

Combined EW+PDF fits

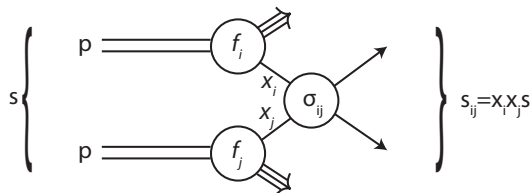
Louis Fayard, **Juan Tafoya**, Zhiqing Zhang

Universite Paris-Saclay

ATLAS - IJCLab

2023.02.24





$$\sigma_{pp \rightarrow X} = \sum_{i,j} \int dx_1 dx_2 \underbrace{f_i^p(x_1, Q^2)}_{PDF} \underbrace{f_j^p(x_2, Q^2)}_{PDF} \times \sigma_{ij}(x_1 x_2 s, \alpha_S(Q^2))$$

HERAPDF style :
$$x f_i(x) = A_i x^{B_i} (1-x)^{C_i} (1 + D_i x + E_i x^2) \underbrace{- A'_g x^{B'_g} (1-x)^{C'_g}}_{\text{term used exclusively for the gluon}}$$

Most studies take the EW and PDF parts as decoupled, using one to find the other:

- EW SM values are assumed in order to fit PDFs
- or
- PDFs are fixed to find the value of EW parameters

AIM → to perform combined EW+PDF fits

Combined fit motivation: W -boson mass uncertainty breakdown

	Stat. Unc.	Muon Unc.	Elec. Unc.	Recoil Unc.	Bckg. Unc.	QCD Unc.	EW Unc.	PDF Unc.	Total Unc.
m_{W^+}	8.9	6.6	8.2	3.1	5.5	8.4	5.4	14.6	23.4
m_{W^-}	9.7	7.2	7.8	3.3	6.6	8.3	5.3	13.6	23.4
m_{W^\pm}	6.8	6.6	6.4	2.9	4.5	8.3	5.5	9.2	18.5

Uncertainty breakdown for m_W in MeV, as obtained in the study [arXiv:1701.07240](https://arxiv.org/abs/1701.07240).

PDF is the main source of systematic uncertainties in EW measurements



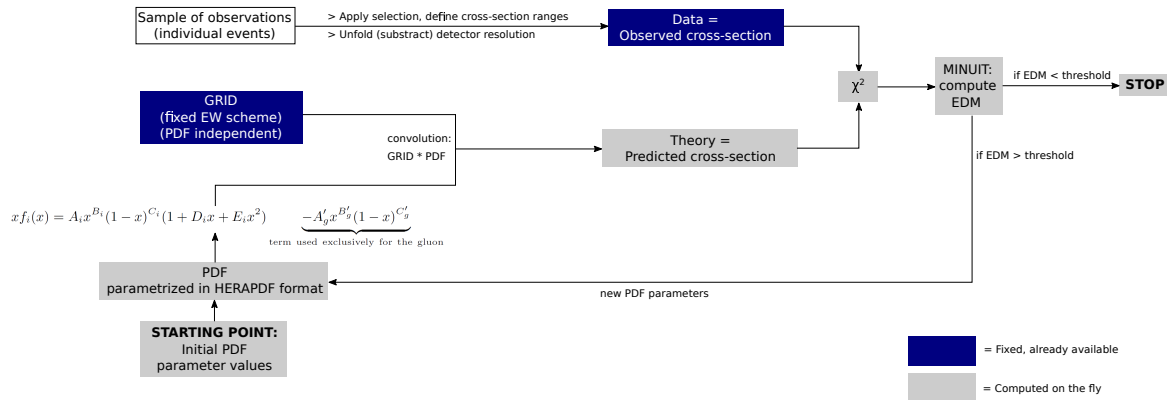
Neglecting the correlation to the PDF uncertainty can translate into an important mis-estimation of the corresponding total uncertainty



We are not particularly interested in improving central EW values

→ Instead, we want to find the correlation between the EW and QCD parts via a simultaneous fit

Fitting PDFs: the standard way using xFitter



Like this, one cannot fit an EW parameter on the fly!

The GRID is computed for fixed EW values → insensitive to these changes (straight out of the box)

We can write the prediction of a differential cross-section with respect to some variable Y (called $\frac{d\sigma}{dY}$) as a function of a PDF set and EW values as follows:

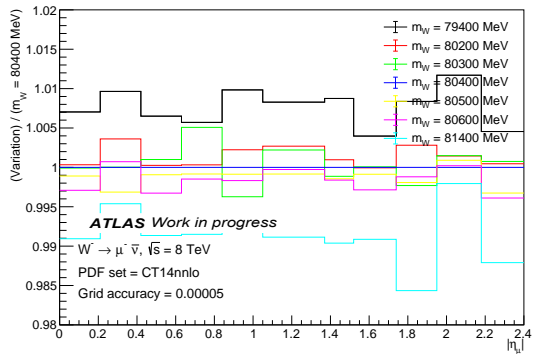
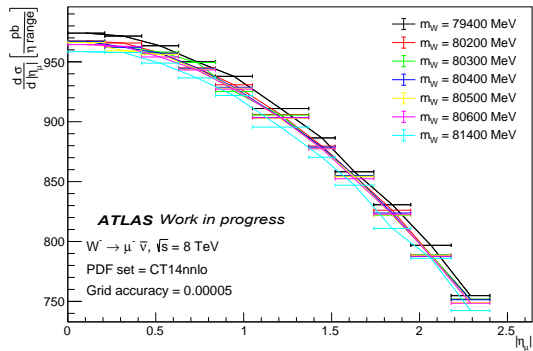
$$\frac{d\sigma}{dY}(\text{PDF, EW variation}) = [\text{GRID}_{\text{NLO}} * \text{PDF}] \times K_{\text{EW,LO} \rightarrow \text{NLO}}^{\text{QCD,NLO} \rightarrow \text{NNLO}} \times f_Y(\text{EW variation})$$

where

- GRID_{NLO} is an APPLgrid at NLO F.O., obtained for a fixed set of EW parameters.
- $\text{GRID}_{\text{NLO}} * \text{PDF}$ is the convolution of the fixed APPLgrid and the PDF being fitted
→ gives the corresponding NLO cross-section prediction.
- $K_{\text{EW,LO} \rightarrow \text{NLO}}^{\text{QCD,NLO} \rightarrow \text{NNLO}}$ are e.g. DYTURBO K-factors
- $f_Y(\text{EW variation})$ is a function (specific to each bin Y), whose parameters are fitted by looking at the relative change of XS by comparing several XS(Various EW values) with the nominal prediction XS(EW nominal). It allows to go from nominal EW values to the ones of interest for the fit.

Bad candidate: m_W in W inclusive $d\sigma/d|\eta_\mu|$ (1904.05631)

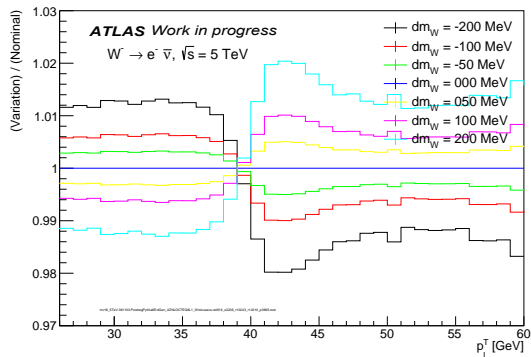
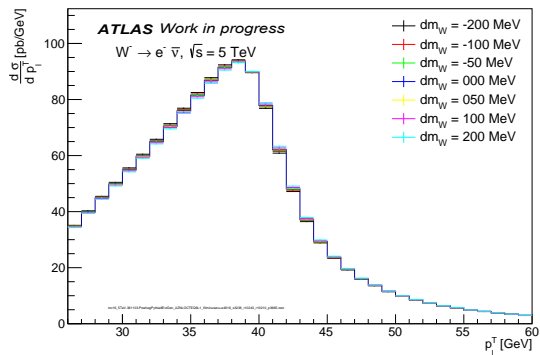
It's already used for PDF fits, but **not sensitive** to changes in m_W



Cross-section predictions obtained with APPLGrids and CT14nnlo

Good candidate: m_W in W inclusive $d\sigma/dp_T^T$ (using Powheg+Pythia samples)

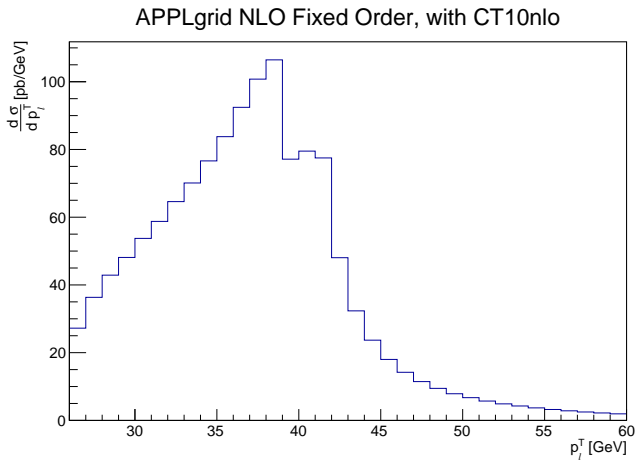
Very sensitive to m_W ! but it hasn't been used in PDF fits, and the corresponding low-pileup study is ongoing...



Cross-section predictions obtained with Powheg+Pythia samples (histogrammed with HistMaker)

Good candidate: m_W in W inclusive $d\sigma/dp_1^T$ (further details)

For a PDF fit, we need an APPLgrid Fixed Order prediction



(this differential cross-section corresponds to $W^- \rightarrow e\bar{\nu}$)

Ultimate goal

Get a combined EW+PDF fit which includes ALL the datasets used for ATLASpdf21 [2112.11266] (to constrain the PDF), plus the EW sensitive dataset (to constrain the EW parameter and the PDF).

That is, to fit:

- epWZ16: HERA [1506.06042, main PDF constraint] + WZ production [1612.03016]
- ttbar8TeV [1511.04716, 4 correlation files]
- Vjets [1711.03296 (W) and 1907.06728 (Z)]
- z3d [1710.05167]
- Wxs8TeV [1904.05631]
- photon [1901.10075]
- top13TeV [1908.07305, 6 correlation files]
- jets [1706.03192]

+ **Dataset sensitive to the EW parameter of interest (m_W):**

- LowPileup: 5 TeV W -boson p_i^T cross-sections [4 channels: $W^{+/-} \rightarrow e\nu$, $W^{+/-} \rightarrow \mu\nu$]

The dataset and prediction sensitive to $m_W \rightarrow d\sigma/dp_l^T$

- **Looking forward** to use the results of the low-pileup W-boson $d\sigma/dp_l^T$ cross-section study (ongoing)
- Data: Simulated differential p_l^T cross-sections (=Pseudodata, $m_W = 80.400$ GeV, $\sqrt{s} = 5$ TeV), 1 GeV bins in the $p_l^T \in [25, 60]$ GeV range (including systematic uncertainties, scaled to the Pseudodata)
- Prediction: APPLgrid Fixed Order grids \times k-factor, where

$$\text{k - factor} = \frac{\text{DYTURBO at NNLO, using PDF, with resummation}}{\text{APPLgrid} * \text{PDF}}$$

\rightarrow implement an NLO \rightarrow NNLO correction, and absorb the lack-of-resummation effect in the same factor

- EW variations: at each p_l^T bin, the differential cross-section change as a function of m_W can be defined as

$$\frac{d\sigma}{dp_l^T}(m_W) = \frac{d\sigma}{dp_l^T}(m_W^{\text{nominal}} = 80.400 \text{ GeV}) \left[1 + \text{factor}_{p_l^T}(m_W - 80.400 \text{ GeV}) \right]$$

where each $\text{factor}_{p_l^T}$ is determined before proceeding to the PDF fit

Useful quantity: given a correlation matrix C , the global correlation of a parameter k is defined as

$$\rho_k^2 = 1 - [C_{kk} * (C^{-1})_{kk}]^{-1}$$

[[https://doi.org/10.1016/0010-4655\(75\)90039-9](https://doi.org/10.1016/0010-4655(75)90039-9), page 356]

Closure tests

Approach

Use only the HERA and LowPileup datasets, where LowPileup = Pseudodata @ $m_W = 80.4\text{GeV}$

The test:

- 1 Do a PDF only fit on both datasets by fixing $M_W\text{fit} = 80.4$ [GeV]
- 2 Fix all the PDF parameters to the value of the fit, release only m_W and re-run fit \rightarrow should recover 80.4

Couple of possible approaches:

Approach 1: HERA + LowPileup data files.

- General enough PDF with physical meaning (i.e. would extrapolate well to other datasets).
- The resulting PDF is a compromise for the best agreement on the HERA and LowPileup data i.e. it could slightly degrade the agreement of the LowPileup prediction.

Approach 2: use only the LowPileup data.

- Should allow to get (almost) perfect agreement between the data and prediction for the LowPileup sets (at least, better than Approach 1).
- Terrible as a PDF. It would work only for this specific channel, but not extrapolate to other datasets.

Closure test, using only the LowPileup statistical uncertainty (agreement to data)

Approach 1: HERA + LowPileup

Dataset	PDFonly(80.4)	mWonly(80.4)
MC XS low-mu Wminusenu pTl 5 TeV	0.92 / 34	0.25 / 34
HERA1+2 NCep 820	56 / 61	56 / 61
HERA1+2 NCep 460	201 / 177	201 / 177
HERA1+2 CCep	49 / 39	49 / 39
MC XS low-mu Wplusenu pTl 5 TeV	2.1 / 34	1.4 / 34
MC XS low-mu Wminusmunu pTl 5 TeV	0.92 / 34	0.26 / 34
HERA1+2 NCem	231 / 159	231 / 159
HERA1+2 CCem	87 / 42	87 / 42
HERA1+2 NCep 575	194 / 221	194 / 221
HERA1+2 NCep 920	359 / 317	359 / 317
MC XS low-mu Wplusmunu pTl 5 TeV	2.1 / 34	1.4 / 34
Correlated χ^2	58	58
Log penalty χ^2	-39.85	-39.78
Total χ^2 / dof	1201 / 1136	1198 / 1151
χ^2 p-value	0.09	0.16

Approach 2: only LowPileup

Dataset	PDFonly	mWonly
MC XS low-mu Wminusenu pTl 5 TeV	0.0100 / 34	0.0100 / 34
MC XS low-mu Wplusenu pTl 5 TeV	0.0100 / 34	0.0100 / 34
MC XS low-mu Wminusmunu pTl 5 TeV	0.0100 / 34	0.0100 / 34
MC XS low-mu Wplusmunu pTl 5 TeV	0.0100 / 34	0.0100 / 34
Correlated χ^2	0	0
Log penalty χ^2	-0.02	-0.02
Total χ^2 / dof	0.040 / 120	0.040 / 135
χ^2 p-value	1.00	1.00

Closure test, using only the LowPileup statistical uncertainty (parameter values)

Approach 1: HERA + LowPileup

Parameter	PDFonly(80.4)	mWonly(80.4)
'Adbar'	0.0959 ± 0.0093	0.09588
'Adv'	1.0000	1.0000
'Ag'	1.0000	1.0000
'Agp'	0.421 ± 0.080	0.4209
'Auv'	1.0000	1.0000
'Bdbar'	-0.173 ± 0.016	-0.1727
'Bdv'	0.562 ± 0.030	0.5623
'Bg'	-0.47 ± 0.13	-0.4677
'Bgp'	-0.558 ± 0.073	-0.5578
'Buv'	0.730 ± 0.023	0.7305
'Cdbar'	6.0 ± 1.7	5.976
'Cdv'	2.97 ± 0.39	2.966
'Cg'	4.25 ± 0.77	4.254
'Cgp'	25.00	25.00
'Cstr'	2.6 ± 1.1	2.563
'Cubar'	8.9 ± 1.1	8.945
'Cuv'	5.05 ± 0.11	5.048
'Dubar'	8.3 ± 3.6	8.326
'Euv'	13.7 ± 1.8	13.69
'Mw_fit'	80.40	80.421 ± 0.013
'rs'	1.51 ± 0.29	1.515

The recovered mass is 21 MeV higher than the pseudodata value! → the mass shift is not produced by the systematics on the unfolding of the LowPileup data.

Approach 2: only LowPileup

Parameter	PDFonly	mWonly
'Adbar'	0.2892 ± 0.0025	0.2892
'Adv'	1.0000	1.0000
'Ag'	1.0000	1.0000
'Agp'	0.182 ± 0.024	0.1815
'Auv'	1.0000	1.0000
'Bdbar'	0.0403 ± 0.0025	0.04028
'Bdv'	0.688 ± 0.014	0.6884
'Bg'	-0.451 ± 0.014	-0.4514
'Bgp'	-0.597 ± 0.013	-0.5974
'Buv'	0.744 ± 0.010	0.7445
'Cdbar'	6.57 ± 0.79	6.572
'Cdv'	3.64 ± 0.11	3.635
'Cg'	2.226 ± 0.062	2.226
'Cgp'	25.00	25.00
'Cstr'	8.6 ± 2.7	8.608
'Cubar'	8.30 ± 0.50	8.299
'Cuv'	4.789 ± 0.043	4.789
'Dubar'	-0.96 ± 0.57	-0.9633
'Euv'	9.85 ± 0.44	9.850
'Mw_fit'	80.40	80.400 ± 0.013
'rs'	0.57 ± 0.12	0.5706

Almost perfectly recovered the nominal value of m_W (although the PDF parameter errors are faulty)

Closure test, including the LowPileup unfolding systematics (agreement to data)

Approach 1: HERA + LowPileup

Dataset	PDFonly	mWonly
HERA1+2 NCep 820	57 / 61	57 / 61
HERA1+2 NCep 460	199 / 177	199 / 177
MC XS low- μ Wplusmunu pTl 5 TeV	1.7 / 34	0.82 / 34
HERA1+2 CCep	47 / 39	47 / 39
HERA1+2 NCem	228 / 159	228 / 159
HERA1+2 CCem	81 / 42	81 / 42
HERA1+2 NCep 575	193 / 221	193 / 221
MC XS low- μ Wminusmunu pTl 5 TeV	1.3 / 34	0.85 / 34
HERA1+2 NCep 920	357 / 317	357 / 317
MC XS low- μ Wplusenu pTl 5 TeV	1.4 / 34	0.55 / 34
MC XS low- μ Wminusenu pTl 5 TeV	1.2 / 34	0.68 / 34
Correlated χ^2	57	57
Log penalty χ^2	-32.65	-32.55
Total χ^2 / dof	1192 / 1136	1189 / 1151
χ^2 p-value	0.12	0.21

Approach 2: only LowPileup

Dataset	PDFonly	mWonly
MC XS low- μ Wplusmunu pTl 5 TeV	0.050 / 34	0.050 / 34
MC XS low- μ Wminusmunu pTl 5 TeV	0.040 / 34	0.040 / 34
MC XS low- μ Wplusenu pTl 5 TeV	0.080 / 34	0.080 / 34
MC XS low- μ Wminusenu pTl 5 TeV	0.040 / 34	0.040 / 34
Correlated χ^2	1.2	1.2
Log penalty χ^2	-2.68	-2.67
Total χ^2 / dof	-1.3 / 120	-1.3 / 135
χ^2 p-value	0.00	0.00

Closure test, including the LowPileup unfolding systematics (parameter values)

Approach 1: HERA + LowPileup

Parameter	PDFonly	mWonly
'Adbar'	0.0938 ± 0.0085	0.09380
'Adv'	1.0000	1.0000
'Ag'	1.0000	1.0000
'Agp'	0.421 ± 0.071	0.4212
'Auv'	1.0000	1.0000
'Bdbar'	-0.174 ± 0.015	-0.1740
'Bdv'	0.597 ± 0.047	0.5971
'Bg'	-0.54 ± 0.11	-0.5415
'Bgp'	-0.603 ± 0.076	-0.6029
'Buv'	0.742 ± 0.025	0.7425
'Cdbar'	3.1 ± 1.6	3.082
'Cdv'	2.96 ± 0.31	2.958
'Cg'	3.71 ± 0.84	3.709
'Cgp'	25.00	25.00
'Cstr'	6.0 ± 3.4	6.028
'Cubar'	8.6 ± 1.0	8.596
'Cuv'	5.014 ± 0.097	5.014
'Dubar'	8.3 ± 3.4	8.313
'Euv'	13.2 ± 1.7	13.18
'Mw_fit'	80.40	80.428 ± 0.016
'rs'	1.66 ± 0.33	1.657

Approach 2: only LowPileup

Parameter	PDFonly	mWonly
'Adbar'	0.283 ± 0.017	0.2832
'Adv'	1.0000	1.0000
'Ag'	1.0000	1.0000
'Agp'	0.239 ± 0.067	0.2387
'Auv'	1.0000	1.0000
'Bdbar'	0.061 ± 0.013	0.06128
'Bdv'	0.649 ± 0.043	0.6486
'Bg'	-0.427 ± 0.037	-0.4267
'Bgp'	-0.491 ± 0.043	-0.4910
'Buv'	0.730 ± 0.024	0.7302
'Cdbar'	6.4 ± 1.5	6.372
'Cdv'	3.63 ± 0.50	3.630
'Cg'	2.55 ± 0.39	2.546
'Cgp'	25.00	25.00
'Cstr'	8.0 ± 4.0	7.952
'Cubar'	9.9 ± 1.2	9.861
'Cuv'	4.57 ± 0.24	4.567
'Dubar'	2.3 ± 1.6	2.275
'Euv'	7.6 ± 1.7	7.613
'Mw_fit'	80.40	80.401 ± 0.016
'rs'	0.75 ± 0.29	0.7498

The recovered mass is 28 MeV higher than the pseudodata value! → Repeating exercise with Pseudodata at $m_W = \{80.3, 80.5\}$ gives the same shift.

Almost perfectly recovered the nominal value of m_W (although the PDF parameter errors are faulty)

- It is possible to perform a successful closure test if it is performed exclusively on the LowPileup data
- Systematics have a minor effect on the central value of the fitted value of m_W , well within uncertainties
- The addition of more datasets, as it will also be shown in the next slides, seems to bias (if that's the correct word?) the value of m_W by an amount larger than the statistical uncertainty
→ this point is not yet well understood, seeking to do additional tests
- Since we are interested in measuring the correlation between the PDF and m_W , how much should we care about the size of the bias?

Combined m_W + PDF fits

Reminder of the final goal

We want to perform a combined fit using the same data as ATLASpdf21 [2112.11266] + the LowPileup data.

Since I have all the relevant data, multiple combined fits can be performed in order to see the behaviour of m_W and its correlation to other parameters.

The fits are done in an incremental way, adding each dataset on top of all the previous ones:

- HERA [1506.06042] + LowPileup
- + WZ production [1612.03016]
- + ttbar8TeV [1511.04716, 4 correlation files]
- + Vjets [1711.03296 (W) and 1907.06728 (Z)]
- + etc...

Combined m_W + PDF fit results (work in progress, pseudodata @ $m_W = 80.4$ GeV)

Parameter	HERA+LowPileup	+epWZ16	+ttbar	+Vjets	+Wxs8TeV	ATLASpdf21+LowPileup
'Adbar'	0.0975 ± 0.0092	0.1029 ± 0.0063	0.1030 ± 0.0057	0.0975 ± 0.0051	0.348 ± 0.062	0.0862 ± 0.0071
'Adv'	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
'Ag'	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
'Agp'	0.413 ± 0.068	0.360 ± 0.059	0.351 ± 0.039	0.318 ± 0.028	0.239 ± 0.027	0.170 ± 0.050
'Aubar'	-	-	-	-	0.0928 ± 0.0048	0.1012 ± 0.0066
'Auv'	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
'Bdbar'	-0.171 ± 0.015	-0.169 ± 0.010	-0.1695 ± 0.0097	-0.1776 ± 0.0092	0.019 ± 0.030	-0.185 ± 0.017
'Bdv'	0.581 ± 0.039	0.679 ± 0.027	0.678 ± 0.027	0.663 ± 0.024	0.531 ± 0.017	0.396 ± 0.065
'Bg'	-0.53 ± 0.12	-0.594 ± 0.090	-0.592 ± 0.074	-0.629 ± 0.059	-0.553 ± 0.069	-0.628 ± 0.053
'Bgp'	-0.596 ± 0.080	-0.644 ± 0.067	-0.644 ± 0.053	-0.675 ± 0.042	-0.650 ± 0.036	-0.709 ± 0.038
'Bstr'	-	-	-	-	0.341 ± 0.080	-0.132 ± 0.040
'Bubar'	-	-	-	-	-0.2184 ± 0.0095	-0.185 ± 0.012
'Buv'	0.735 ± 0.023	0.737 ± 0.012	0.735 ± 0.011	0.723 ± 0.011	0.710 ± 0.011	0.713 ± 0.018
'Cdbar'	4.4 ± 2.1	2.46 ± 0.50	2.45 ± 0.48	1.84 ± 0.33	27.7 ± 3.6	1.88 ± 0.14
'Cdv'	2.84 ± 0.24	3.69 ± 0.25	3.70 ± 0.25	3.72 ± 0.21	2.66 ± 0.15	4.75 ± 0.21
'Cg'	3.67 ± 0.85	2.69 ± 0.67	2.61 ± 0.46	2.24 ± 0.35	2.11 ± 0.31	4.65 ± 0.36
'Cgp'	25.00	25.00	25.00	25.00	25.00	25.00
'Cstr'	5.1 ± 2.7	10.3 ± 1.7	10.6 ± 1.7	11.4 ± 1.4	7.0 ± 1.0	9.3 ± 1.0
'Cubar'	8.2 ± 1.0	6.0 ± 1.3	6.0 ± 1.3	7.0 ± 1.0	4.39 ± 0.38	4.44 ± 0.36
'Cuv'	5.015 ± 0.096	4.848 ± 0.076	4.852 ± 0.076	4.926 ± 0.084	4.904 ± 0.078	4.781 ± 0.094
'Ddv'	-	-	-	-	-	11.2 ± 5.0
'Dg'	-	-	-	-	-	9.6 ± 4.1
'Dubar'	7.7 ± 3.3	2.0 ± 1.8	2.0 ± 1.8	3.5 ± 1.7	-	-
'Duv'	-	-	-	-	-	0.06 ± 0.36
'Euv'	13.7 ± 1.7	10.92 ± 0.92	11.04 ± 0.92	11.97 ± 0.93	12.54 ± 0.95	11.05 ± 0.95
'Mw_fit'	80.446 ± 0.020	80.432 ± 0.018	80.432 ± 0.018	80.421 ± 0.017	80.454 ± 0.017	80.420 ± 0.018
rs	1.54 ± 0.51	1.567 ± 0.666	1.566 ± 0.666	1.571 ± 0.666	2.72 ± 0.65	1.55 ± 0.57
Fit status	converged	converged	converged	converged	converged	converged
Uncertainties	migrad-hesse	migrad-hesse	migrad-hesse	pos-def-forced	migrad-hesse	migrad-hesse

Global Correlation 0.61816 0.46116 0.45988 0.41180 0.44756 0.48443

- We seek to use the results of the ongoing W-boson low-pileup, which will give differential cross-sections as a function of p_i^T
- We have a working framework (based on xFitter) which is capable of minimizing the a PDF along m_W given a parametrization determined off-line
 - for now, using pseudodata @ $m_W = 80.4$ GeV
- Preliminary closure tests are capable of recovering the input mass if fitting ONLY the p_i^T cross-sections
- Addition of further datasets seems to bias the mass value by 20-50 MeV (with an uncertainty of 18 MeV)
- The correlation between m_W and the rest of the PDF parameters seems relatively stable, with values around 45%

Backup

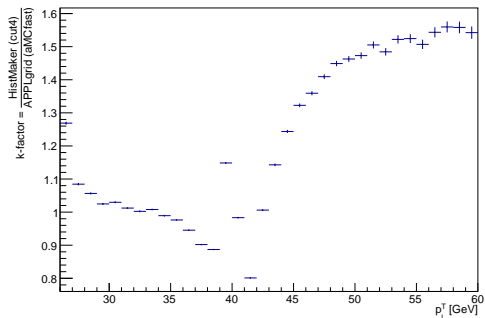
Combined fit motivation: effective leptonic weak mixing angle uncertainty breakdown

Channel	ee_{CC}	$\mu\mu_{CC}$	ee_{CF}	$ee_{CC} + \mu\mu_{CC}$	$ee_{CC} + \mu\mu_{CC} + ee_{CF}$
Central value	0.23148	0.23123	0.23166	0.23119	0.23140
Uncertainties					
Total	68	59	43	49	36
Stat.	48	40	29	31	21
Syst.	48	44	32	38	29
Uncertainties in measurements					
PDF (meas.)	8	9	7	6	4
p_T^Z modelling	0	0	7	0	5
Lepton scale	4	4	4	4	3
Lepton resolution	6	1	2	2	1
Lepton efficiency	11	3	3	2	4
Electron charge misidentification	2	0	1	1	< 1
Muon sagitta bias	0	5	0	1	2
Background	1	2	1	1	2
MC. stat.	25	22	18	16	12
Uncertainties in predictions					
PDF (predictions)	37	35	22	33	24
QCD scales	6	8	9	5	6
EW corrections	3	3	3	3	3

Uncertainty breakdown for $\sin^2 \theta_W$, as obtained in the study [ATL-CONF-2018-037](#).

m_W in W inclusive $d\sigma/dp_T^T$ at 5 TeV: k-factor and m_W dependence

k-factors
(these values remain constant)



m_W dependence
(used to fit the value of factor p_T^T in each bin)

