Measurement of the W boson mass in proton proton collisions at $\sqrt{s}$ = 13 TeV

The CMS Collaboration

Krzysztof Doroba
Physics Dept.
University of Warsaw

Based on seminar by Josh Bendavid CERN, Oct 17 2024

15 Nov 2024 ZCiOF seminar

A high precision measurement of the W mass at CMS

Josh Bendavid (MIT)
on behalf of the CMS Collaboration

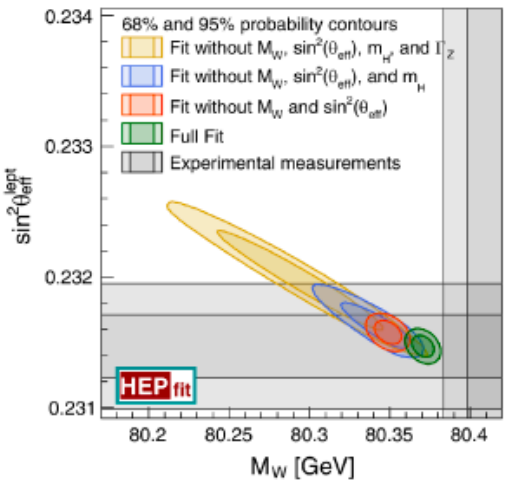
Sept. 17, 2024

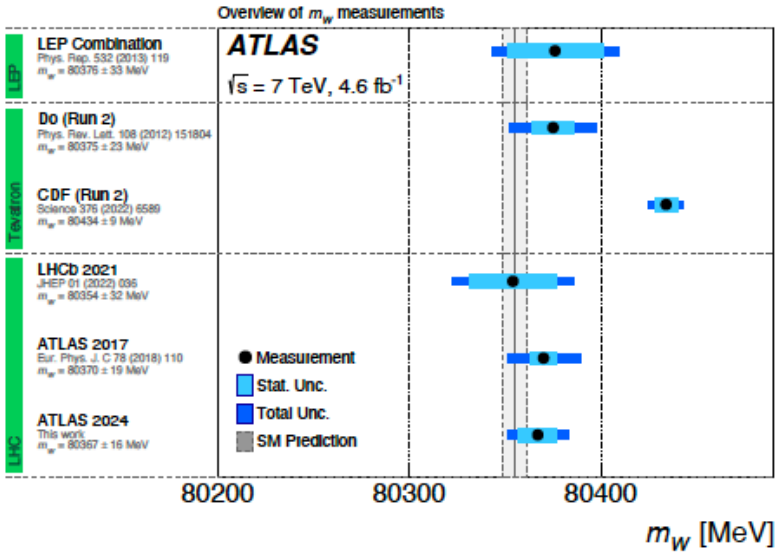
Outline of the talk

- Motivation
- Possible Methods
- Data
- Main Problem – Accuracy
- Most Important Contributions to Measurement Error
- Measurement (maximum likelihood template fit)
- The Result
- Summary

Introduction

Phys.Rev.Lett. 129 (2022) 27, 271801





WHY TO MAESURE m_W ?

- The discovery of the Higgs and the precise measurement of its mass provides the complete set of inputs needed to overconstrain the Standard Model
- Recent CDF measurement in significant tension with SM prediction and other measurements

HOW TO MEASURE m_W?

W→ $l\nu$ ($q\bar{q}$ $-huge\ QCD\ background, jet\ energy\ scale$)

$W \rightarrow \mu\nu$ is the choice

but W can not be fully reconstructed because of neutrino ☹

Transverse mass

$$m_W^T = \sqrt{2p_T^{\mu} p_T^{miss}(1 - cos\Delta\phi_{l\nu})}$$

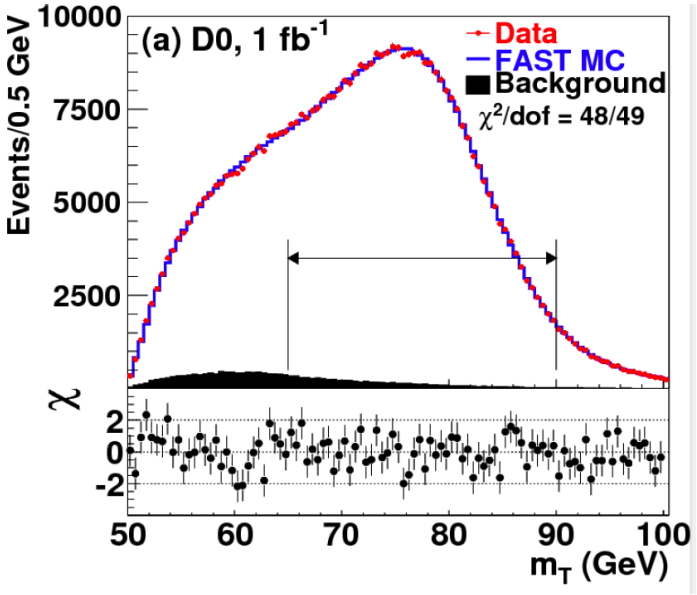


Figure 3.10: An example of transverse mass Jacobian peak measured at Tevatron.

p_T^{miss} is a problem at LHC because of a large number of pp collisions

p_T^{μ} distribution and the Jacobian peak is the solution !!!

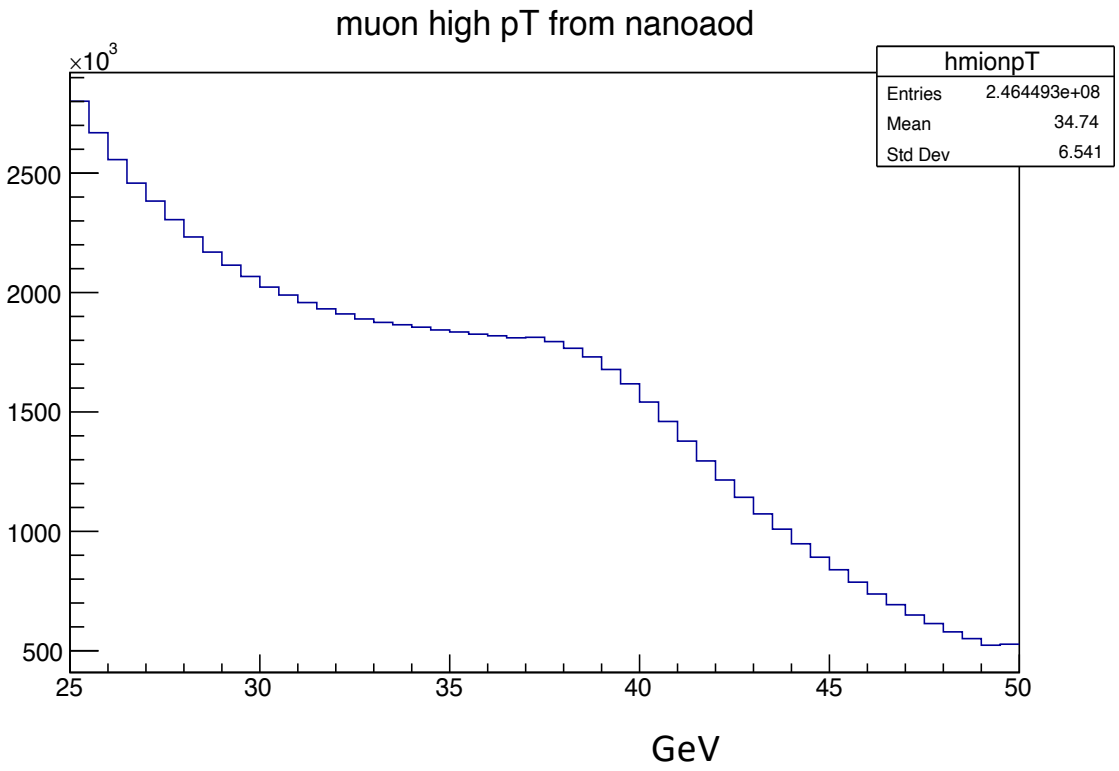
Jacobian peak – in colliding $q\bar{q}$ CM frame

E>>m_μ, $E_\mu = E_\nu = 2p^\mu = m_W$

$$\frac{dN}{dp_T^\mu} = \frac{dN}{d\theta^*}\frac{d\theta^*}{dp_T^\mu}$$

$$p_L^\mu = p^\mu cos\theta^* = \frac{dp_T^\mu}{d\theta^*} = \sqrt{(p^\mu)^2 - \left(p_T^\mu\right)^2} = \sqrt{\left(\frac{m_W}{2}\right)^2 - \left(p_T^\mu\right)^2}$$

$$\frac{dN}{dp_T^\mu} = \frac{1}{\sqrt{\left(\frac{m_W}{2}\right)^2 - \left(p_T^\mu\right)^2}}\frac{dN}{d\theta^*}$$



CMS measurement of m_W is based on a fit to the measured $\left(p_T^{\mu}, \eta^{\mu}, q^{\mu}\right)$ distributions using simulated templates of $W \to \mu\nu$ signal and most background processes

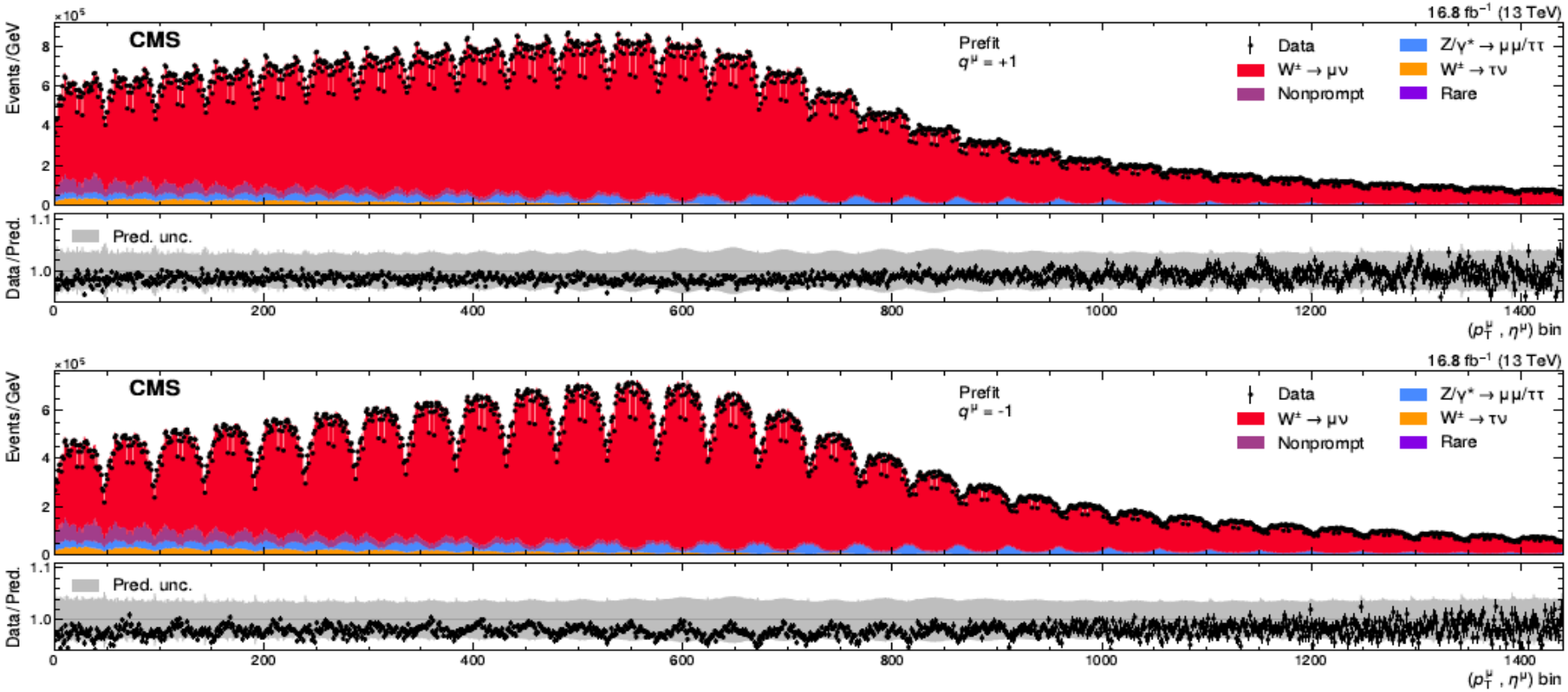


Figure A.16: Measured and predicted $(p_{\mathrm{T}}^{\mu}, \eta^{\mu})$ distributions for positively (upper) and negatively (lower) charged muons. The two-dimensional distribution is "unrolled" such that each bin on the x-axis represents one $(p_{\mathrm{T}}^{\mu}, \eta^{\mu})$ cell. The gray band represents the uncertainty in the prediction, before the fit to the data.

The real challenge for this m_W measurement is its accuracy. To achieve it CMS:

- Use 2016 well understood CMS 13 TeV pp data : 16.8 fm^{-1} (~25 interactions per crossing)
- Measure p_T^μ distribution with highest precision possible
- Prepare with Monte Carlo the most precise and credible templates for η^μ, p_T^μ distributions
- Use 4B fully simulated events, > 100M selected W events, special computing
- Extracted m_W from profile likelihood fit to muon $(\eta^\mu, p_T^\mu, charge)$ distributions (48 η bins [-2.4,2.4] and 30 p_T bins 1 GeV wide) with precise and credible error determination !!!

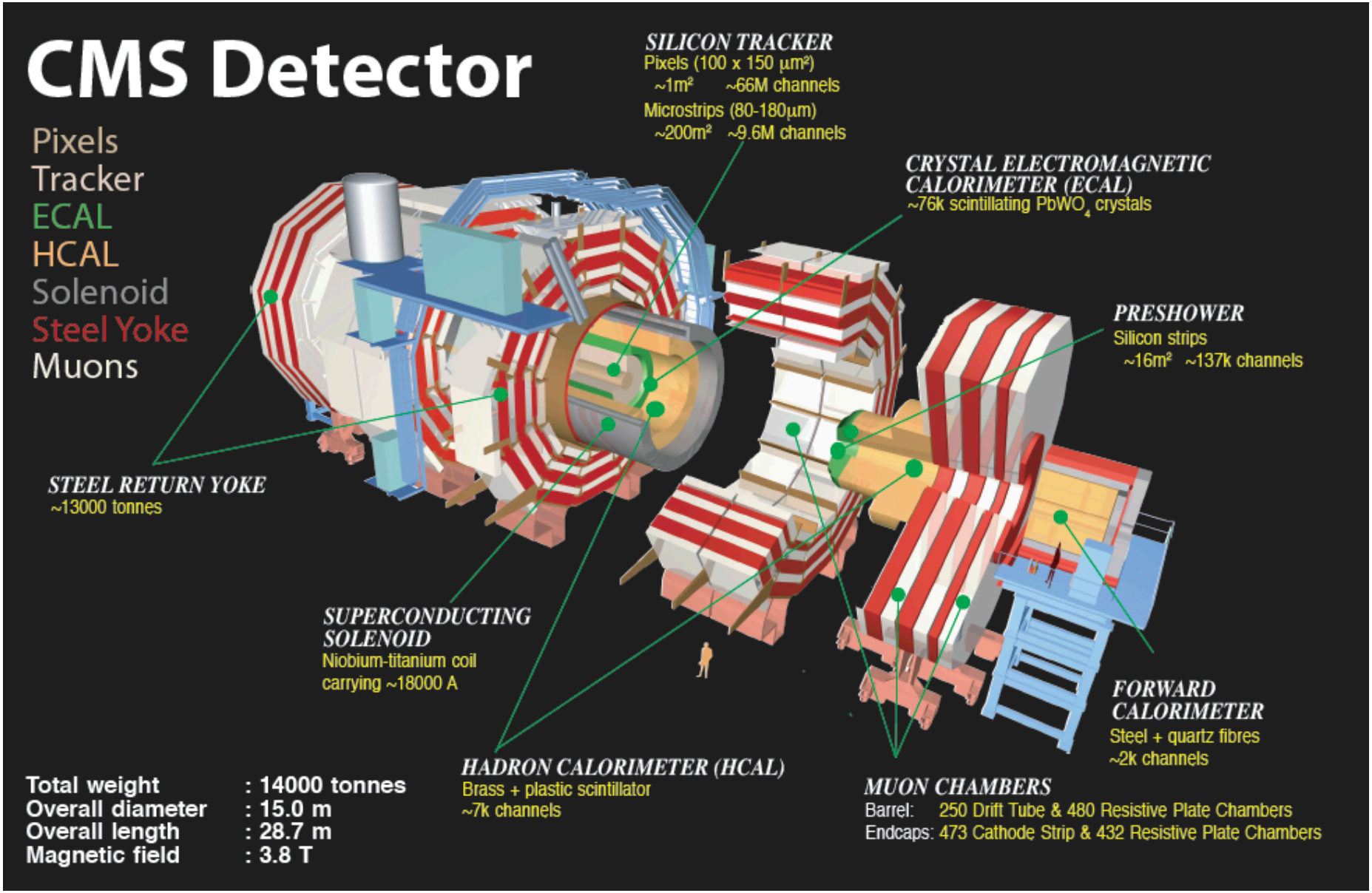
Which means special care was paid to following subjects :

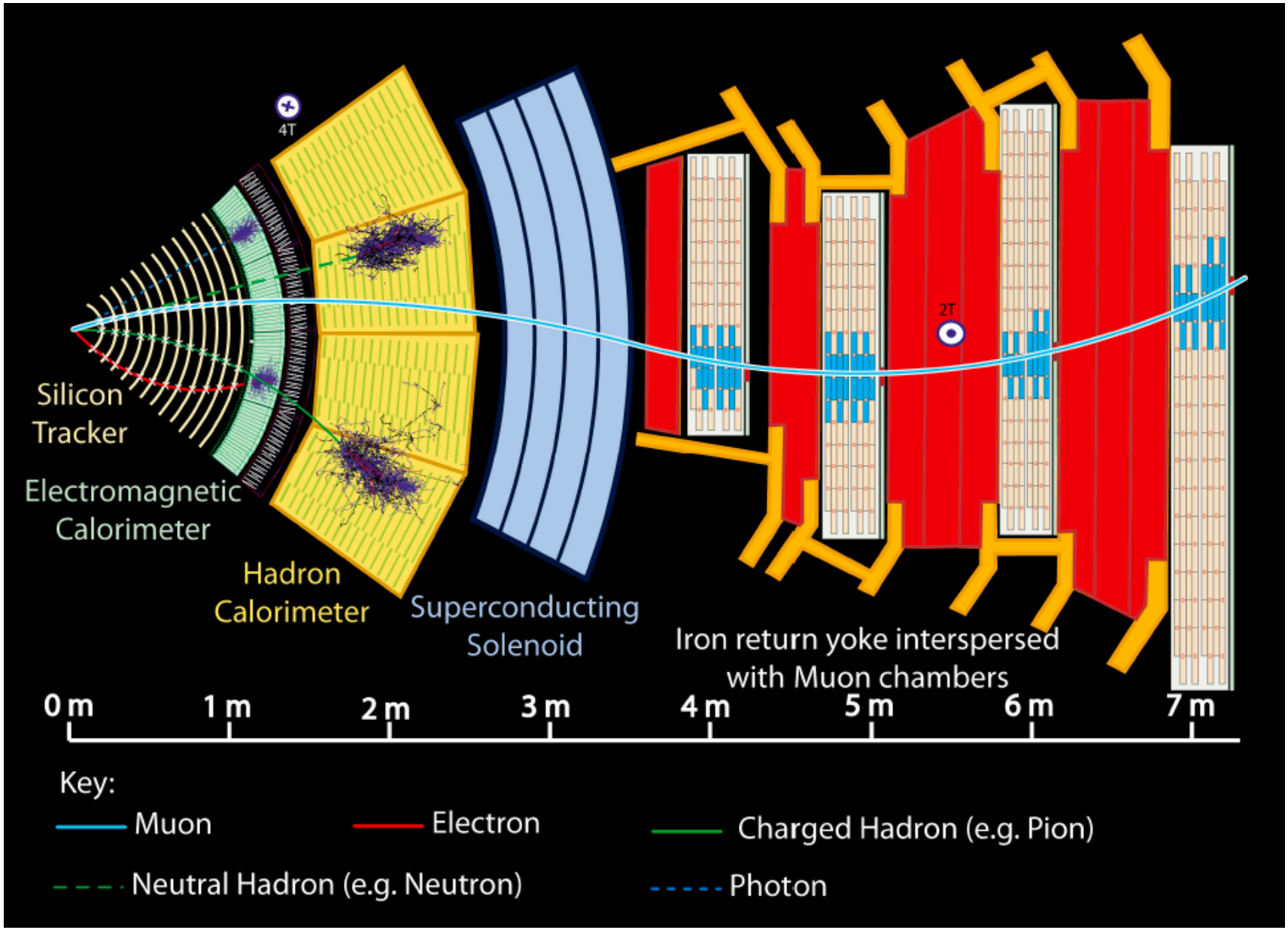
- Event samples and selection criteria
- Efficiency corrections
- Hadronic recoil calibration
- Non-prompt background determinations
- Muon momentum calibration
- Modeling of the W and Z transverse momentum distributions
- Modeling of the angular distribution in W and Z leptonic decays
- Parton distributions functions
- Impact of missing higher order EW corrections
- Additional validations of theoretical modeling ./.

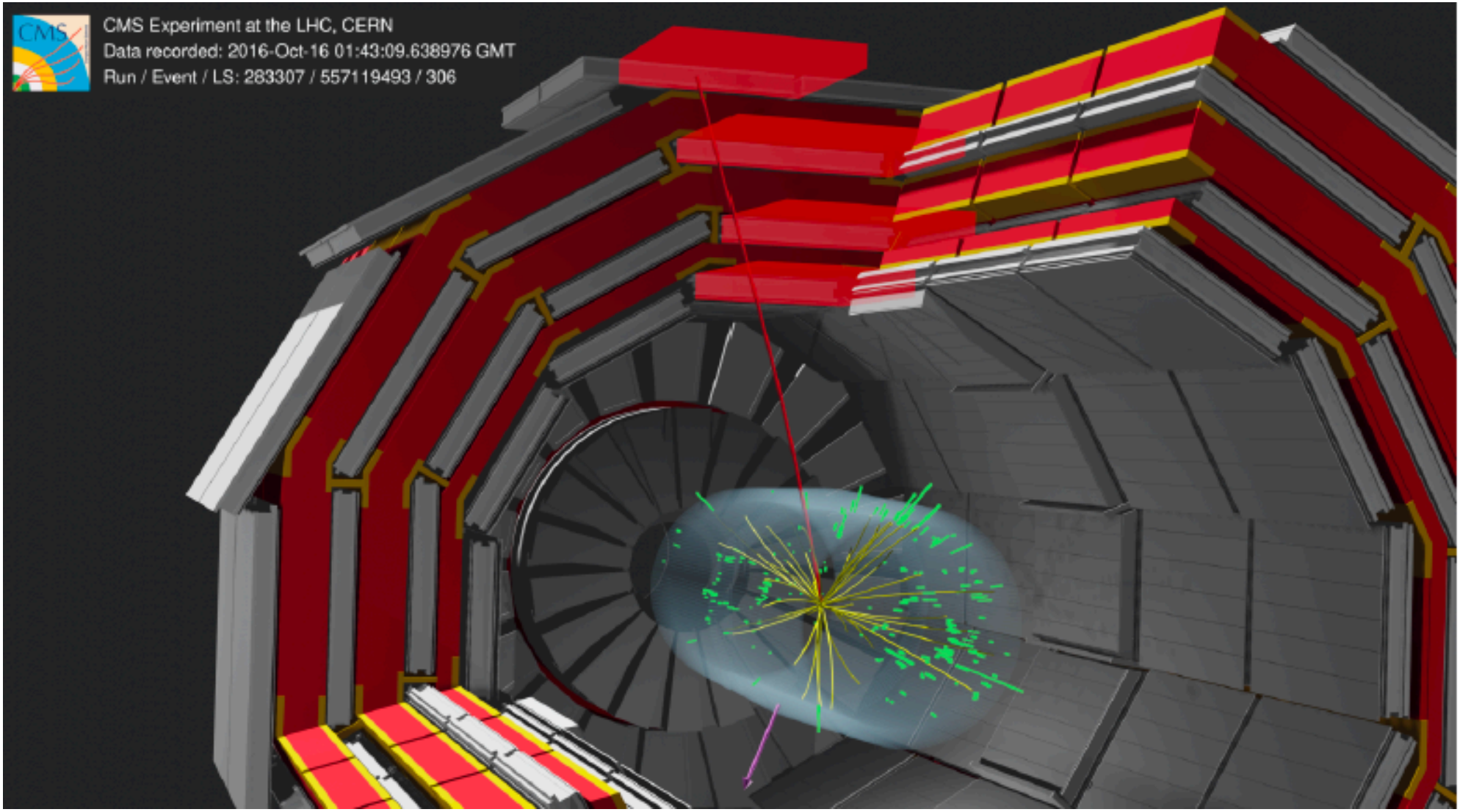
- Helicity fit
- $W-like\ Z$ and W mass measurements
- Measurements of $m_{W^+} - m_{W^-}$
- Results with alternative parton distribution functions

All this was done with:

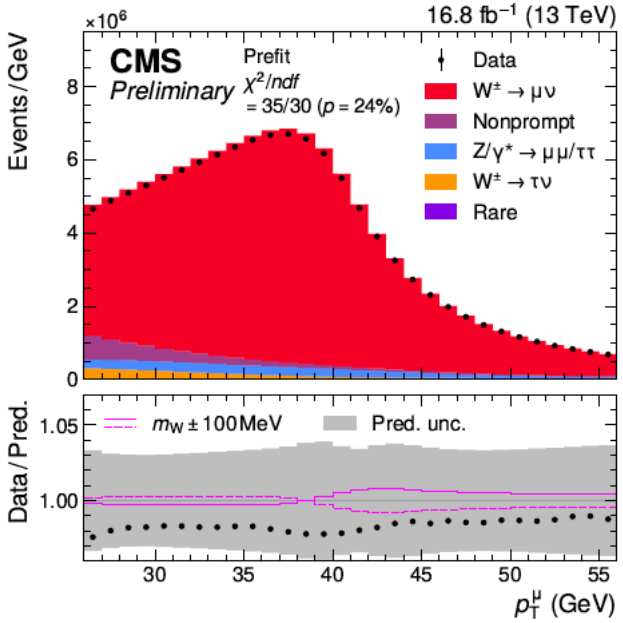
- Sufficient computer power
  - no CERN GRID, two devoted computers (MiT+Italy?)
  - TENSORFLOW software package
    to make the m_W and $W-like\, m_Z$ fits computationally feasible and numerically stable with more then 2000 bins and 4000 nuisance parameters
  - …….. (See CERN Data Science Seminar by David Walter on 16.10.2024)


CMS Detector
Pixels
Tracker
ECAL
HCAL
Solenoid
Steel Yoke
Muons
SILICON TRACKER
Pixels (100 x 150 µm²)
~1m² ~66M channels
Microstrips (80-180µm)
~200m² ~9.6M channels
CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
~76k scintillating PbWO4 crystals
PRESHOWER
Silicon strips
~16m² ~137k channels
STEEL RETURN YOKE
~13000 tonnes
SUPERCONDUCTING SOLENOID
Niobium-titanium coil
carrying ~18000 A
FORWARD CALORIMETER
Steel + quartz fibres
~2k channels
HADRON CALORIMETER (HCAL)
Brass + plastic scintillator
~7k channels
MUON CHAMBERS
Barrel: 250 Drift Tube & 480 Resistive Plate Chambers
Endcaps: 473 Cathode Strip & 432 Resistive Plate Chambers
Total weight : 14000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T


4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon


CMS Experiment at the LHC, CERN
Data recorded: 2016-Oct-16 01:43:09.638976 GMT
Run / Event / LS: 283307 / 557119493 / 306

Event selection



- Straightforward single muon selection: track quality criteria, loose transverse impact parameter cut, and isolation
- Selected events are about 90% $W \to \mu\nu$
- Nonprompt background from data-driven estimate
  - Mostly from B and D decays with smaller contribution from π or K decay-in-flight
- Prompt backgrounds from simulation with all relevant corrections/uncertainties
  - $W \to \tau\nu$, $Z \to \mu\mu$ (mostly with one muon out-of-acceptance), $Z \to \tau\tau$, top, diboson

This is

BLINDFOLD ANALYSIS

- First determine what the errors will be
- Then extract m_W with the fit to the data

So look at the sources of errors
and try to find the way to diminish them

Hadronic recoil calibration

How to improve modeling of p_T^{miss} ?

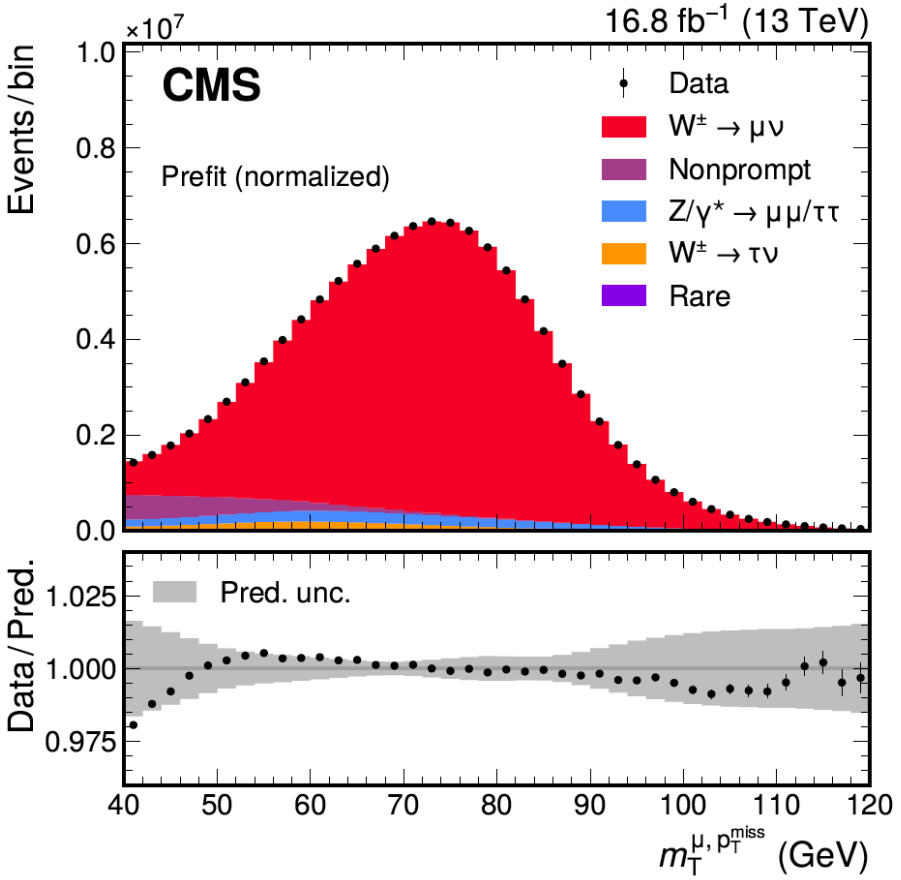
- With DEPMET (ML)

p_T^{μ} measurement and m_T distribution?

- define

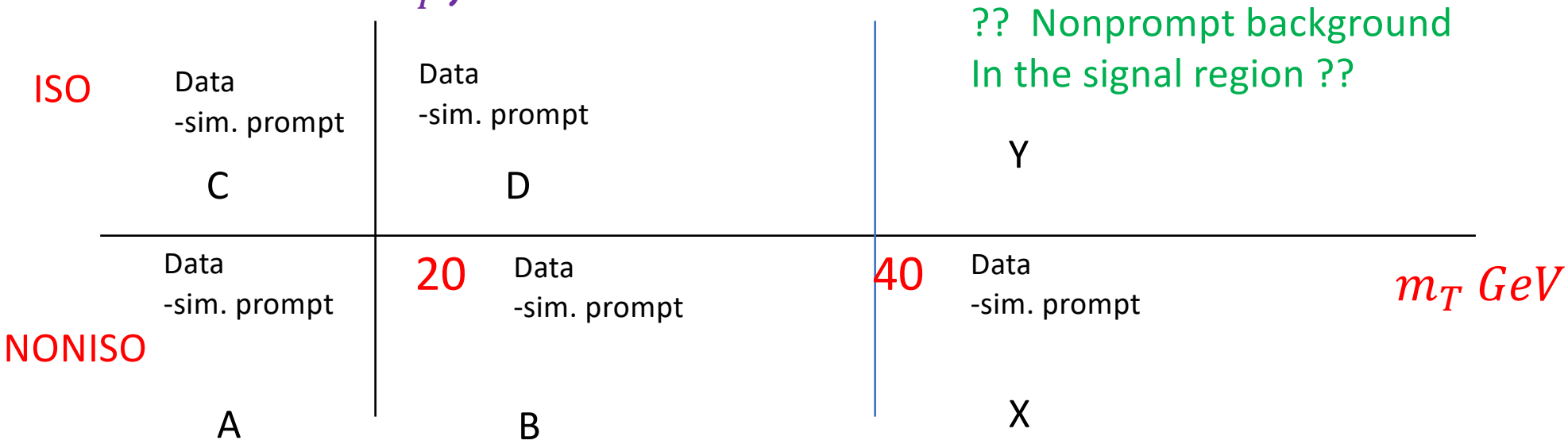
  $\perp and \parallel$ components of hadronic recoil parametrized as a function of the reconstructed $p_T^{\mu\mu}$ (both for reconstructed and simulated events)

- Use measured $Z \to \mu\mu$ events
- Map simulation to the Z data, derive the corrections
- Apply these corrections to $W \to \mu\nu$ events



Non-prompt background determination

- Main source in-flight decays of heavy flavor hadrons
- Evaluation – with data from sideband regions (in μ isolation and inverse m_T)



- Y is defined by ratios of A,B,C,D,X in each bin of p_T^{μ}, η^{μ} and $charge$

- Nonprompt contribution in each sideband is smoothed in p_T^{μ} with exponential of third order Chebyshev polynomial. (polynomial coefficients are used in error determination.)
- Validate above procedure with simulated non-prompt events in W region. Agree with ABCD result within 0.8 factor
- Finally validate ABCD with events enriched with non-prompt from secondary vertex. 2% agreement between data and simulations.

- Uncertainty in m_W from non-prompt background estimated 3.2 MeV

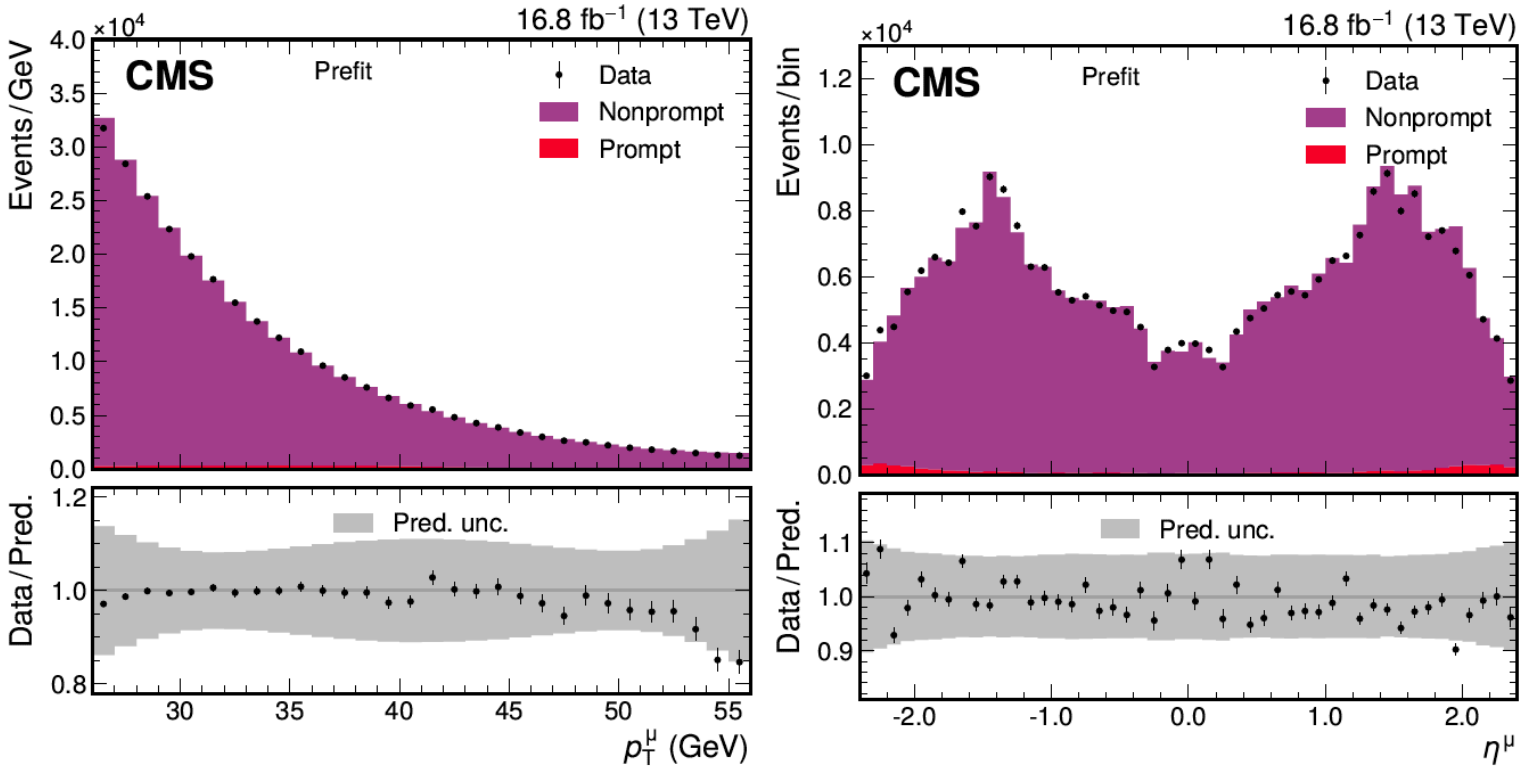
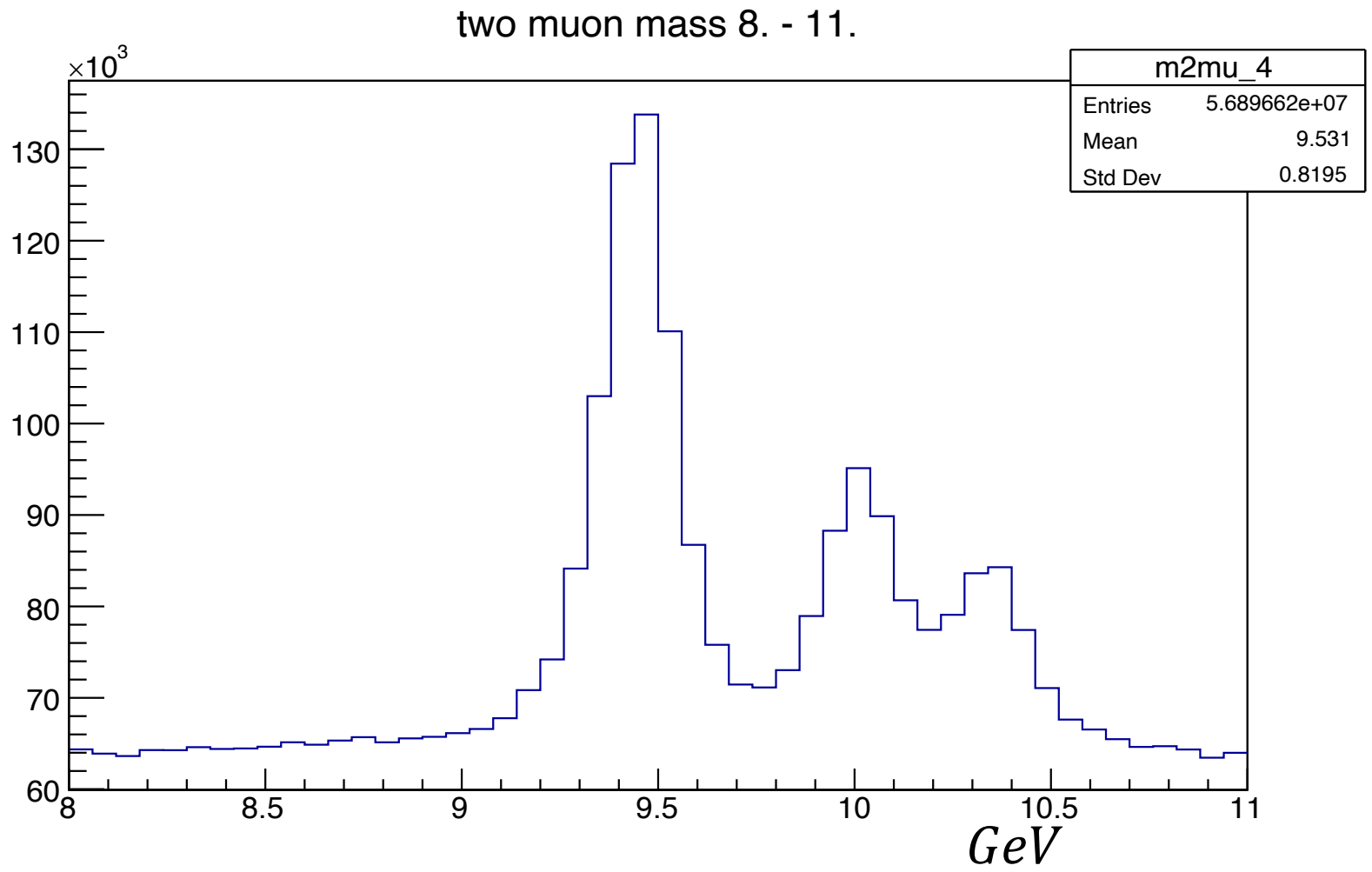


Figure A.3: The observed data and the prediction of the extended ABCD method before the maximum likelihood fit, for the p_T^μ (left) and η^μ (right) distributions, in a region enriched in events with nonprompt muons obtained by selecting muons compatible with being produced in a secondary vertex. Small contributions from events with a prompt lepton, evaluated using simulated samples, are shown by the red histogram. The total uncertainties (statistical and systematic) are represented by the gray bands.

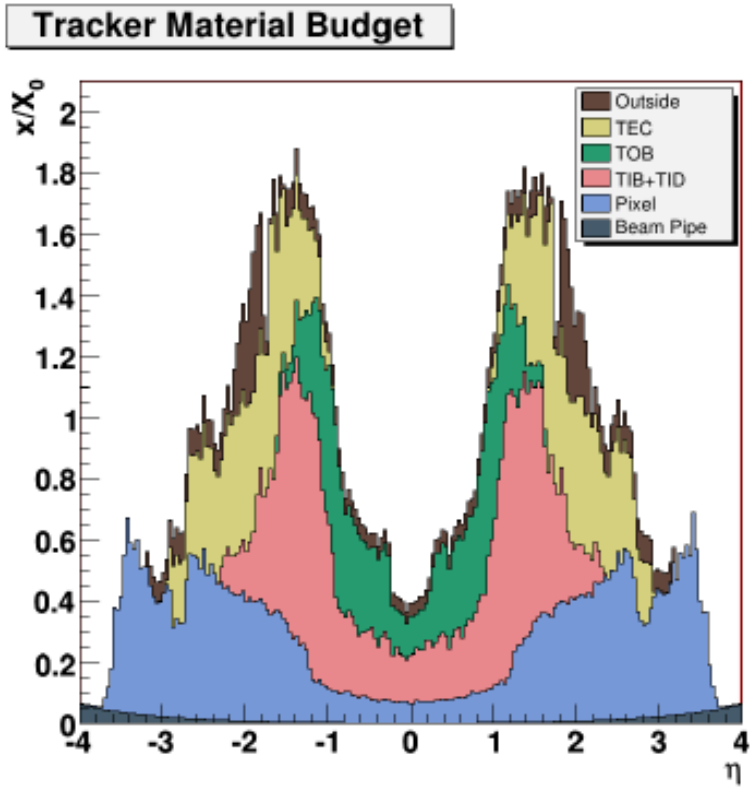
Muon momentum calibration

Standard CMS reconstruction
→
as on nanoAOD



Muon momentum measurement

- Muon must have isolated, reconstructed track in tracker and muon detectors
- Muon momentum inferred from the track curvature.
- For $m_W$measurement track curvature is determined by silicon pixel and strip detectors
- Muon system used only for triggering and identification
- Strategy: calibrate with quarkonia, validate with Z^0 mass.



Muon momentum calibration

Alignment part:

1. standard pattern recognition + Kalman filter fit
2. Refit with Continuous Variable Helix
   - Continuous energy loss
   - Multiple scattering from finite material elements
   - Material model + GEANT4 propagator
   - Measured magnetic field map (on the ground level)

+Calibration part:

3. Generalized global correction procedure
   - Further small corrections to above
   - 6 parameters per tracker module, fit to sample of $J/\psi \rightarrow \mu\mu$

Muon momentum calibration/validation

4. Residual corrections
- include "weak modes" (subdetector skewness)
- include non Gaussian J/ψ shape

fits to $J/\psi \to \mu\mu$ dimuon mass distribution binned in 4-dim space + - p_T^μ η^μ

→corrections to p_T^μ in η^μ bins

5. Validation
- use scale from global correction as for $J/\psi \to \mu\mu$, then

fit to $Z \to \mu\mu$ and apply the corrections to simulated muons

Validation of muon momentum calibration

$$k = {}^{1}/_{p_T} \qquad \frac{\delta k}{k} = A_{i\eta} - \epsilon_{i\eta} k + q\frac{M_{i\eta}}{k}$$

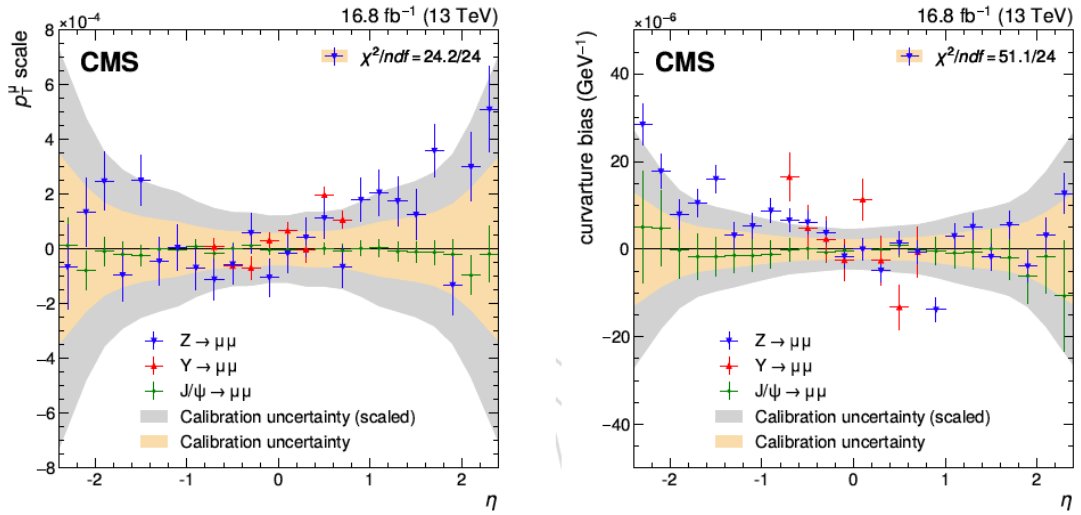


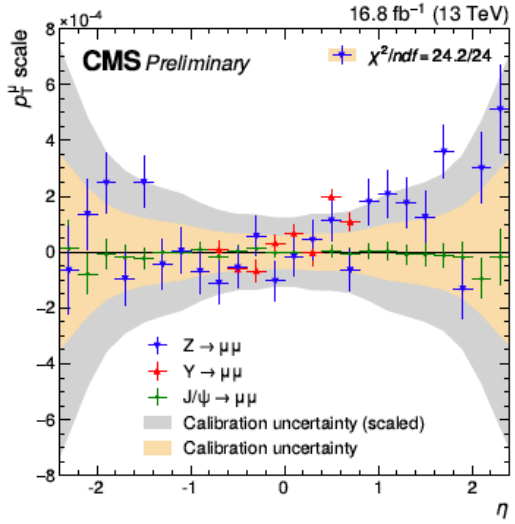
Figure A.4: Charge-independent (left) and charge-dependent (right) closure results from using J/ψ, Υ(1S), and Z events. The charge-independent closure plot shows an equivalent magnetic field scale factor, while the charge-dependent closure plot shows an equivalent misalignment term. The points with error bars represent the scale and statistical uncertainty associated with the closure test, while the yellow band represents the corresponding statistical uncertainty in the calibration parameters themselves, from the J/ψ calibration sample. The filled gray band shows the scaled statistical uncertainty as described in the text. The calibration uncertainties are fully uncorrelated from the Z and Υ(1S) closure uncertainties, but very strongly correlated with the J/ψ closure uncertainties, since they use the same data.

$A_{i\eta}$ - correction to magnetic field

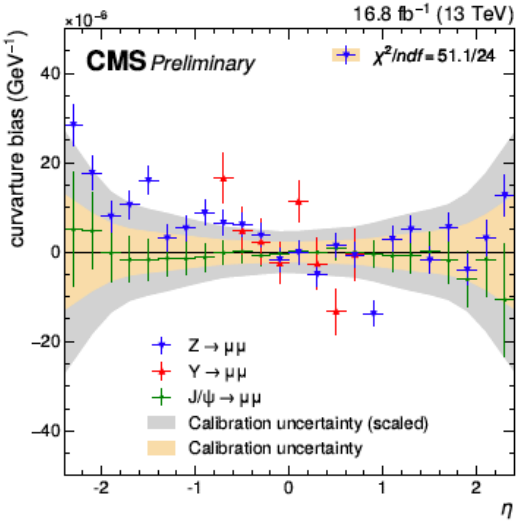
$\epsilon_{i\eta}$ - energy loss

$M_{i\eta}$ - alignment

charge-independent



charge-dependent



- Calibration is validated with $\Upsilon_{1S} \to \mu\mu$ and $Z \to \mu\mu$ in terms of B-field and alignment-like residual parameters
- B-field-like term for Z is consistent with zero within statistical uncertainties, alignment-like almost so
- Statistical uncertainty on calibration parameters from J/ψ scaled by 2.1 to cover all possible correlated patterns of bias across the detector from any not-explicitly-accounted-for systematic effects

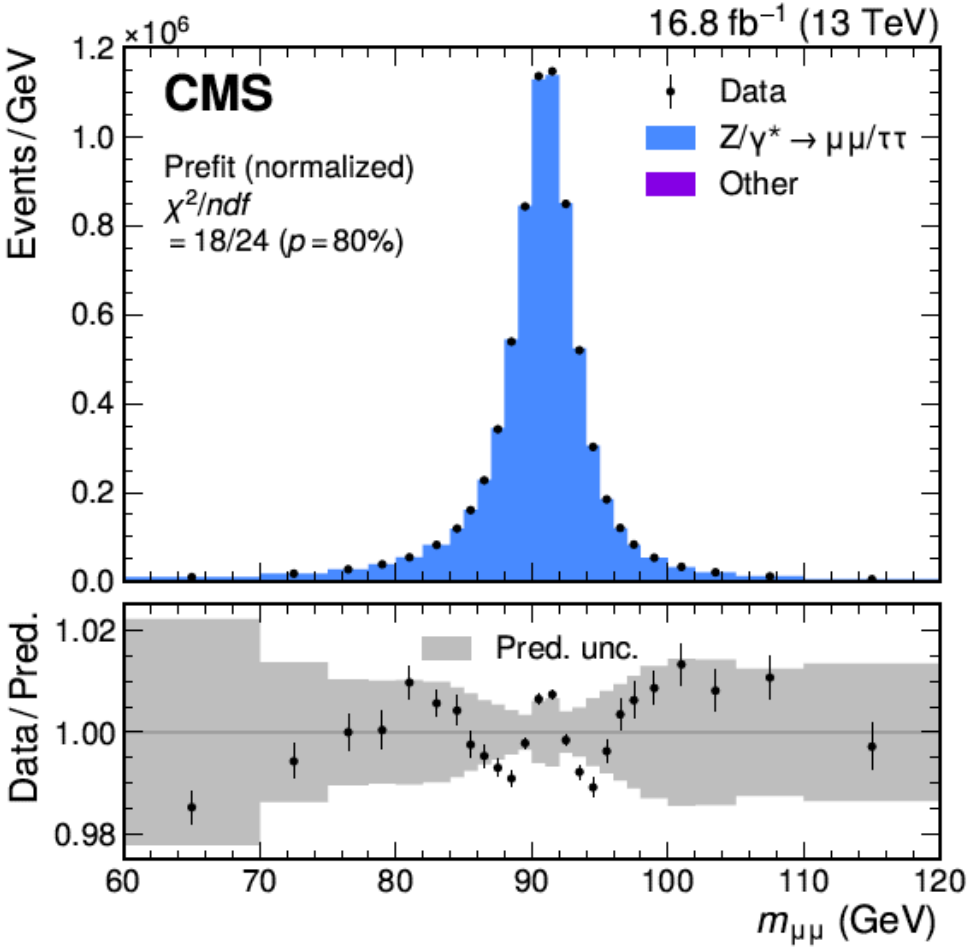


Table A.1: Breakdown of muon calibration uncertainties.

Source of uncertainty	Nuisance parameters	Uncertainty in m_W (MeV)
J/ψ calibration stat. (scaled ×2.1)	144	3.7
Z closure stat.	48	1.0
Z closure (LEP measurement)	1	1.7
Resolution stat. (scaled ×10)	72	1.4
Pixel multiplicity	49	0.7
Total	314	4.8

Simulation framework for p_T^V spectra:

- MiNNLO$_{PS}$ –state of art calculations in QCD including resummation of log enhanced contributions at small p_T^V and model for nonperturbative effects there
- SCETLIB – performs p_T^V resummation using soft-collinear effective theory
- DY-TURBO – Drell-Yan calculations NNLL combined with NNLO

SCELTLIB matched with DY-TURBO allows for N^{3}LL+NNLO accuracy

Result: statistical power of the MiNNLO$_{PS}$ while improving its accuracy at small p_T^V

Simulation framework for p_T^V spectra:

Calculations with other libraries:
DYTURBO, MATRIX+RadISH and CuTe-MCFM
predict shifts in m_W within SCETLIB+DYTUTBO uncertainties

REMARK !!!

CMS is the first experiment which for m_W measurement deduces p_T^W spectra from data validated p_T^W MC rather then from measured p_T^Z spectrum

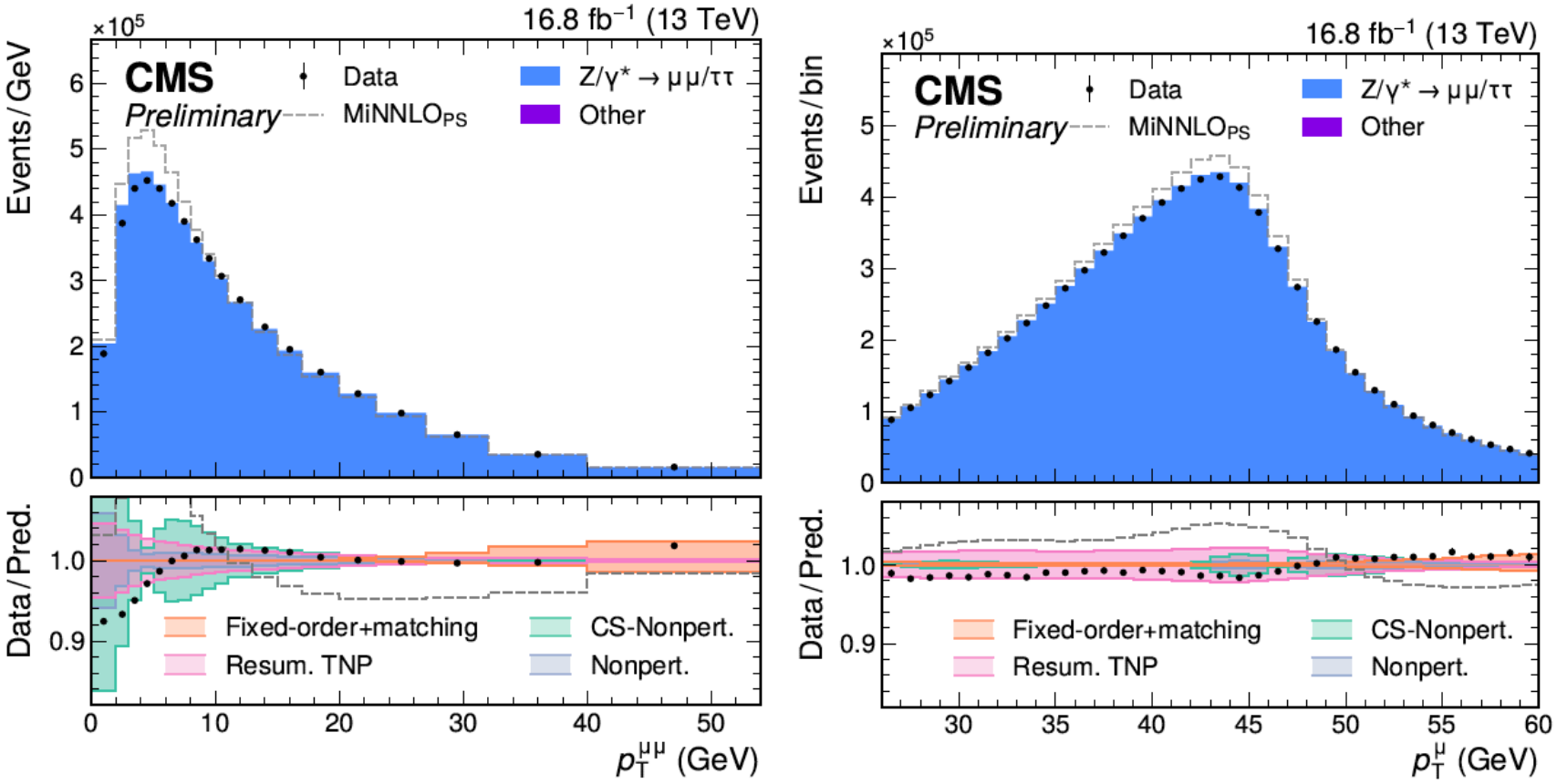
Estimation of the theory uncertainties

But calculations are approximative and one has to determine their uncertainty

- For fixed order calculations – changing SCENTLIB+DYTURBO parameters (μ_R, μ_Fscales) (correlation not included)
- For perturbative uncertainties in resummed predictions – with Theory Nuisance Parameters (TNS) (correlations included)
- ……..

TNS exploits known all order perturbative structure of resummed calculations

Estimation of the theory uncertainties



Nonperturbative and resummation uncertainties have the largest impact on p_T^V distribution in the Jacobian peak region, most sensitive one to m_W value

What We Should be Doing.

$\mathrm{N}^{1+1}\mathrm{LO}$: $f^{\text{predicted}}(\alpha) = f_0 + f_1\,\alpha + f_2(\theta_2)\,\alpha^2$

Parametrize the (leading) source of uncertainty $f_2 \equiv f_2(\theta_2)$

- In terms of unknown but well-defined parameters θ_n, which are the *theory nuisance parameters (TNPs)*
  - Simplest: Use f_2 itself: $f_2(\theta_2) \equiv \theta_2$
  - Better: Account for known internal structure of f_2 (color, partonic channels, ...)
- Include the full parametrized next term in the prediction $+\,f_2(\theta_2)\,\alpha^2$
  - Sufficient to include the next term, we always assume that expansion converges, so f_3 is not yet relevant
    ($\rightarrow$ We assume a renormalon-free series)
  - When f_2 becomes known (or tightly constrained), need to include f_3

    $\mathrm{N}^{2+1}\mathrm{LO}$: $f^{\text{predicted}}(\alpha) = f_0 + f_1\,\alpha + f_2\,\alpha^2 + f_3(\theta_3)\,\alpha^3$

Finally: Vary all θ_i to account for correctly correlated theory uncertainty

Theory Uncertainties via TNPs.

As before, except that now prediction depents on θ_i (ignoring exp. NPs now)

ML fits: $$L(y, \theta_i) = P(d|y, \theta_i) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(u_i - \theta_i)^2}{2\sigma_i^2}\right]$$

χ^2 fits: $$\chi^2(y, \theta_i) = \sum_d \frac{[d - f_d^{\text{predicted}}(y, \theta_i)]^2}{\sigma_d^2} + \sum_i \frac{(u_i - \theta_i)^2}{\sigma_i^2}$$

Important to distinguish what we need to estimate

- We *do not* need a precise estimate of the true value of θ_i (or f_2)
  - Often our best estimate for θ_i will be $u_i = 0$
  - We can (and will) have $f_2(\theta_2 = 0) \neq 0$
  - If we obtain nontrivial information on θ_i we can include it via $u_i \neq 0$
- We *do* need to estimate σ_i of u_i (the systematic "theory uncertainty")
  - i.e., how is θ_i allowed to vary
  - For this purpose, it is sufficient to understand the *generic size* of f_2, e.g. we can consider the most dominant contributions or limits

Parton distribution functions

CD18Z, NNPDF3.1, NNPDF4.0, CT18, MSHT20, PDF4LHC21 considered

How to chose PDT ?

- For given PDF make the m_W, m_Z analysis with pseudo-data from another PDF.
- Check if obtained m_W is within uncertainty of the tested PDF

Only CT18Z, CT18 and PDF4LHC passes this test

CT18Z fits best to the data

Uncertainty from PDF in $m_w = 4.4\ MeV$

Parton distribution functions

CT18Z PDF is the choice

Uncertainty from PDF in $m_w = 4.4\ MeV$

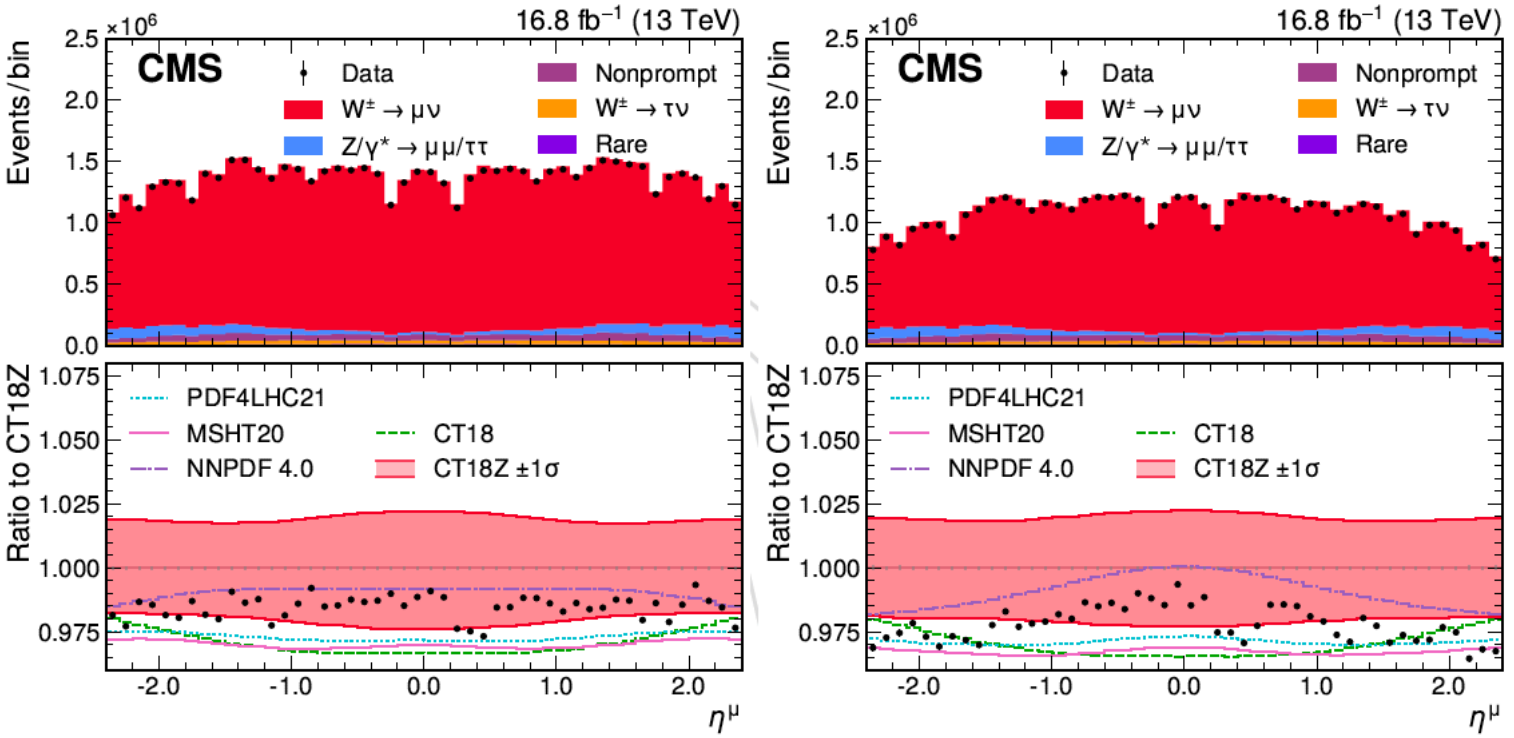


Figure A.8: Measured and predicted η^{μ} distributions for selected W^+ (left) and W^- (right) events. The nominal prediction, obtained with the CT18Z PDF set, is shown in filled light red. The uncertainty, evaluated as the sum of the eigenvector variation sets, is represented by the filled band in the lower panel. The predictions using the PDF4LHC21, MSHT20, NNPDF4.0, and CT18 sets are also shown (without uncertainty bands).

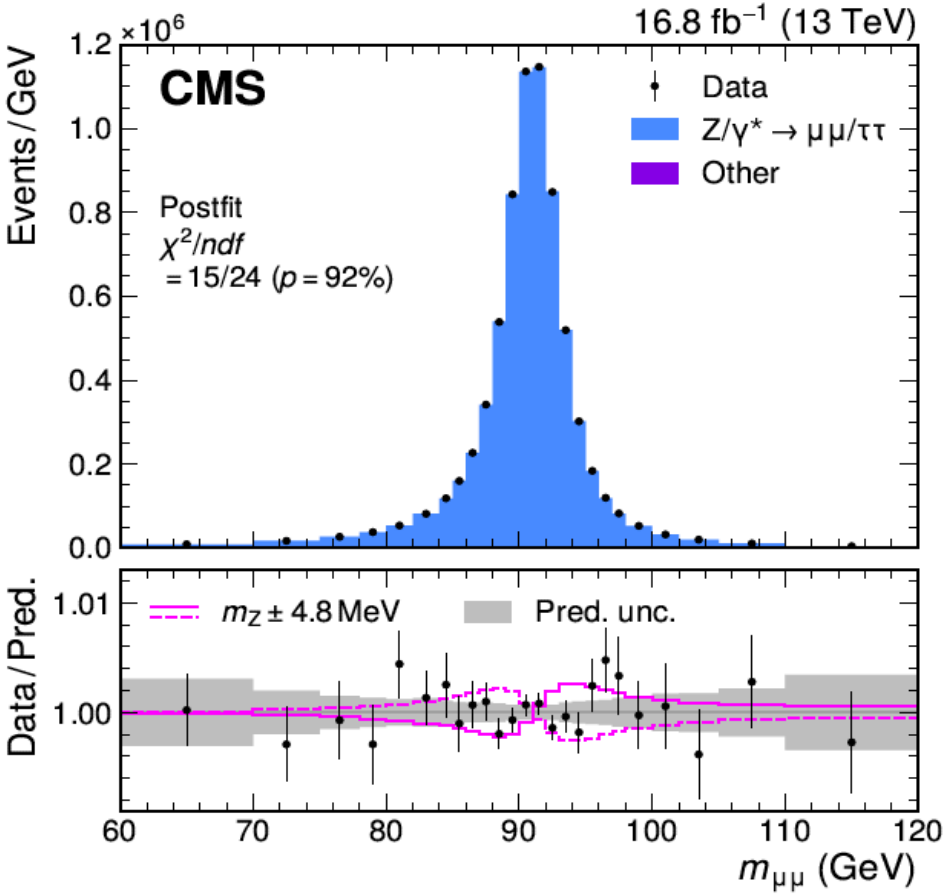
Z boson mass from dimuon mass distribution

Binned maximum likelihood fit to the dimuon mass distribution

Final validation of calibration/ uncertainties

$$m_Z^{\mu\mu} - m_Z^{PDG} = 2.2 \pm 4.8\ MeV$$

As $Z \to \mu\mu$ involved in calibration validation, this result does not qualify for word averages



$W-like$ measurement of the Z boson mass

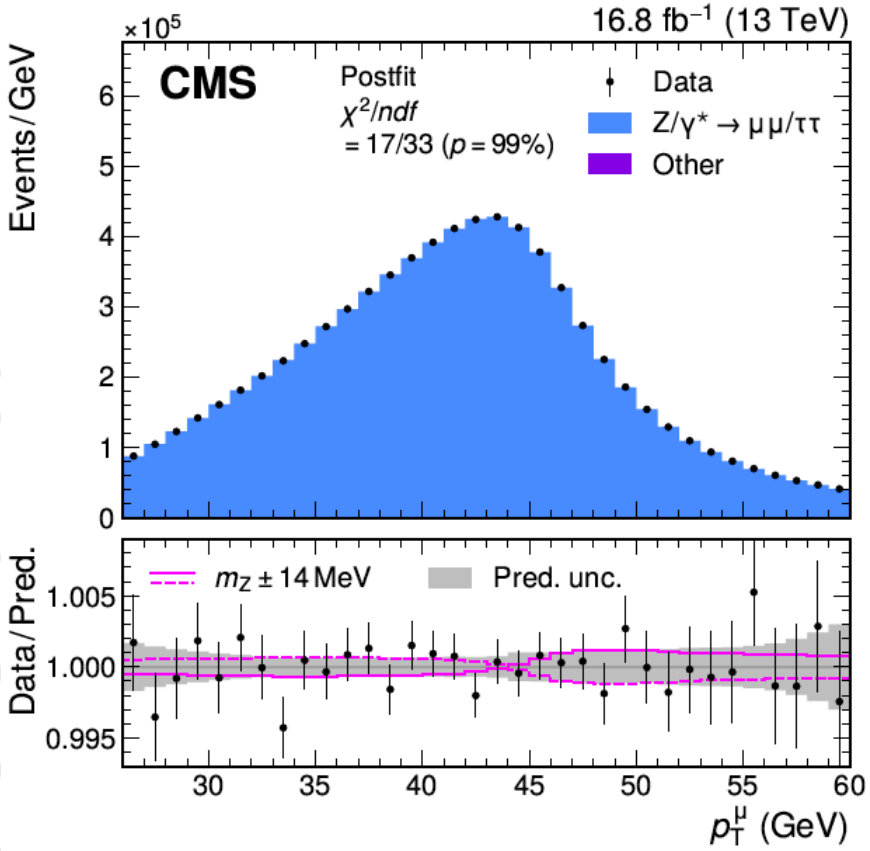
From binned maximum likelihood fit to $\left(p_T^{\mu}, \eta^{\mu}, q^{\mu\mu}\right)$ distribution of selected muons

$$m_Z^{W-like} = 91182 \pm 7(stat) \pm 12(syst) = 91\,182 \pm 14$$

$$m_Z^{PDG} = 91\,188 \pm 2.0 MeV$$

Figure A.17: Measured and simulated p_{T}^{μ} distributions, with the prediction adjusted according to the best fit values of nuisance parameters obtained from the maximum likelihood fit of the W-like m_{Z} analysis. The solid and dashed purple lines represent, respectively, an increase and decrease of m_{Z} by 9.9 MeV. The uncertainties in the predictions, after the systematic uncertainty profiling in the maximum likelihood fit, are shown by the shaded band.

14.



- For the next step : extract the best fit values of nuisance parameters from p_T^{μ} modeling

Validation of theory model

- From this plot → extract systematic uncertainty
- Verify that they are **consistent** with those ↑ describing the p_T^Z modeling from the $W-like\ m_Z$ measurements

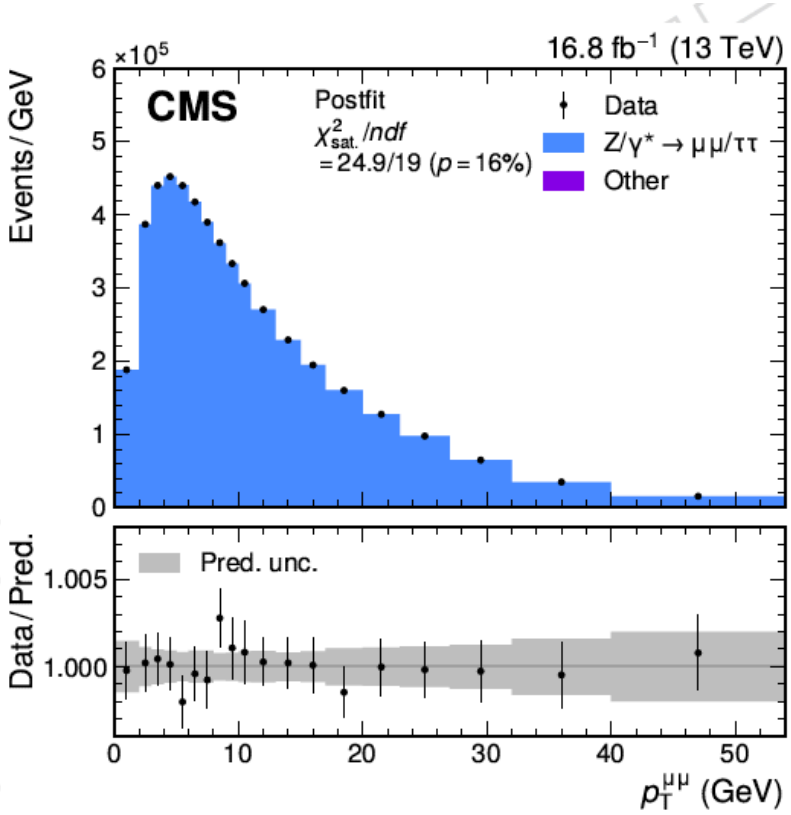


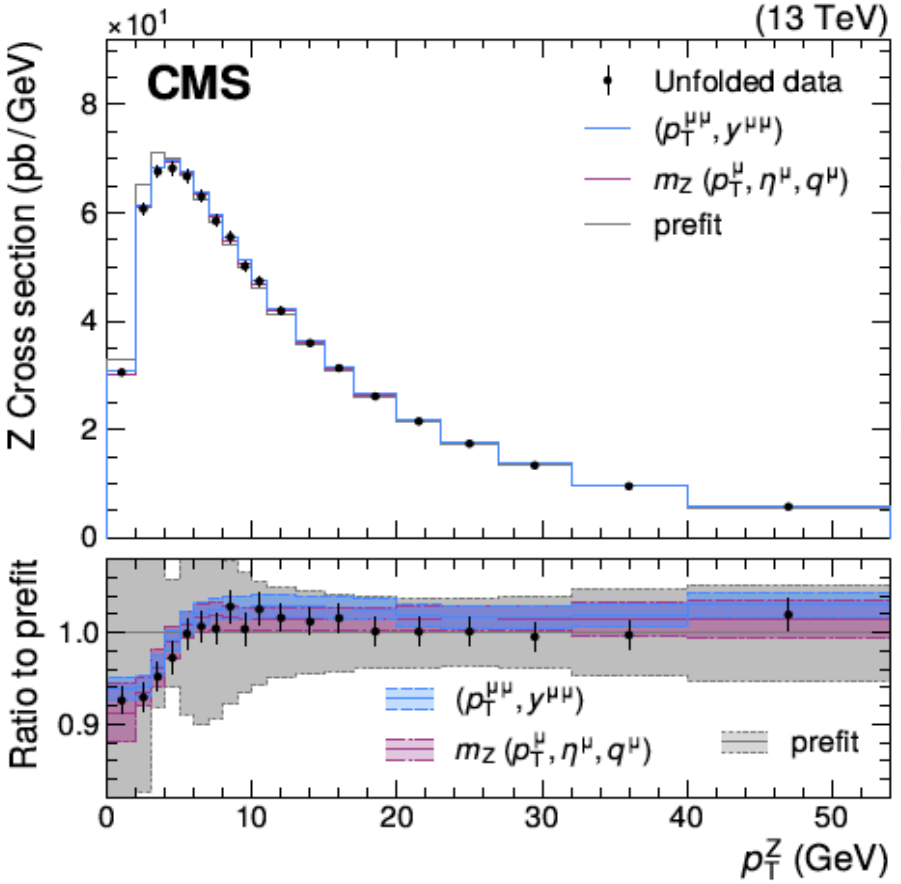
Figure A.10: Measured and simulated $p_T^{\mu\mu}$ distributions in the Z boson events, with the norm ization and uncertainties of the prediction set to the post-fit values. The gray band represe the total systematic uncertainty.

Figure 2: **Validation of the theory model**: Unfolded measured $p_{\mathrm{T}}^{\mathrm{Z}}$ distribution compared with the generator-level SCETLIB + DYTURBO predictions before (prefit, gray) and after adjusting the nuisance parameters to the best fit values obtained from the W-like m_{Z} fit (magenta) or from the direct fit to the $p_{\mathrm{T}}^{\mu\mu}$ distribution (blue).

Consistency of the distributions obtained:

- From the direct $p_T^{\mu\mu}$ fit
- From the $W-like\, m_Z$ fit
- From the data (both of them)

Confirms the robustness of the predic-ions and uncertainty model as well as ability of the (p_T^{μ}, η^{μ}) distribution to constrain the p_T^V modeling in situ



Validation of theory model

Consistency above supports using same treatment for p_T^W distribution in the m_W extraction

$$p_T^{\mu\mu} \not\Rightarrow p_T^W$$

$$(p_T^\mu, \eta^\mu) \Rightarrow p_T^W$$

Measurement of the W boson mass

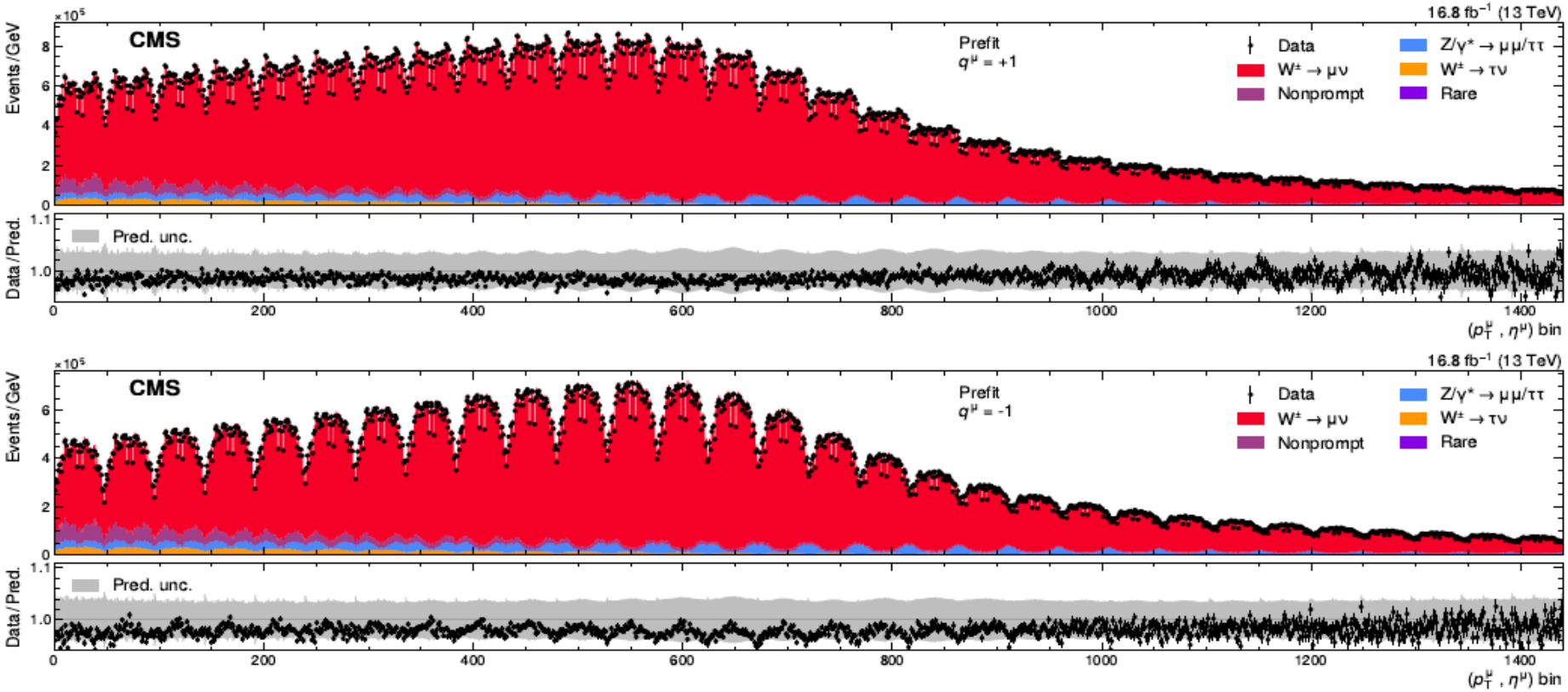
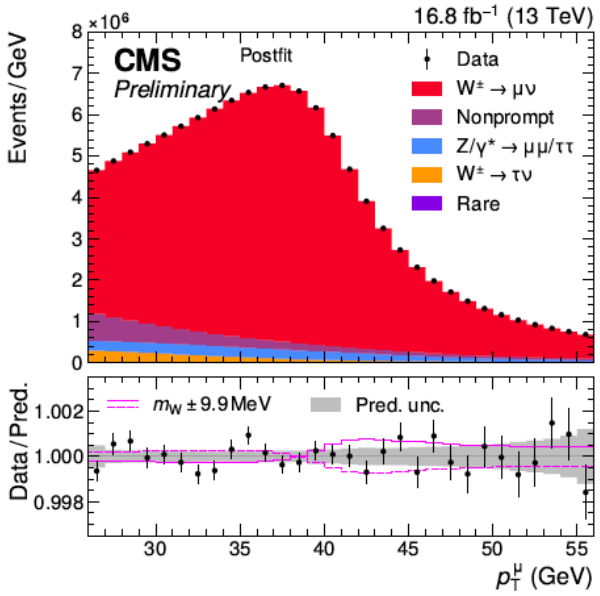


Figure A.16: Measured and predicted $(p_{\mathrm{T}}^{\mu}, \eta^{\mu})$ distributions for positively (upper) and negatively (lower) charged muons. The two-dimensional distribution is "unrolled" such that each bin on the x-axis represents one $(p_{\mathrm{T}}^{\mu}, \eta^{\mu})$ cell. The gray band represents the uncertainty in the prediction, before the fit to the data.

m_W Measurement

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



Source of uncertainty	Impact (MeV) Nominal	Global
Muon momentum scale	4.8	4.4
Muon reco. efficiency	3.0	2.3
W and Z angular coeffs.	3.3	3.0
Higher-order EW	2.0	1.9
p_T^V modeling	2.0	0.8
PDF	4.4	2.8
Nonprompt background	3.2	1.7
Integrated luminosity	0.1	0.1
MC sample size	1.5	3.8
Data sample size	2.4	6.0
Total uncertainty	9.9	9.9

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

Measurement of the W boson mass

From binned maximum likelihood template fit to the $\left(p_T^{\mu}, \eta^{\mu}, q^{\mu}\right)$ distribution

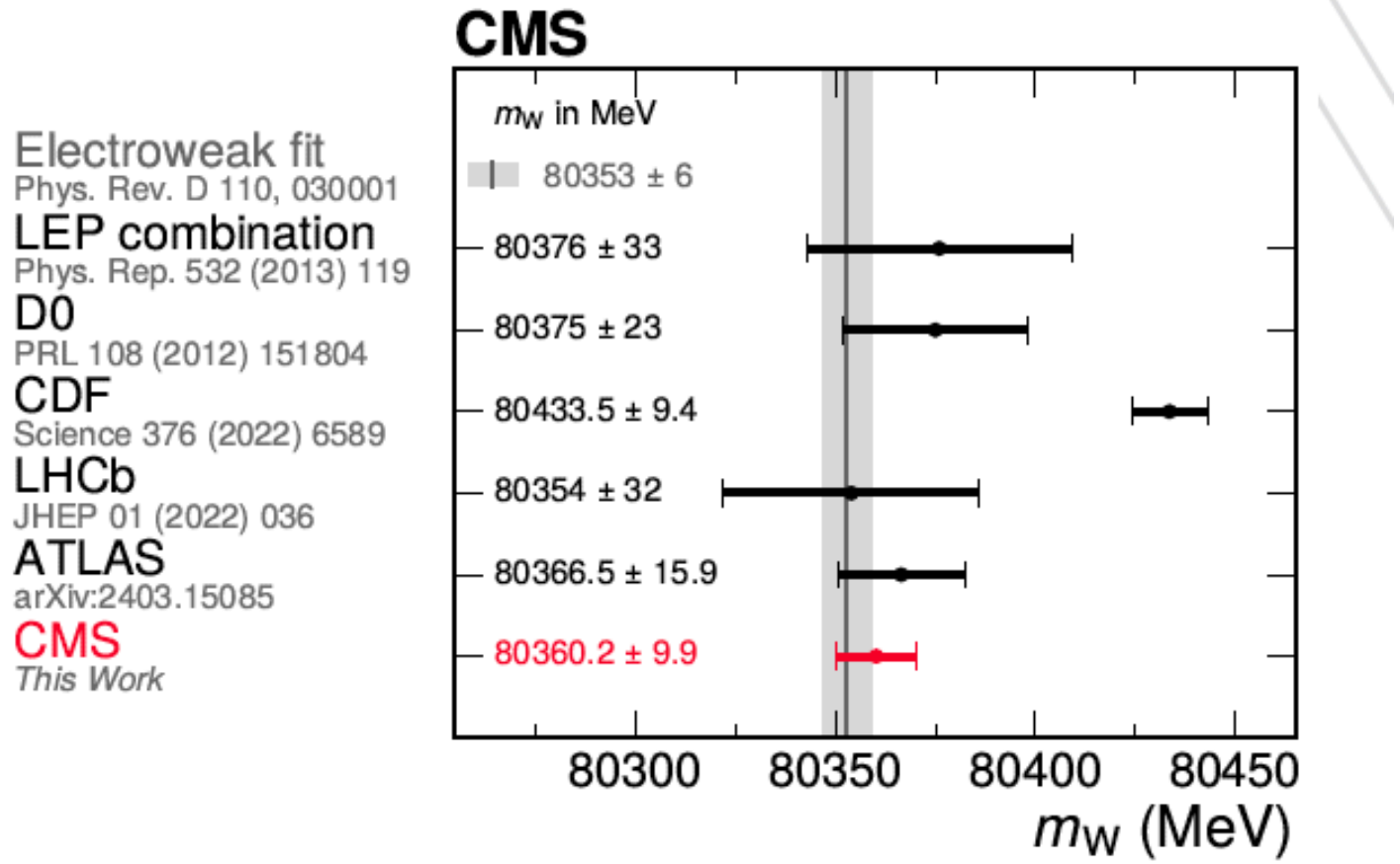
$$m_W = 80\,360.2 \pm 2.4(stat.) \pm 9.6(syst.)\ MeV$$

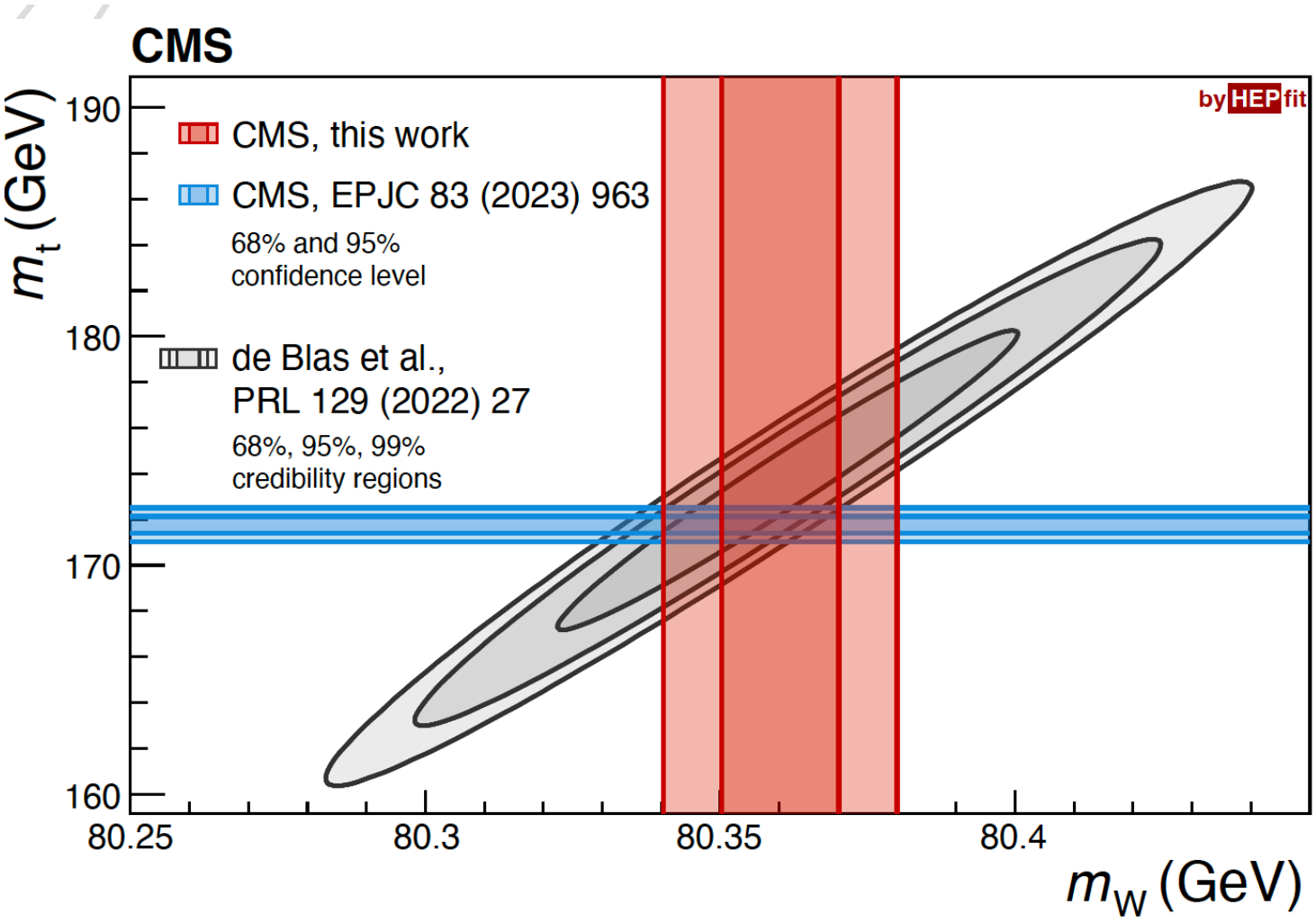
$$m_W = 80\,360.2 \pm 9.9\ MeV$$

in agreement with the EW fit and all experimental results except CDF

Summary

The precise W boson mass measurement as a test of the SM




CMS
CMS, this work
CMS, EPJC 83 (2023) 963
68% and 95%
confidence level
de Blas et al.,
PRL 129 (2022) 27
68%, 95%, 99%
credibility regions
byHEPfit
m_t (GeV)
190
180
170
160
80.25
80.3
80.35
80.4
m_W (GeV)

Source of uncertainty	Impact (MeV)			
	Nominal		Global	
	in m_Z	in m_W	in m_Z	in m_W
Muon momentum scale	5.6	4.8	5.3	4.4
Muon reco. efficiency	3.8	3.0	3.0	2.3
W and Z angular coeffs.	4.9	3.3	4.5	3.0
Higher-order EW	2.2	2.0	2.2	1.9
p_T^V modeling	1.7	2.0	1.0	0.8
PDF	2.4	4.4	1.9	2.8
Nonprompt background	–	3.2	–	1.7
Integrated luminosity	0.3	0.1	0.2	0.1
MC sample size	2.5	1.5	3.6	3.8
Data sample size	6.9	2.4	10.1	6.0
Total uncertainty	13.5	9.9	13.5	9.9

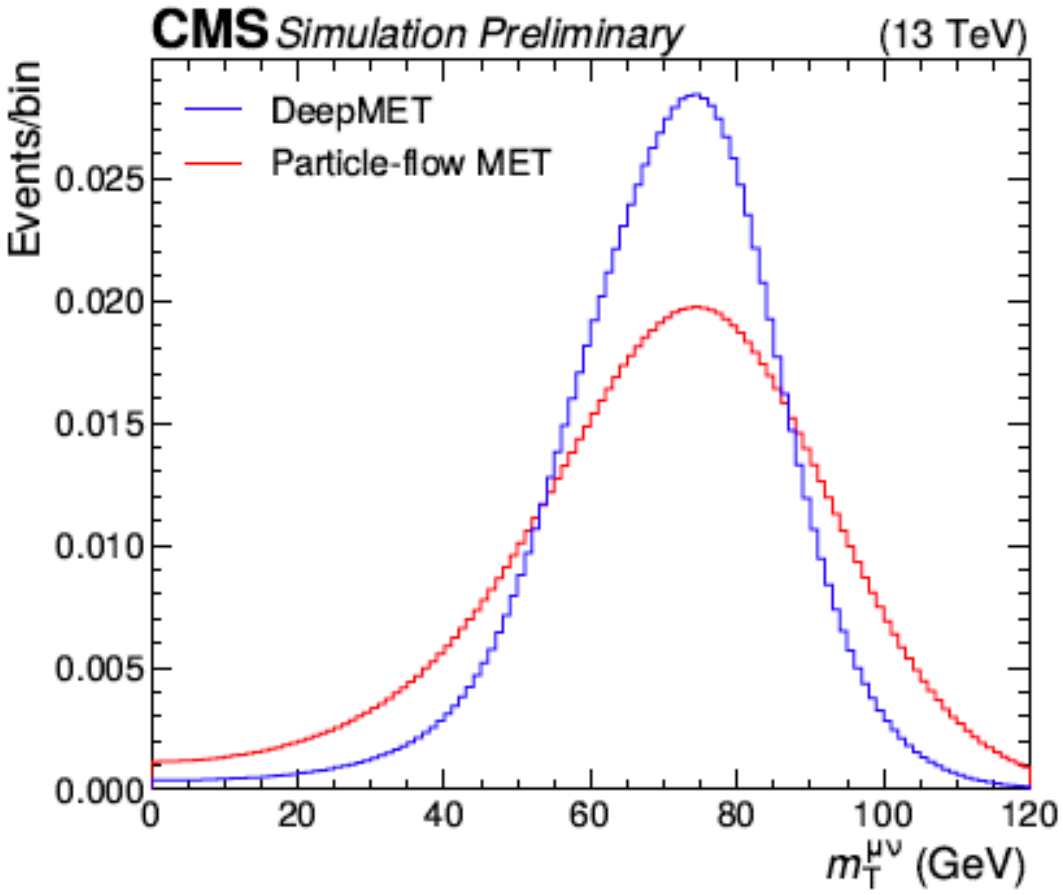
$$m_W = 80\,360.2 \pm 9.6\ MeV$$

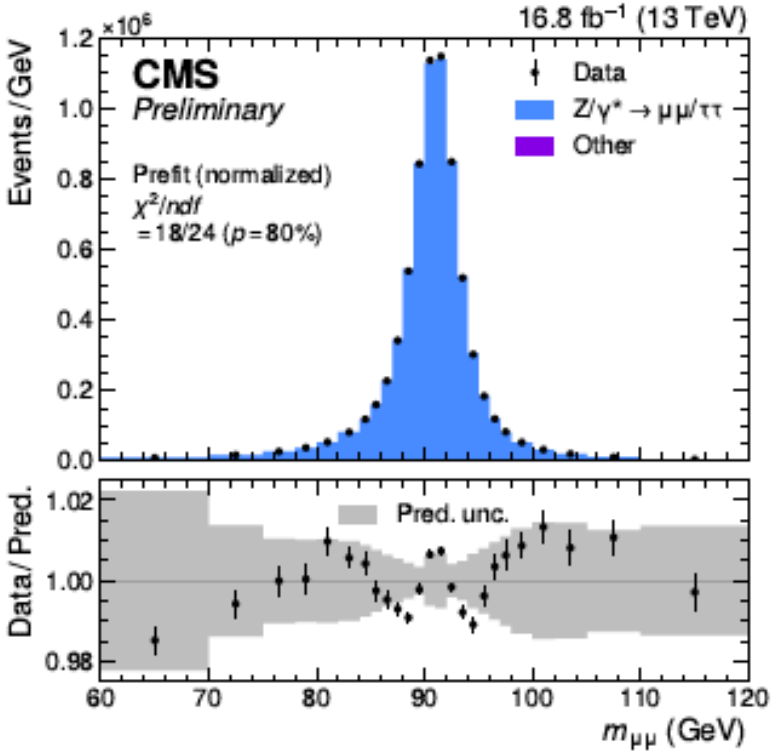
m_W measurement in agreement with SM prediction !!!

This is most precise m_W measurement at LHC up to now

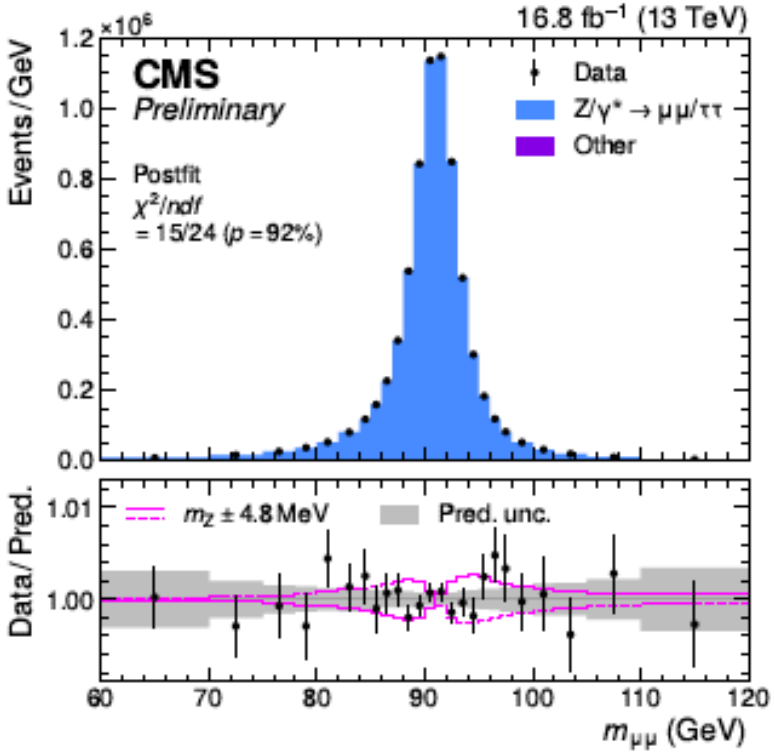
THANK YOU !!!

BACKUP

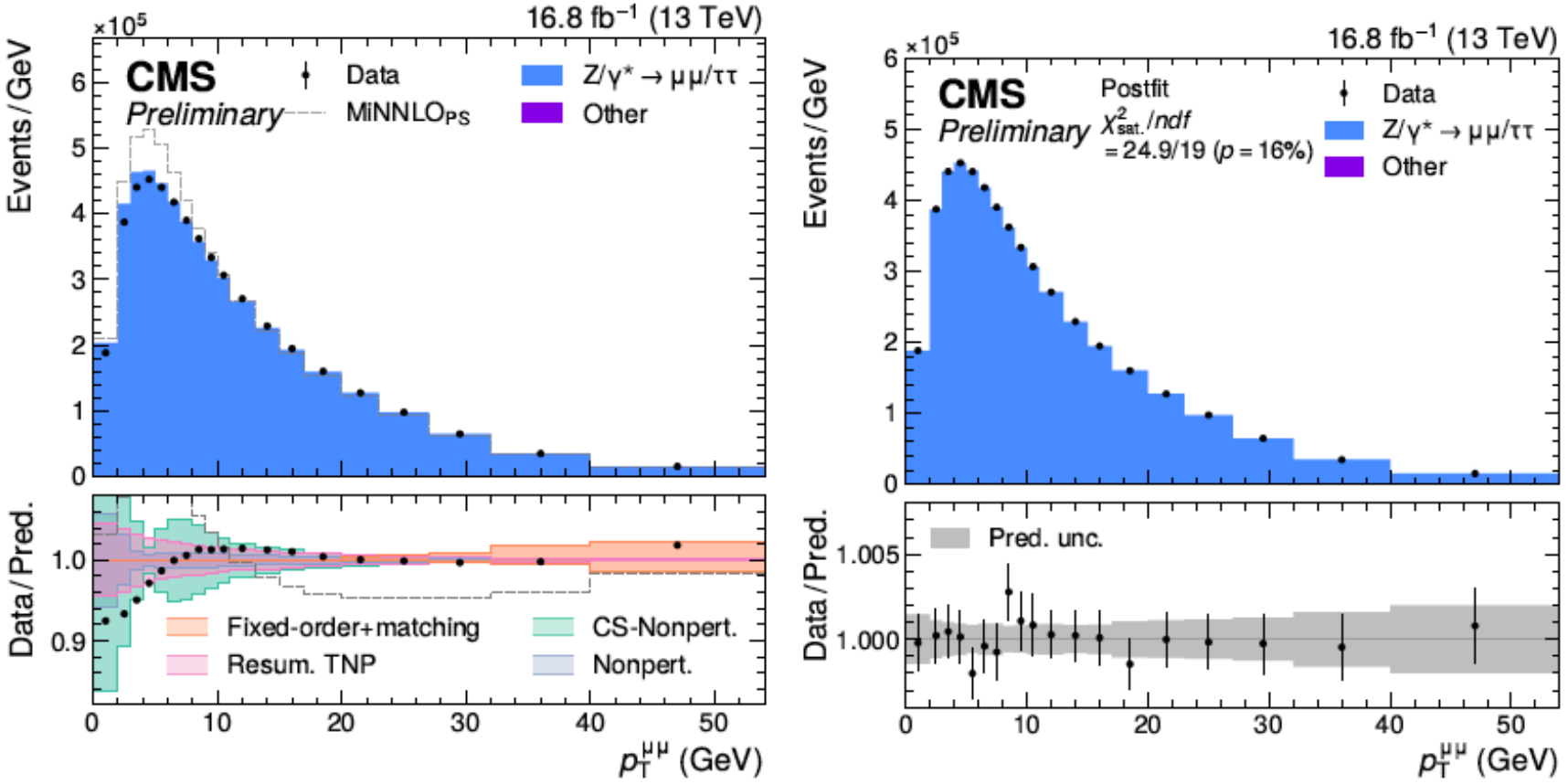

CMS Simulation Preliminary
(13 TeV)
DeepMET
Particle-flow MET
Events/bin
0.025
0.020
0.015
0.010
0.005
0.000
0
20
40
60
80
100
120
$m_T^{\mu\nu}$ (GeV)



(a) prefit (normalised)

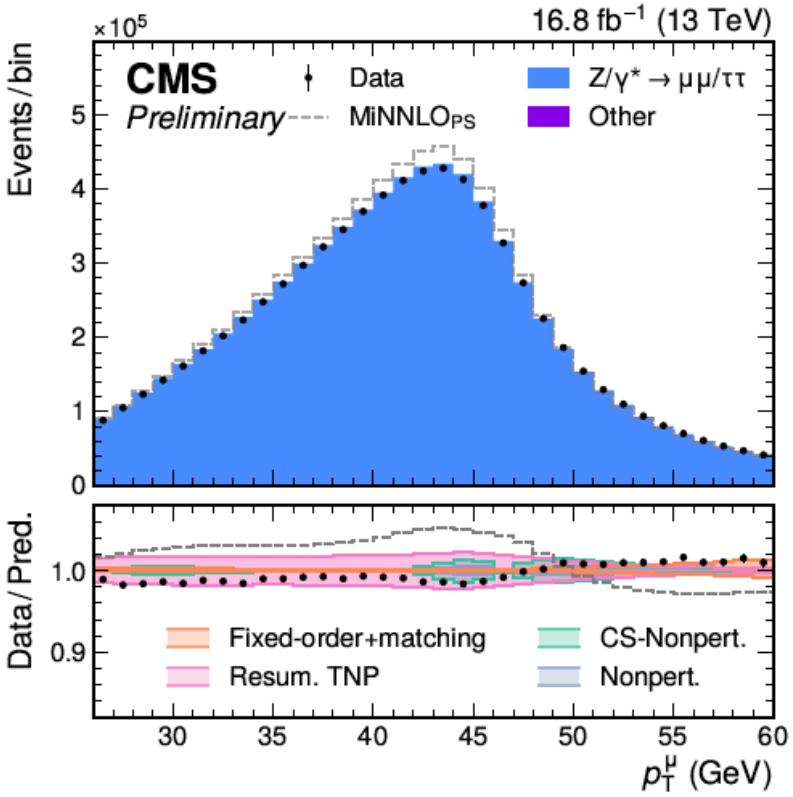


(b) postfit

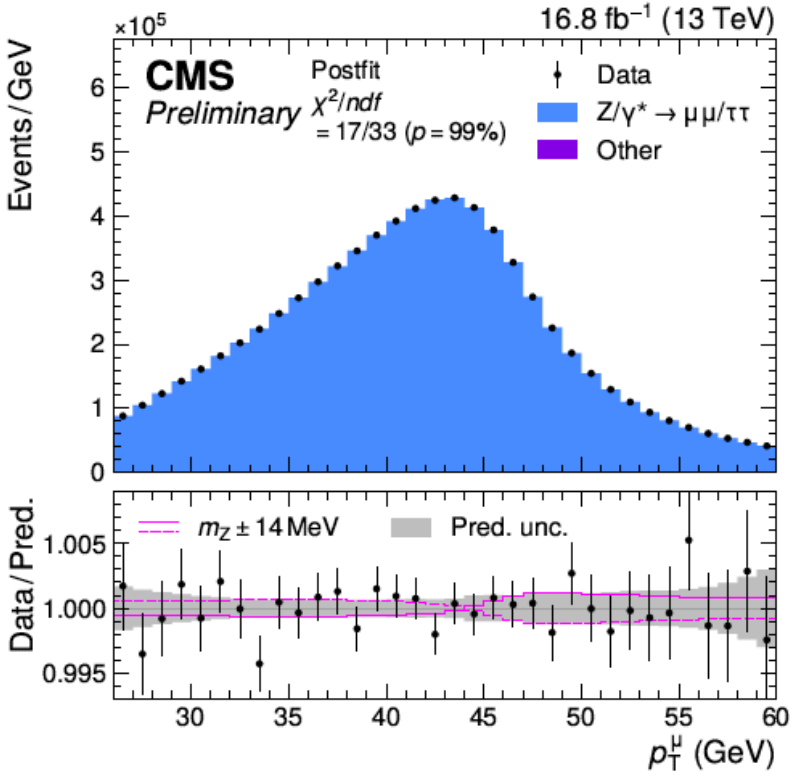


(a) prefit

(b) postfit



(a) prefit

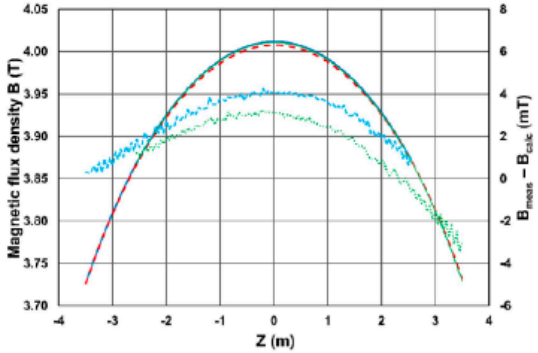


(b) postfit

Magnetic Field Model

- High granularity (33,840 space points) 3D field map taken in 2006 (but on the surface and without much of the detector)
  - NMR probes with relative accuracy better than 5e-5 and calibrated hall probes with accuracy of ~3e-4
- TOSCA model+parameterization used for track reconstruction reproduces field map data to +-0.1% with some variation vs z
- Possible future improvement: use the (interpolated) field map data directly
- Several NMR probes inside the solenoid (but outside the tracking volume) for monitoring
- **Magnetic field in tracking volume known to 0.1% a priori**
  - Residual corrections at this level not-unexpected
  - Uniformity could possibly be improved with direct use of field map data

JINST 5:T03021,2010
Symmetry 14 (2022) 169



Model vs field map data at R = 0.1m (surface)

Source	Field	Δ (rel.)
Surface NMR (2006)	3.9176T	-8e-4
In-situ NMR (2008)	3.9206T	0
In-situ Model Prediction	3.9181T	-6e-4

Model vs NMR Measurements at R = 2.91m, z = -0.01m

oun *mathematics, physics* A procedure to obtain a finite result from a divergent sum (series) of functions, involving the integral transformation of another (convergent) function in which the individual terms defining the original function are rescaled.

In the SCET formalism used here there are three perturbative ingredients in the $p_{\mathrm{T}}^{\mathrm{V}}$ resummation: the "hard function" that describes the hard virtual corrections for W and Z production, the "proton beam functions" that describe standard nonperturbative PDFs as well as perturbative collinear radiation, and the "soft function" describing soft radiation. All these functions share a system of renormalization group equations whose solution yields the all-order resummation of logarithms of $p_{\mathrm{T}}^{\mathrm{V}}/m_{\mathrm{V}}$. In the TNP approach, the minimal independent set of ingredients that would be needed at the next perturbative order are identified and parametrized in terms of common nuisance parameters. Specifically, there are six sources of TNPs: the three fixed-order boundary conditions of each of the hard, soft, and beam functions, and three anomalous dimensions governing their renormalization group evolution, namely the cusp anomalous dimension (Γ_{cusp}) and the virtuality and rapidity noncusp anomalous dimensions (γ_μ and γ_ν). The TNPs of the hard and soft functions and the three anomalous dimensions are numerical constants.

we can expand eq. (7.133) in λ,

$$\frac{L_i(q,p_T^\ell)}{L_{\pm(i)}(q^2)} = \frac{3}{4\pi}\frac{1}{\sqrt{2Q}}\int_0^{2\pi} \mathrm{d}\varphi\, \frac{\theta(q_T\, c_\varphi + p_L)}{\sqrt{q_T\, c_\varphi + p_L}}\, g_i\big(\pi/2,\varphi\big) \times \big[1+\mathcal{O}(\lambda)\big]\,. \tag{7.137}$$

This vanishes for all g_i odd in φ, which only leaves $i=-1,0,2,3$. This should be contrasted with the naive LP result in eq. (7.134), which only receives contributions from $i=-1,0$. The $i=2$ contribution is proportional to the double Boer-Mulders effect, which we can neglect, see the discussion below eq. (7.89). For $i=3$ we have $W_3\sim\mathcal{O}(\lambda)$, see table 7.1, which thus yields a linear power correction. Hence, we find the interesting effect that the proximity to the Jacobian peak induces sensitivity to new hadronic structure functions at $\mathcal{O}(\lambda)$, which do not contribute at $\mathcal{O}(\lambda)$ away from the peak region.

From eq. (7.137) it is evident that naively expanding in q_T near the Jacobian peak would amount to expanding in q_T/p_L, which is not allowed. However, eq. (7.137) is only valid near the peak, because by counting $p_L/Q\sim\lambda$ we have expanded away the dependence on $\kappa = 1+\mathcal{O}(\lambda)$, which is not allowed away from the peak. Hence, to cover the full range of p_T^ℓ, we must not expand in p_L, while near the peak we must count $q_T\sim p_L$ to avoid inducing uncontrolled leptonic power corrections in q_T/p_L. Clearly, the simplest way to satisfy both requirements is to not expand at all and keep the exact result corresponding to eq. (7.133).

Finally, note that the breakdown of the naive power expansion around $p_T^\ell = Q/2$ does not immediately affect the leptonic tensor if we only consider a fiducial cut $p_T^\ell \geq p_T^{\text{min}}$, since we can evaluate it as

$$\Phi_L(q,p_T^{\text{min}}) = \int_{p_T^{\text{min}}} \mathrm{d}p_T^\ell \frac{\mathrm{d}\Phi_L(q)}{\mathrm{d}p_T^\ell} = \frac{1}{8\pi} - \int_0^{p_T^{\text{min}}} \mathrm{d}p_T^\ell \frac{\mathrm{d}\Phi_L(q)}{\mathrm{d}p_T^\ell}\,. \tag{7.138}$$

Thus, the leptonic power corrections in this case scale as $q_T/(Q-2p_T^{\text{min}})$, and so as long as $p_T^{\text{min}}\ll Q/2$, the effect of p_T^{min} can be treated as a linear fiducial power correction as discussed for the q_T spectrum with fiducial cuts in section 7.4.2.

Numerical results

There are two key insights from our analysis of the differential p_T^ℓ phase space. First, the p_T^ℓ spectrum near the Jacobian peak is directly sensitive to the small transverse momentum q_T of the decaying vector boson. This causes fixed-order predictions to become unreliable in this region, which is a well-known effect. Second, the strict $q_T\to 0$ limit by itself cannot describe the p_T^ℓ spectrum in this region, which means the strict LP q_T resummation is also insufficient. Both problems are cured simultaneously by combining the exact leptonic tensor, which encodes the exact decay kinematics and automatically retains all leptonic power corrections, with the q_T-resummed hadronic tensor, thus allowing us to obtain physical predictions around the Jacobian peak.

We illustrate this in figure 7.10 for the p_T^ℓ spectrum in $W^+\to\ell^+\nu_\ell$ decays, where we show the spectrum both at fixed order (left) and after resummation including fiducial

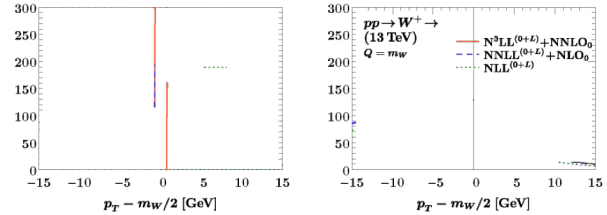


Figure 7.10: Lepton transverse momentum spectrum for on-resonance W^+ production at the LHC at fixed order (left) and including the resummation of fiducial power corrections to N^{3}LL (right). The horizontal axes shows the distance to the Jacobian peak at $p_T^\ell = m_W/2$.

power corrections (right). In both panels, the horizontal axis shows the distance of p_T^ℓ to the Jacobian peak at $p_T^\ell = m_W/2$, and to avoid smearing out the peak we consider the spectrum at a fixed point $Q=m_W$. The fixed-order spectrum (left) is shown at LO$_0$ (green dotted), NLO$_0$ (blue dashed), and NNLO$_0$ (red solid). The LO$_0$ result corresponds to Born kinematics and clearly shows the kinematic edge at $p_T^\ell = Q/2$. Starting at NLO$_0$, the W boson can have nonvanishing q_T, which opens up the phase space beyond the edge. However, in the vicinity of the edge, the fixed-order predictions become unstable due to the sensitivity to small q_T, which is clearly visible by the diverging NLO$_0$ and NNLO$_0$ curves, and in particular by the sign change between NLO$_0$ and NNLO$_0$ at $p_T^\ell\sim Q/2$.

In the right panel in figure 7.10, we show the resummed p_T^ℓ spectrum at NLL$^{(0+L)}$ (green dotted), NNLL$^{(0+L)}$+NLO$_0$ (blue dashed), and N^{3}LL$^{(0+L)}$+NNLO$_0$ (red solid). The resummation including leptonic power corrections cures the unphysical behaviour of the fixed-order results, yielding a well-behaved spectrum in the full p_T^ℓ range, with a resummed Sudakov shoulder at $p_T^\ell\sim m_W/2$. Note that the cross section beyond the edge is already populated at NLL$^{(0+L)}$ without any fixed-order matching. We stress that without including the exact leptonic tensor, the resummation would only affect the region $p_T^\ell < m_W/2$, and not cure the peak region. In fact, the results with strict LP resummation would look very similar to the pure fixed-order results, with the N^{3}LL$^{(0)}$+NNLO$_0$ essentially indistinguishable from the pure NNLO$_0$ result.

This is the first time that resummed N^{3}LL results for the p_T^ℓ spectrum are presented, and we observe extremely good perturbative convergence, with the results at NNLL$^{(0+L)}$+NLO$_0$ and N^{3}LL$^{(0+L)}$+NNLO$_0$ falling on top of each other. We leave a more detailed phenomenological analysis of the p_T^ℓ spectrum to future work.

7.4.4 ϕ^* spectrum

The ϕ^* observable was first proposed in ref. [497], extending earlier work on the a_T observable [498, 499]. Both observables are sensitive to small q_T, but promise better experimental

perturbation theory contains physical information, namely, that the most divergent contribution from an additional emission arises from the limit where it becomes simultaneously soft and collinear.

Resummation

The result in eq. (1.6) is problematic because for $\tau_{\text{cut}} \ll 1$, the double logarithms of τ_{cut} become large and overcome the suppression by α_s, so the perturbative series *diverges* and cannot be truncated. To arrive at any meaningful prediction, we must find a way to reorganize, or *resum*, the perturbative series. To do so, we retain to first approximation only the most singular *leading logarithmic* (LL) terms $m = 2n$. Using the fact that eq. (1.4) holds recursively for every subsequent emission, one can show that the coefficients satisfy $C_{n,2n} = (-C)^n/n!$ with $C > 0$ a constant that depends on the charge of the primary emitter. (A pedagogical derivation can be found in ref. [38].) These are the coefficients of an exponential series, so we find

$$\sigma(\tau_{\text{cut}}) = \sigma^{(0)} \exp\Big[-C\,\alpha_s \ln^2 \tau_{\text{cut}}\Big] + (\text{terms with } m < 2n) + \mathcal{O}(\tau_{\text{cut}})\,. \tag{1.7}$$

This result is known as the Sudakov form factor [39], and is perfectly convergent for $\tau_{\text{cut}} \to 0$, where it tends to 0. It is the basis of the most elementary type (C) predictions: These so-called *parton showers* [40] recursively add emissions in a Markov process using eq. (1.7) as the probability for no emission to occur above a given cut. It is important to realize that eq. (1.7), despite being an all-order result in α_s, is only the first term in a systematic expansion to higher logarithmic order, and on its own is essentially as precise (or imprecise) as a leading-order calculation in the cases where fixed-order perturbation theory converges.

To increase the precision of this result, we should include the next-to-leading logarithmic (NLL) terms $m = 2n-1$, the next-to-next-to-leading logarithmic (NNLL) terms $m = 2n-2$, and so forth. Many methods exist to extend the analytic resummation to subleading logarithmic orders. They all rely on the principle of *factorization*, i.e., a systematic separation of the dynamics at the low energy scale $\tau_{\text{cut}}Q$, where soft and collinear radiation is emitted, from the hard production process that occurs at the scale $Q \gg \tau_{\text{cut}}Q$. Specifically, we will make use of *effective field theory* in this thesis to make the separation of scales manifest at the level of the cross section.

The drawback of all these methods is that they only apply to type (A) predictions in the presence of, typically, a single experimental constraint. A common theme of this thesis is the extension of these analytic resummation methods to more differential observables. In this way, realistic experimental observables that could previously only be computed using parton showers can now systematically be computed to higher logarithmic orders for the first time, extending the range of observables that can benefit from precise analytically resummed type (A) predictions. In other cases, a component of the calculation that would otherwise have to be calculated in an expensive type (B) calculation can now be evaluated in much shorter time using our results.

+ Add image

Fixed order QCD and resummation

- In the perturbative regime these logarithms can become as large as (breakdown of the PT below this limit)

$$L \sim \frac{1}{\alpha_s}$$

- This makes “higher order” corrections as large as leading order ones, i.e. $(\alpha_s L)^n L \sim \alpha_s L^2$

- The PT series breaks down and the probability of the reaction diverges logarithmically in the large L limit instead of being suppressed

- The resummation of the large logarithms to all perturbative orders restores the correct physical (Sudakov) suppression and rescues the predictive power of perturbation theory

What We Should be Doing.

$$f(\alpha) = f_0 + f_1\,\alpha \underbrace{+ f_2\,\alpha^2 + f_3\,\alpha^3 + \mathcal{O}(\alpha^4)}_{\text{source of the theory uncertainty}}$$

Identify the actual source of uncertainty: unknown $f_2, f_3, \ldots$

- $f(\alpha)$ is only a function of α $\Rightarrow$ f_n are just numbers
  - Simplest case
- $f(\alpha) = f(\alpha; x)$ $\Rightarrow$ $f_n(x)$ are functions
  - If dominant/leading x dependence is known $\rightarrow$ can be reduced to previous
- $f(\alpha) = f(\alpha; x_1, x_2, \ldots)$ $\Rightarrow$ $f_n(x_1, x_2, \ldots)$ are N-dim. functions
  - Need honest/realistic estimate of an unknown function ...

KALMAN FILTER

w_k fluctuation parameters
$\in_k$ measurement errors
m_k measurement of x_k
x_k track parameters

One can understand the basic idea of the Kalman filter in the following way. If there is an estimate of the state vector at time (location) t_{k-1}, it is extrapolated to time t_k by means of the system equation. The estimate at time t_k is then computed as the weighted mean of the predicted state vector and of the actual measurement at time t_k, according to the measurement equation. The information contained in this estimate can be passed back to all previous estimates by means of a second filter running backwards or by the smoother.

The main formulas for our linear dynamic system are the following:

System equation:

$$\mathbf{x}_k = F_k \cdot \mathbf{x}_{k-1} + \mathbf{w}_k \tag{23}$$

$$E\{\mathbf{w}_k\} = 0, \quad \mathrm{cov}\{\mathbf{w}_k\} = Q_k \quad (1 \leq k \leq N) \tag{24}$$

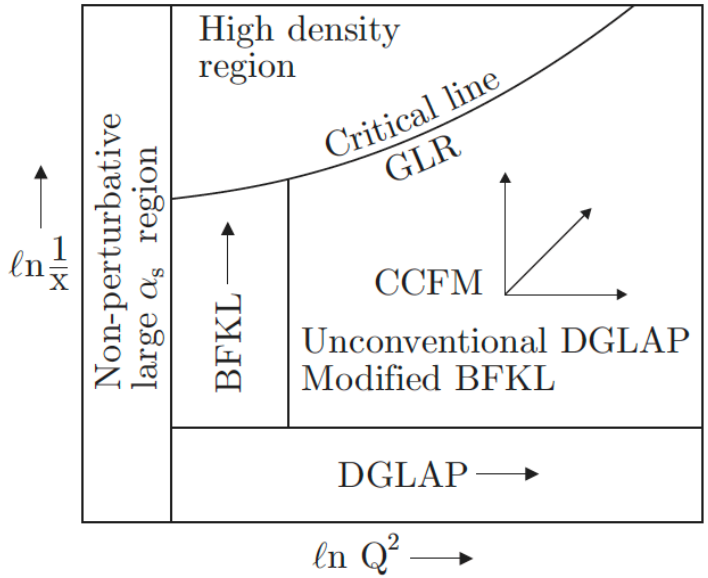
Measurement equation:

$$m_k = H_k \cdot \mathbf{x}_k + \varepsilon_k \tag{25}$$

$$E\{\varepsilon_k\} = 0, \quad \mathrm{cov}\{\varepsilon_k\} = V_k = G_k^{-1} \quad (1 \leq k \leq N) \tag{26}$$

where the matrices Q_k and V_k represent the process noise (multiple scattering, bremsstrahlung, etc.) and measurement noise (detector resolution) respectively. The details of Q_k calculation for the parameterization adopted in NOMAD can be found in Ref. [10].

W ogólności funkcje struktury można przedstawić w postaci szeregów potęgowych z parametrem rozwinięcia α_s. Szeregi te zawierają zarówno człony proporcjonalne do $\ln \frac{Q^2}{Q_0^2}$ jak i do $\ln \frac{1}{x}$. W rzędzie wiodącym standardowej ewolucji DGLAP (LO DGLAP) następuje resumacja wkładów logarytmicznych $(\alpha_s \ln \frac{Q^2}{Q_0^2})^n$ (patrz podrozdział 2.4.2). W rzędzie następnym do wiodącego NLO DGLAP (*Next-to-Leading Order*) sumowane są człony $\alpha_s(\alpha_s \ln \frac{Q^2}{Q_0^2})^{n-1}$ [28, 29], które pojawiają się gdy pędy poprzeczne dwóch sąsiednich partonów wymienianych w kanale t nie są silnie uporządkowane lecz porównywalne, co powoduje utratę czynnika $\ln \frac{Q^2}{Q_0^2}$. W ostatnich latach przeprowadzono obliczenia poprawek w rzędzie następnym do niewiodącego NNLO DGLAP (*Next-to-Next-to-Leading Order*) dla funkcji struktury F_1, F_2 i F_3 [30].



Radiacja gluonów kreuje niezerowy pęd poprzeczny partonów wymienianych w kanale t; każda emisja daje przyczynek proporcjonalny do $\alpha_s \int \frac{dk_t^2}{k_t^2}$. Oznacza to, że wkład do funkcji struktury od diagramu drabinkowego o n szczeblach wynosi:

$$\alpha_s^n \int_{Q_0^2}^{Q^2} \frac{dk_{nt}^2}{k_{nt}^2} \cdots \int_{Q_0^2}^{k_{3t}^2} \frac{dk_{2t}^2}{k_{2t}^2} \int_{Q_0^2}^{k_{2t}^2} \frac{dk_{1t}^2}{k_{1t}^2} = \frac{\alpha_s^n}{n!} \ln^n \frac{Q^2}{Q_0^2} \quad . \tag{2.29}$$

Duże logarytmy w Q^2 wynikają z wycałkowania po uporządkowanych pędach poprzecznych partonów wymienianych w kanale t. W każdym rzędzie n duże logarytmy $\ln(Q^2/Q_0^2)$ kompensują malejącą logarytmicznie z Q^2 silną stałą sprzężenia α_s, $\alpha_s^n \ln^n(Q^2/Q_0^2) \sim 1$. Dlatego też, aby otrzymać skończony wynik w rozwinięciu perturbacyjnym należy wysumować wszystkie diagramy z n rozciągającym się do ∞. Przedstawiony schemat obliczeń nosi nazwę przybliżenia wiodących logarytmów (*Leading Log Approximation, LLA*), ponieważ każdej potędze n w α_s towarzyszy taka sama potęga w $\ln(Q^2/Q_0^2)$. Obliczenia w przybliżeniu LLA są poprawne dla dużych wartości Q^2 i niezbyt małych wartości x, zdefiniowanych warunkiem $\alpha_s(Q^2) \ln \frac{1}{x} \ll \alpha_s(Q^2) \ln \frac{Q^2}{Q_0^2}$. Przybliżenie LLA nazywane jest również przybliżeniem współliniowym (*collinear approximation*). Zarówno wirtualności jak i pędy poprzeczne partonów wymienianych w kanale t są zaniedbywalne w porównaniu z twardą skalą Q^2 określoną przez wirtualność fotonu. Partony możemy więc traktować jako bezmasowe i poruszające się w tym samym kierunku (współliniowo) co proton.

Przy wyborze skali $\mu_r = \mu_f = \mu$ różniczkowo-całkowe równania DGLAP, opisujące ewolucję w μ^2 rozkładu partonów $f(x, \mu^2)$, można zapisać symbolicznie:

$$\frac{\partial f}{\partial \ln \mu^2} \sim \frac{\alpha_s(\mu^2)}{2\pi} (P \otimes f), \tag{2.30}$$

skali renormalizacji i faktoryzacji, $\mu_r^2 = \mu_f^2 = \mu^2 = Q^2 + 4p_t^2$,

tion criteria. For ATLAS analyses, a common muon isolation criterion has been defined as follows[13]:

- number of tracks in cone $0.2 < 2$
- energy deposit in the electromagnetic calorimeter $< 2GeV$, inner cone: 0.075, outer cone 0.15
- energy deposit in the hadronic calorimeter $< 10GeV$, inner cone: 0.15, outer cone 0.30.

CMS

defined as

$$I_{\mathrm{PF}} = \frac{1}{p_{\mathrm{T}}} \left(\sum_{\mathrm{h}^{\pm}} p_{\mathrm{T}}^{\mathrm{h}^{\pm}} + \sum_{\gamma} p_{\mathrm{T}}^{\gamma} + \sum_{\mathrm{h}^{0}} p_{\mathrm{T}}^{\mathrm{h}^{0}} \right), \tag{5.4}$$

where the sums run over the charged hadrons ($\mathrm{h}^{\pm}$), photons (γ), and neutral hadrons (h^{0}) with a distance ΔR to the lepton smaller than either 0.3 or 0.5 in the (η, φ) plane.

Short answer: The eigenvector with the largest eigenvalue is the direction along which the data set has the maximum variance. Meditate upon this.

Long answer: Let's say you want to reduce the dimensionality of your data set, say down to just one dimension. In general, this means picking a unit vector u, and replacing each data point, x_i, with its projection along this vector, $u^T x_i$. Of course, you should choose u so that you retain as much of the variation of the data points as possible: if your data points lay along a line and you picked u orthogonal to that line, all the data points would project onto the same value, and you would lose almost all the information in the data set! So you would like to maximize the *variance* of the new data values $u^T x_i$. It's not hard to show that if the covariance matrix of the original data points x_i was Σ, the variance of the new data points is just $u^T \Sigma u$. As Σ is symmetric, the unit vector u which maximizes $u^T \Sigma u$ is nothing but the eigenvector with the largest eigenvalue.