill the new histogram with these weights
- Repeat this procesure 1000 times. Now every bin will have entries distributed according to a Poisson distribution centred around the measured number of events.
- For every replica ($i = 1, \ldots, 1000$), do:
- 1. Get the reweighted data distribution
  2. Build a reweighted background-subtracted data using Eq. (7.15)
  3. Unfold the reweighted background-subtracted data to the truth with the nominal response
  4. Store the entries of the new unfolded distribution
- Once the loop is completed, the uncertainty in each bin j from the 1000 variations is cpmputed as the RMS of all the variations, as per Eq. (7.16), where $N = 1000$, x_i^j are the entries from the i-th variation in the j-th bin, and $\bar{x}^j$ refers to the entries in the j-th bin of the nominal unfolded distribution.

$$\sigma_U^j = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i^j - \bar{x}^j)^2} \qquad (7.16)$$

Dim-6 SMEFT interpretation

- Plan to do EFT interpretation on **CP-even/odd dim-6 operators**
  - Set limits on 4 Wilson coefficients: **cW, cHWB, cWtil, CHWBtil**
  - mc23a/d/e samples requested for **linear, quadratic and cross terms**
  - MC request in EW subgroup and the production JIRAticket
    - mc23a/d are done, waiting for some tails in mc23e to finish

- The cross-section prediction in each bin:

$$f^{b}_{\text{pred}}\ (\boldsymbol{c},\boldsymbol{\theta}) = f^{b}_{\text{pred, SM}}\left(1+\sum_i \underbrace{A_{bi}}_{\text{linear}}\frac{c_i}{\Lambda^2}+\sum_i \underbrace{B_{bi}}_{\text{quadratic}}\frac{c_i^2}{\Lambda^4}+\sum_{i<j}\underbrace{C_{bij}}_{\text{Cross}}\frac{c_i c_j}{\Lambda^4}\right)\prod_n^{n_{\text{theory,syst}}}(1+u_n\theta_n)$$

- The Likelihood function for EFT fit

$$L(\mathbf{x}\mid\mathbf{c},\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^{n_{\text{bins}}}\det(V)}}\exp\left(-\frac{1}{2}\Delta\mathbf{x}^{\top}(\mathbf{c},\boldsymbol{\theta})\,V^{-1}\,\Delta\mathbf{x}(\mathbf{c},\boldsymbol{\theta})\right)\times\prod_{i=1}^{n_{\text{theo syst}}} f_i(\theta_{\text{theo syst},i})\times\prod_{i=1}^{n_{\text{exp syst}}} f_i(\theta_{\text{exp syst},i})$$

EFT interpretation: plan

- Fitting tool: EFT-fun Fit setup:
  - Unfolded distributions of **mll** (only mc23a used for this preliminary study)
    - Not particularly sensitive to EFT effects, used mainly to validate the framework and because we expect low sensitivity
  - EFT formulation only linear effect
  - Stat uncertainty on unfolded data - embedded in covariance matrix *V*
  - Experimental uncertainties on unfolded data - treated as NPs
  - Theoretical uncertainties on prediction - treated as NPs
- Will perform EFT study on $\Delta\phi(\ell\ell)$ which gives us sensitivity to CP-even and CP-odd operators
  - Will also add ONN which is expected to improve this even further

CPV in EFT

- EFT extensions to SM add CP-violating operators to the SM lagrangian
- The leading order BSM term in the matrix element is the interference between SM and EFT

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \sum_i \frac{c_i}{\Lambda^2}\tilde{\mathcal{O}}_i \longrightarrow |\mathcal{M}_{\mathrm{BSM}}|^2 = |\mathcal{M}_{\mathrm{SM}}|^2 + 2\mathrm{Re}\{\mathcal{M}_{\mathrm{SM}}\mathcal{M}^{\star}_{\mathrm{d6}}\} + |\mathcal{M}_{\mathrm{d6}}|^2$$

- CP-odd EFT operators => interference is also CP-odd
- Interference can be constructive or destructive
- Kinematic similarities between constructive and destructive interference events means that if the variable is insensitive to CP-odd nature (ex. pT), the interference effects cancel out
- Sensitivity can be “resurrected” by studing CP-odd sensitive observables arxiv.708.07823

Neural Network variable

- arxiv.2112.05052 and arxiv.2209.05143: ML models can be used to construct complex observables that are sensitive to CP-odd asymmetries
- By training a multi-class NN to classify constructive, destructive and SM events, a CP-sensitive observable can be defines as:
  - $O_{NN} = P_+ - P_-$