PDFs and recent CMS precision measurements

Josh Bendavid (MIT)

arXiv:2408.07622 CMS-PAS-SMP-23-002

Nov. 27, 2024 LHC EW precision sub-group meeting

Theoretical Considerations

- \bullet W (and Z) production at hadron colliders described by PDFs $+$ perturbative QCD and Electroweak calculations
	- Small additional non-perturbative effects from "intrinsic $k\tau$ "
- Relatively large theoretical uncertainties due to large logarithms at low W or Z p_T
- Usual strategy is to use precise $Z \rightarrow \ell \ell \, p_{\tau}$ spectrum from data to tune the theoretical prediction

Phys.Lett.B 845 (2023) 138125 Phys. Rev. D 107, L011506, 2023

Theoretical Considerations

- **•** PDFs are a challenge: In recent precision measurements at hadron colliders often a significant spread in measured values depending on the choice of PDF set
- Angular dependence of W and Z production can be decomposed in terms of angular coefficients/helicity cross sections:
- **•** This can be a useful way to factorize theoretical corrections and uncertainties

arXiv:2408.07622, arXiv:2309.12986

$$
\frac{d^5 \sigma}{dq_T^2 dy dm d\cos\theta d\phi} = \frac{3}{16\pi} \frac{d^3 \sigma^{U+L}}{dq_T^2 dy dm} [(1 + \cos^2\theta) + \frac{1}{2}A_0(1 - 3\cos^2\theta) + A_1 \sin 2\theta \cos\phi + \frac{1}{2}A_2 \sin^2\theta \cos 2\phi + A_3 \sin\theta \cos\phi + A_4 \cos\theta + A_5 \sin^2\theta \sin 2\phi + A_6 \sin 2\theta \sin\phi + A_7 \sin\theta \sin\phi]
$$

The CMS Detector

The CMS Detector

m_W Measurement at CMS

- Use well-understood subset of 13 TeV data: 16.8 fb⁻¹ from later part of 2016 run (\sim 30 mean interactions per crossing)
- **•** Focus on muon channel and kinematics
	- Larger experimental systematics for electrons and hadronic recoil, especially with higher pileup
- **General strategy:** Exploit large dataset, accurate modeling of uncertainties for maximal in-situ contraints on theoretical modeling

- **•** Reserve Z data as an independent cross-check as much as possible:
- Muon calibration from J/ψ , validated with Z
- **o** In-situ constraints on theory modeling from W data itself, independent validation with Z

m_W Measurement at CMS

- m_W extracted from profile likelihood fit to muon $(\eta, p_T, \text{charge})$
	- Thousands of bins and systematic variations
	- **Optimized Tensorflow-based fitting framework**
- **•** Building on experimental techniques, tools, and experience from W-like m_Z measurement (2016) and W rapidity-helicity measurement (2020) which established strong in-situ constraints on PDFs from charged lepton kinematics
- 4B fully simulated MC events, >100M selected W candidates
	- Significant computing/technical challenges for a measurement of this complexity

m_W Measurement at CMS

- **Enabling feature of the measurement:** Systematic variations in W p_T , rapidity, decay angles from QCD uncertainties, PDFs, have a different effect on the muon kinematics as compared to a change in m_W
- PDF and boson p_T modeling uncertainties are strongly constrained in-situ by the data

- \bullet $Z \rightarrow \mu\mu$ events are also selected with very similar selection
- One muon removed and treated as neutrino \bullet
- To avoid statistical correlations, apply trigger and use kinematics of positive (negative) muons for even (odd) numbered events
- \bullet Z mass can be extracted from single muon (η, p_{τ}) , charge) distribution as for W case
- Validates all aspects of the actual W measurement except for non-prompt and $Z \rightarrow \mu\mu$ background
- **•** Theory uncertainties are similar (but not identical) to final m_W measurement

Theoretical Modeling

- **Overall strategy:** construct the best possible theoretical model for the W and constrain in-situ directly with the W data
- Z data is "only" used for validation
- Nominal Theory uncertainties:
	- **•** Perturbative QCD
	- a PDF_s
	- Additional non-perturbative QCD (e.g. transverse momentum of partons within proton)
	- **A** Electroweak effects
- **In addition:** Helicity cross section fit is used as a cross-check which augments or replaces the theory uncertainties by directly varying the different components of the angular decomposition
	- Reduced theory/model-dependence at the cost of increased statistical uncertainty

Parton Distribution Functions

- **Good: PDF sets are accompanied** by uncertainty models with well defined correlations across phase space and between processes
- **Bad: Different PDFs don't** necessarily agree within their uncertainties
- **•** Missing higher order uncertainties, resummation corrections in predictions usually not included
	- Partly mitigated by tolerance factors, etc

Parton Distribution Functions

- **Strategy:** Scale prefit PDF uncertainties to ensure consistency between sets for measured m_W value
- **•** This procedure does not prove that e.g. NNPDF4.0 uncertainty is underestimated, only that it's too small to cover the central value of the other sets
- **CT18Z** is chosen as the nominal since it covers the others without scaling and with small uncertainty
	- But note that this set is amongst the largest in terms of nominal uncertainty

Parton Distribution Functions

- **Strategy:** Scale prefit PDF uncertainties to ensure consistency between sets for measured m_W value
- **•** Scaling factors are determined with analysis still blind by using pseudodata generated from each PDF set and fitting with every other PDF set and its uncertainty
- **•** n.b. symmetrization procedure is applied for asymmetric uncertainties which tends to increase the uncertainty for CT18 and MSHT
- **This procedure does not prove that e.g. NNPDF4.0 uncertainty is** underestimated, only that it's too small to cover the central value of the other sets
- CT18Z is chosen as the nominal since it covers the others without inflation and small uncertainty

PDF Compatibility with Data

MSHT20aN3LO

NNPDF3.1

NNPDF4.0

99.0/116

99.1/116

99.7/116

87

87

86

Saturated likelihood test statistic from simultaneous fit to Z $y^{\mu\mu}$ and W η^μ distributions for each PDF set

122.8/116

105.5/116

104.3/116

• No strong discriminating power/all sets give a good p-value (possibly due to other conservative uncertainties)

31

75

77

82

51

46

101.9/116

115.0/116

116.7/116

Angular Distributions

- **•** Missing higher order uncertainties propagated to angular coefficients through variations of μ_r and μ_f in MINNI OPS
- **While MiNNLOPS predicts** angular coefficients consistent with fixed order calculations. Pythia intrinsic k_T treatment actually modifies them somewhat
	- \bullet In particular A_1 and A_3 at low boson p_T due to isotropic smearing
- This effect may or may not be physical \rightarrow propagate the full difference as an additional uncertainty Eur.Phys.J.C 82 (2022) 8, 693

Helicity Cross Section Fit

- Theory model represents our best understanding of QCD and proton structure
- As an additional test of its validity, or in case of BSM physics in W production or decay, a less model-dependent measurement of m_W is useful
- **Basic strategy:** Parameterize theory uncertainty explicitly in terms of the 9 helicity cross sections $\sigma_i \equiv \sigma^{U+L} A_i$ instead of the PDF and non-perturbative models $+$ perturbative uncertainty, and fit the helicity cross sections (double-differential in W rapidity and p_T) together with m_W
- In this way theoretical uncertainties are "traded" for larger statistical uncertainties

$$
\frac{d^5 \sigma}{dq_T^2 dy dm d \cos \theta d\phi} = \frac{3}{16\pi} \frac{d^3 \sigma^{U+L}}{dq_T^2 dy dm} [(1 + \cos^2 \theta) + \frac{1}{2} A_0 (1 - 3 \cos^2 \theta) + A_1 \sin 2\theta \cos \phi + \frac{1}{2} A_2 \sin^2 \theta \cos 2\phi + A_3 \sin \theta \cos \phi + A_4 \cos \theta + A_5 \sin^2 \theta \sin 2\phi + A_6 \sin 2\theta \sin \phi + A_7 \sin \theta \sin \phi
$$

- With current data/observables not possible to simultaneously constrain all of the relevant helicity components, so cross sections are regularized via constraints to the nominal prediction
- Relevant theory uncertainties are retained since they have different correlations

W-like m_z result

 \bullet Nominal W-like result:

$$
m_Z - m_Z^{\rm PDG} = -6 \pm 14 \text{MeV}
$$

Even-odd event selection reversed (nearly statistically independent sample)

$$
m_Z - m_Z^{\rm PDG} = 8 \pm 14 \text{MeV}
$$

• All extracted m_Z values in agreement with the LEP/PDG value

W-like m_z result: Uncertainty Breakdown

- Largest uncertainties are statistical, muon calibration, angular coefficients
- Total uncertainty is well defined, but several different ways of decomposing statistical and systematics uncertainties
- When uncertainties are constrained in-situ, "global" impacts (used e.g. for ATLAS 2024 m_W measurement) tends to count them as part of the statistical uncertainties

Nuisance Parameters

m_W Measurement

• Now with all elements in place, on to the m_W measurement:

- **•** For the nominal measurement, total uncertainty is 9.9MeV
- Most precise measurement at the LHC and comparable to CDF precision

$\overline{m_W}$ result

$m_{W} = 80360.2 \pm 9.9$ MeV

- Compatible with the Standard Model expectation and with other \bullet measurements
- In clear tension with the CDF measurement \bullet

Helicity Fit Result: $m_W = 80360.8 \pm 15.2$ MeV

Validation: Simultaneous dilepton+W fit

- Nominal result is from fit to muon (n, p_{τ}) , charge) for W candidates alone
- **Interesting to compare with simultaneous** fit to $p_T^{\mu\mu}$ distribution from Z events
- \bullet Fit results propagated to inclusive W p_T distribution as for Z case shown previously
- Postfit W p_T distribution broadly consistent and with strong constraints from data
- $\Delta m_W = +0.6$ MeV with respect to nominal, uncertainty would decrease to 9.6 MeV
- \bullet But additional complications for W/Z correlations, so the nominal W only fit is more robust and is the nominal result

PDF Dependence of Result

• Scaling of prefit PDF uncertainties reduces the dependence on PDF set and brings the variations within the quoted PDF uncertainties

Additional Theory Cross Checks

Result is stable under variations of the TNP model and not very sensitive to changes in the initial prediction within the uncertainties

m_W result: Validation checks

- **O** Consistent results when extracting 48 independent m_W parameters split in charge and 24 η bins
- η -sign difference: $m_W^{\eta > 0} - m_W^{\eta < 0} = 5.8 \pm 12.4$ MeV

• Charge difference:
$$
m_W^+ - m_W^- = 57 \pm 30 \text{MeV}
$$

m_W result: Closer look at charge difference

- $m_W^+ m_W^- = 57 \pm 30$ MeV, p-value 6.0%
- \bullet Uncertainty on charge difference much larger than nominal m_W uncertainty
- Strong anti-correlations due to experimental uncertainties (alignment) and theory uncertainties related to W polarization (opposite-parity coupling of W to μ^+ and $\mu^-)$
- Correlation between charge difference and m_W itself is only 2%

m_W result: Closer look at charge difference

• Reminder: For W-like m_z fit:

 $m_Z^+ - m_Z^- = 31 \pm 32$ MeV (nominal)

 $m_Z^+ - m_Z^- = 6 \pm 32$ MeV (reversed even-odd event selection)

- No conclusive evidence for a systematic problem $(< 2\sigma)$
- **•** Statistical fluctuations from finite data and MC samples at the level of 16 MeV for $m_W^+ - m_W^-$
- Even extreme variations of the related systematics lead to small variations in m_W (< 1MeV), within associated uncertainties
- \bullet Possible/plausible scenario: $\sim 1\sigma$ off on alignment and A_i's plus $\sim 1\sigma$ statistical fluctuation corresponds to totally negligible effect on m_W (0.1MeV)

A_3 Variations By Charge

- A correlated variation of A_3 between W^+ and W^- produces an \bullet anti-correlated variation for the charged lepton kinematics
- The variation corresponding to switching off pythia intrinsic k_T for the angular coefficients mixes effects from A_1 and A_3

Higher order corrections for A_3

Coefficient A₃ for W⁺ (inclusive rapidity)

Weak Mixing Angle Measurement

 $\sin^2\theta_{\mathrm{eff}}^\ell$ extracted from unfolded A_4 or detector-level weighted A_{FB} in $Z/\gamma^* \to ee/\mu\mu$ events differential in dilepton rapidity and mass

Weak Mixing Angle: PDF Profiling

- Different regions of phase space have different sensitivity to PDFs vs sin $^2\theta$
- **•** PDF profiling (or numerically equivalent weighting) used to reduce PDF uncertainty (and dependence!)

Weak Mixing Angle: PDF Dependence

Weak Mixing Angle: PDF Constraints

Weak Mixing Angle: PDF Constraints

Backup

Cross checks for mW charge difference

- Key numbers to compare to:
	- α Calibration uncertainty on mM' : \triangle R MaV
	- Calibration uncertainty on α $mWh + mWh - 21.3 MMV$
	- Non-perturbative uncertainty \sim on angular coeffs (pythia shower kT) for mW: 1 MeV
	- Non-perturbative uncertainty \circ on angular coeffs (pythia shower kT) for mW+ - mW-: 14 MeV
	- Data+MC stat uncertainty \sim (global impacts) on mW+ mW-: 15.8 MeV
- N.b alternate calibrations don't necessarily have fully consistent/complete uncertainty models
- \bullet Various plausible shifts of systematic uncertainties can give large variations on mW+ - mW- but small variations on mW itself Always within the corresponding uncertainties for both cases
- \bullet Even extreme brute force variation of alignment parameters leads to very small change in mW
- No smoking gun, not possible to identify a single "cause" \bullet
- Likely a combination of a few systematic effects (alignment, angular coefficients) at the 1 sigma level, combined with a statistical \bullet fluctuation

slide to be reformatted/refined

Charge Difference Impacts: Nominal

Charge Difference Impacts: Global

Table M_W re:

Leptor Leptor Recoil Recoil Leptor

QED ra W bos Total

Theoretical Modeling: Technical Details

- **•** Fully coherent theoretical treatment for W and Z (both μ and τ decays)
- **•** Fully simulated MC samples with $MINDOPS + Pythia 8 + Photos$
	- ${\cal O}(\alpha_s^2)$ accuracy (also for angular coefficients), but limited logarithmic accuracy for W/Z p_T modeling from POWHEG emissions and shower

- σ^{U+L} is corrected double (triple) differentially for W (Z) production using resummed SCETLIB prediction matched to fixed order DYTurbo prediction $(N^3LL + NNLO$ for nominal predictions)
- Angular coefficients are left as-is (validated against MCFM and DYTurbo fixed order predictions)*

$$
\frac{d^5 \sigma}{dq_T^2 dy dm d \cos \theta d\phi} = \frac{3}{16\pi} \frac{d^3 \sigma^{U+L}}{dq_T^2 dy dm} [(1 + \cos^2 \theta) + \frac{1}{2} A_0 (1 - 3 \cos^2 \theta) + A_1 \sin 2\theta \cos \phi + \frac{1}{2} A_2 \sin^2 \theta \cos 2\phi + A_3 \sin \theta \cos \phi + A_4 \cos \theta + A_5 \sin^2 \theta \sin 2\phi + A_6 \sin 2\theta \sin \phi + A_7 \sin \theta \sin \phi]
$$

Boson p_T Modeling Uncertainties

- Non-perturbative: Intrinsic momentum of partons (TMD PDF), non-perturbative uncertainties in resummation
- **Resummation** (perturbative): "Theory Nuisance Parameters" corresponding to coefficients in resummed calculation
- Matching: Variation in matching scale \bullet
- **Fixed order:** Missing higher orders in α_s assessed through μ_r, μ_f variations

Boson p_T Modeling Uncertainties: Non-perturbative effects

- Empirical model inspired by TMD PDFs: ∼Gaussian smearing of \bullet parton momentum, with additional freedom to account for possible x and flavour dependence
- The associated parameters cannot be predicted a priori, but must be determined from data (or lattice calculations)
- \bullet Initial values are somewhat arbitrary, with large uncertainties applied \rightarrow intended to be constrained from data

Boson p_T Modeling Uncertainties: Non-perturbative effects

- CS kernel is related to matching of non-perturbative model to resummation and is "universal" (fully correlated between W and Z)
- The rest of the NP model is taken as decorrelated between W^+ , W^- and Z, and with an additional rapidity-dependent term for the degree of smearing to account for possible x and flavour dependence

Boson p_T Modeling Uncertainties: Resummation

- Use "Theory Nuisances Parameters" corresponding to the terms appearing in the resummed calculation
- **In contrast to scale variations, this provides a well defined correlation** model across phase space (and between W and Z) and therefore better suited to profiling (see e.g. talk from F. Tackmann [here](https://indico.fis.ucm.es/event/20/contributions/529/attachments/342/600/2024-04_16_SCET_TNPs.pdf))
- **Propagating the uncertainty in this way facilitates constraining the theory** from W data alone, but also makes the correlation model between W and Z more robust for a simultaneous fit/tuning

Boson p_T Modeling Uncertainties: Heavy Quark Mass **Effects**

- **Impact of heavy quark mass effects at least partly evaluated by varying** charm and bottom thresholds in MSHT20 PDF set
- Contribution to uncertainty on m_W : 0.6 MeV
- \bullet Somewhat different effects on W vs Z \rightarrow More delicate for combined $W+Z$ fit

Electroweak Uncertainties

- **O** Most important electroweak effect is from QED FSR, included in nominal MC prediction through PHOTOS
	- Includes higher order corrections and pair production
- **•** Residual uncertainties for QED FSR (and ISR) very small, $<$ 0.5MeV contribution for m_W
- **C** Largest electroweak uncertainty from virtual corrections, ∼ 2MeV on m_W

Charge Difference with Helicity Fit

Charge difference also very similar between nominal and helicity fit, and stable under changes in prefit uncertainties for the helicity cross sections

Validation of boson p_T modeling with $Z \rightarrow \mu\mu$

- Fit theory model to dilepton p_T spectrum directly to validate that it can describe the data
- \bullet O(10%) level discrepancy due to untuned non-perturbative parameters at low p_T fully reabsorbed
- \bullet Postfit description of the spectrum at 0.1% level

Validation of boson p_T modeling with W-like $Z \rightarrow \mu\mu$

• When running the full W-like fit to single muon (η, p_{τ}) , charge) the theory model is also able to accommodate the muon p_T distribution very precisely

Validation of boson p_T modeling with $Z \rightarrow \mu\mu$

- **O** Detector level fit results can be propagated to predictions for unfolded Z p_T spectrum
	- For both direct fit to $p_T^{\mu\mu}$ and W-like fit to single muon $(\eta, p_{\tau}, \text{charge})$
- **•** Strong and consistent constraints from **both** fits, and in agreement with unfolded data
- Direct fit to $p_T^{\mu\mu}$ has stronger constraints but W-like fit is able to correctly disentangle m_Z from the Z p_T spectrum
- \bullet m_W can be measured without tuning the p_T spectrum to the Z

W-like m_z result: Validation checks

- **O** Consistent results when extracting 48 independent m_Z parameters split in charge and 24 η bins
- \bullet η -sign difference: $m_Z^{\eta > 0} - m_Z^{\eta < 0} = 35 \pm 20$ MeV
- **O** Charge difference: $m_Z^+ - m_Z^- = 31 \pm 32$ MeV
- **•** Charge difference with reversed even-odd event selection:

 $m_Z^+ - m_Z^- = 6 \pm 32$ MeV