QCD + EW fits with xFitter

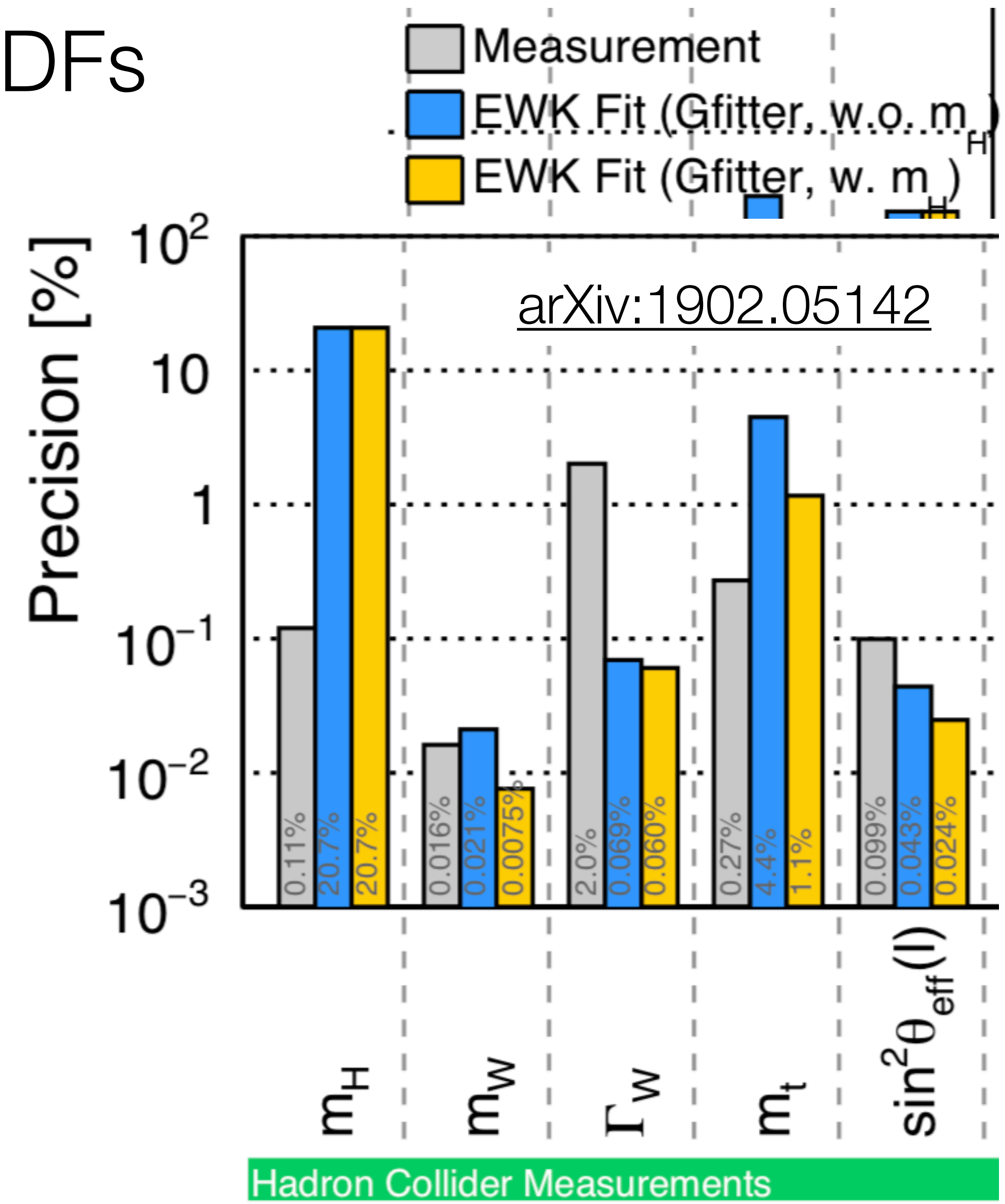
xFitter workshop, DESY-Hamburg

Feb. 27th, 2020

Simone Amoroso

PDFs treatment in EW measurements

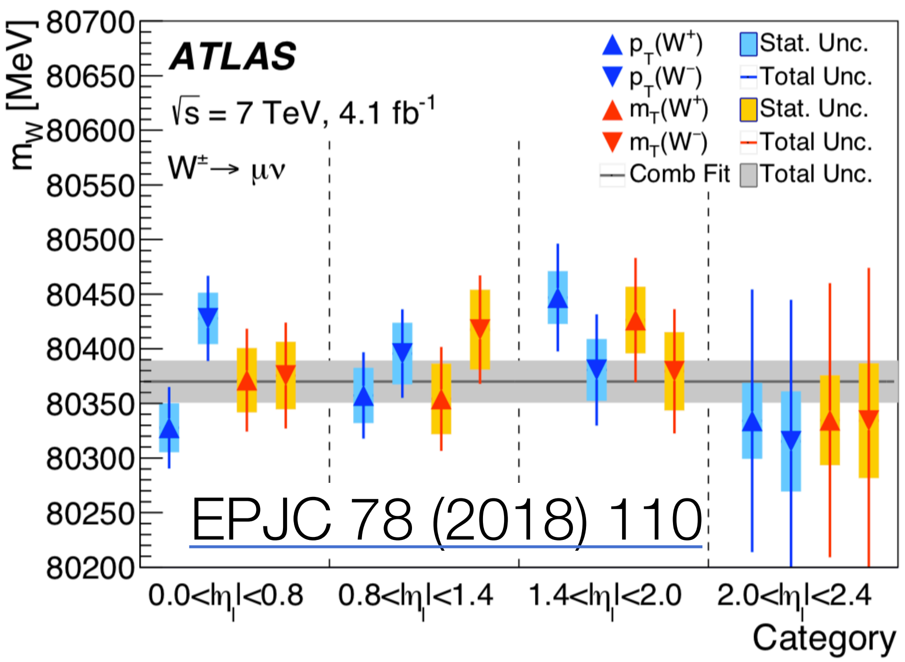
* At the level of precision the LHC has achieved PDF uncertainties are dominating precision EW measurements
* Most of these measurements use some form of in-situ constraint of the PDFs
* PDFs profiling has been used in recent measurement to reduce PDFs uncertainties on EW parameters
* I will argue why PDF profiling will not be enough for the level of precision we need to achieve at the HL-LHC



W-mass and PDFs

- ATLAS W-mass measurement significantly affected by PDFs
  - Large W polarisation uncertainty impacting p_T^l distribution
- BLUE combination of different categories reduced PDF uncertainties as W^+ and W^- are anti-correlated

W-boson charge	W^+		W^-		Combined	
Kinematic distribution	p_T^ℓ	m_T	p_T^ℓ	m_T	p_T^ℓ	m_T
δm_W [MeV]						
Fixed-order PDF uncertainty	13.1	14.9	12.0	14.2	8.0	8.7
Total	15.9	18.1	14.8	17.2	11.6	12.9



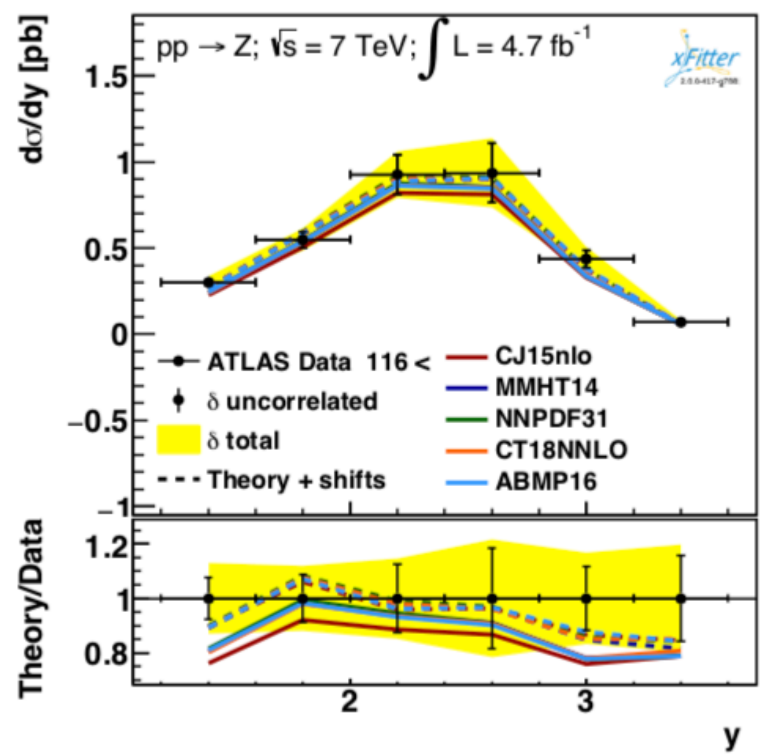
EPJC 78 (2018) 110

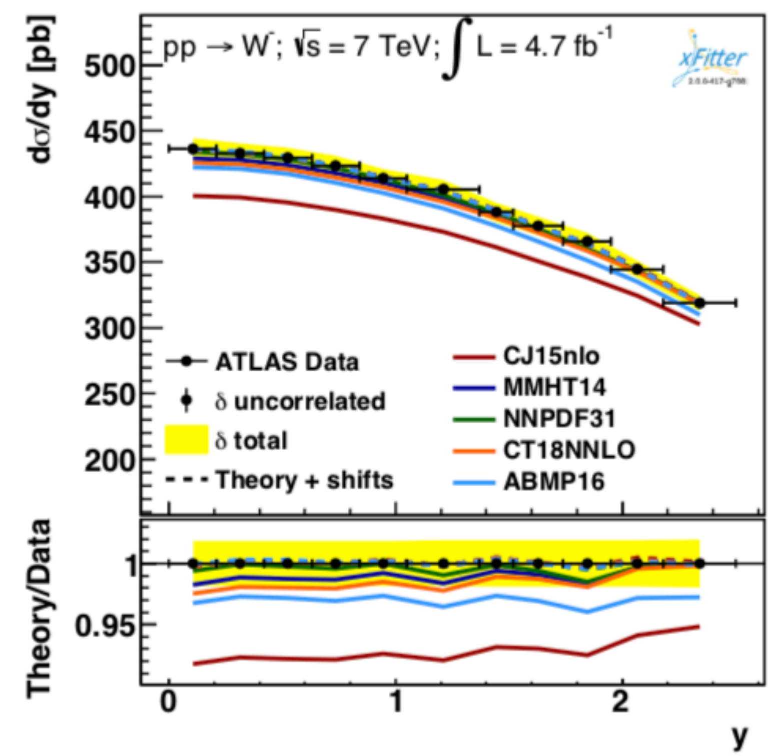
- Avoided an explicit PDF profiling including additional data
  - Issues with mixing constraints from unfolded (Z-rapidity) and reco-level (W p_T^l, m_T) distributions and impact of resummation corrections

ATLAS-Tevatron W-mass

- Ongoing efforts towards an ATLAS+Tevatron m_W combination
- Trying to move measurements to a common PDF set
- Several PDF sets will be quoted
- Which one should be taken as nominal?
- Started an xFitter benchmarking of modern PDFs against all Drell-Yan data in xFitter

1612.03016

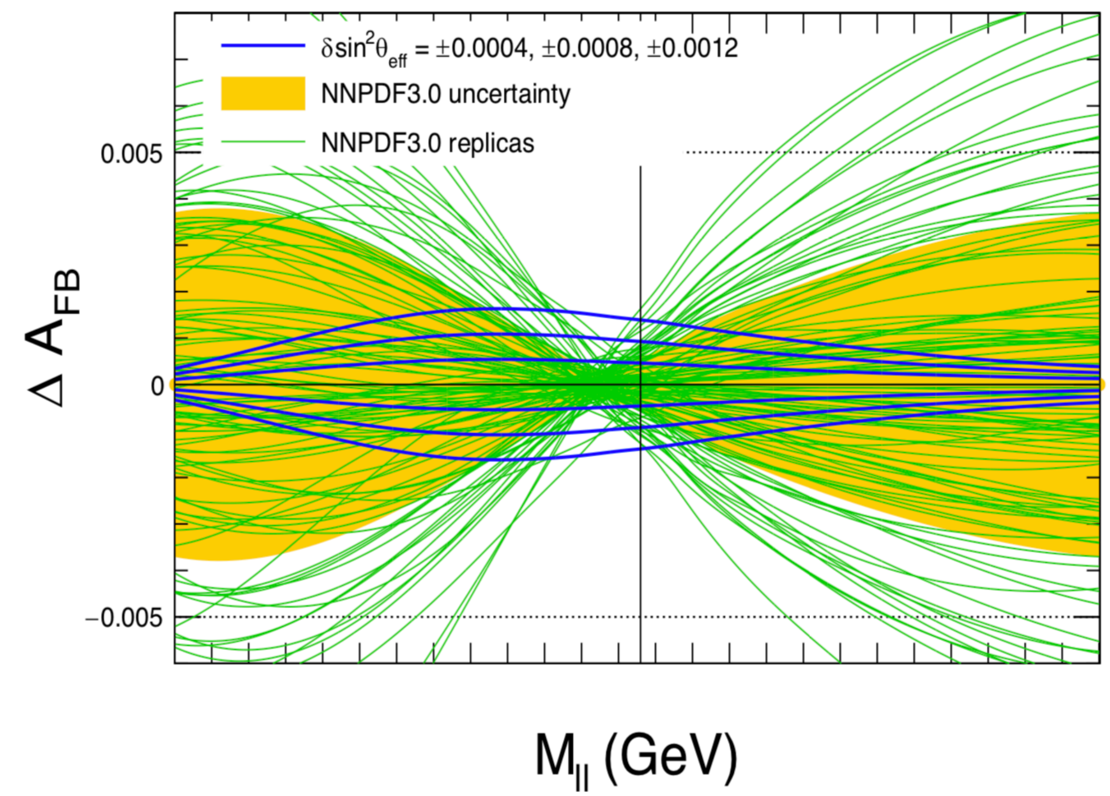




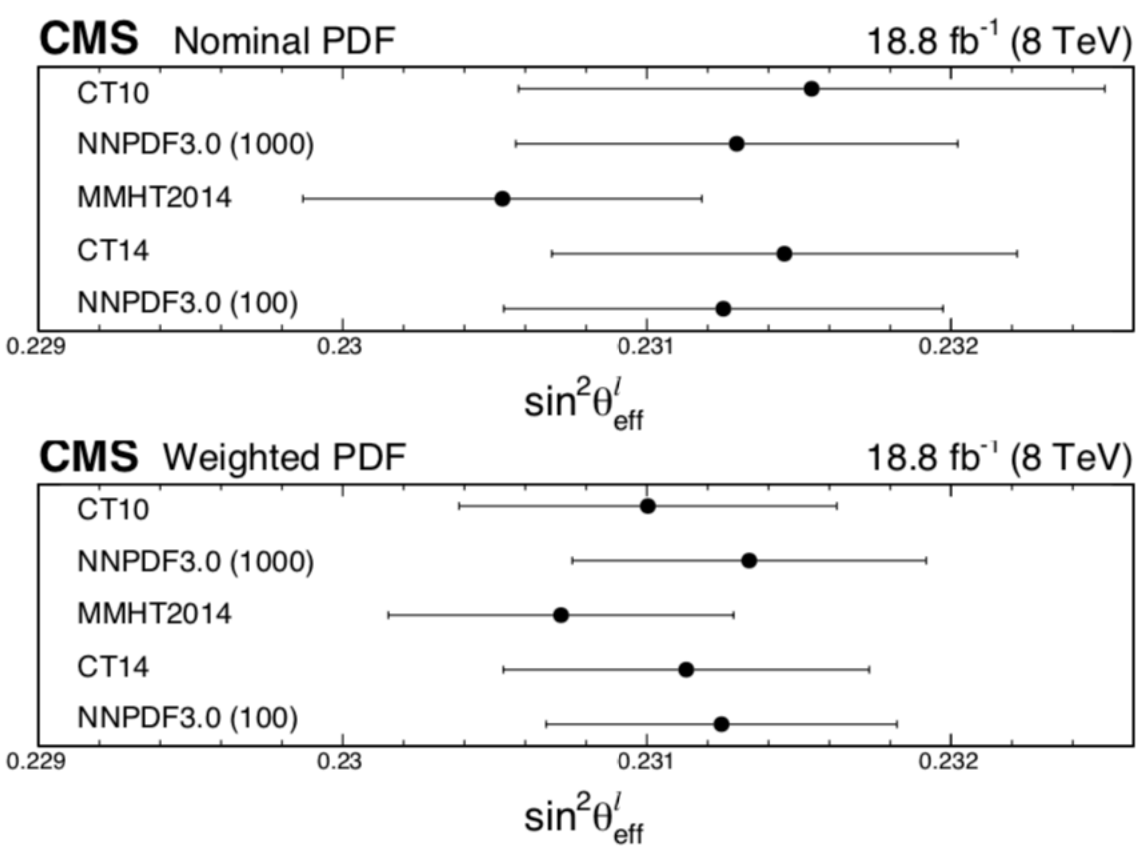
Dataset	CJ15nlo	MMHT14	NNPDF31	CT18NNLO	ABMP16
ATLAS low mass Z rapidity 2011	26 / 6	18 / 6	14 / 6	12 / 6	21 / 6
ATLAS peak CC Z rapidity 2011	52 / 12	21 / 12	12 / 12	16 / 12	24 / 12
ATLAS peak CF Z rapidity 2011	16 / 9	11 / 9	11 / 9	10 / 9	9.2 / 9
ATLAS high mass CC Z rapidity 2011	7.7 / 6	6.1 / 6	5.8 / 6	5.9 / 6	6.1 / 6
ATLAS high mass CF Z rapidity 2011	4.6 / 6	5.5 / 6	4.7 / 6	4.8 / 6	4.5 / 6
ATLAS W- lepton rapidity 2011	17 / 11	8.4 / 11	8.7 / 11	9.1 / 11	10 / 11
ATLAS W+ lepton rapidity 2011	16 / 11	11 / 11	11 / 11	10 / 11	13 / 11
Correlated χ^2	118	50	31	40	50
Log penalty χ^2	-9.09	-3.32	-2.45	-3.66	-4.22
Total χ^2 / dof	247 / 61	127 / 61	95 / 61	104 / 61	134 / 61
χ^2 p-value	0.00	0.00	0.00	0.00	0.00

Weak mixing angle

- PDF uncertainties constrained in A_{FB}, A_4 interpretation exploiting correlations in m^{ll}, y^{ll}
  - CMS - Bayesian reweighing
  - ATLAS - Hessian profiling
- Yet PDFs remain the largest source of uncertainty



Eur. Phys. J. C (2018) 78: 701



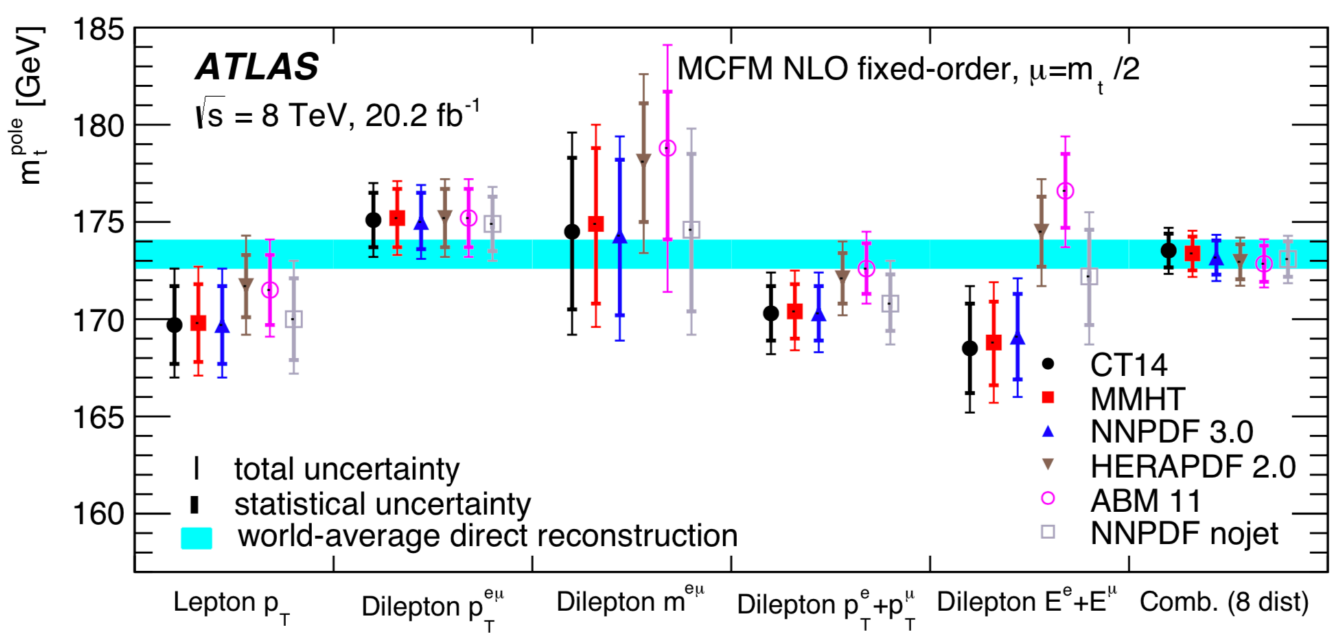
ATLAS-CONF-2018-037

	CT10	CT14	MMHT14	NNPDF31
$\sin^2\theta^\ell_{\text{eff}}$	0.23118	0.23141	0.23140	0.23146
	Uncertainties in measurements			
Total	39	37	36	38
Stat.	21	21	21	21
Syst.	32	31	29	31

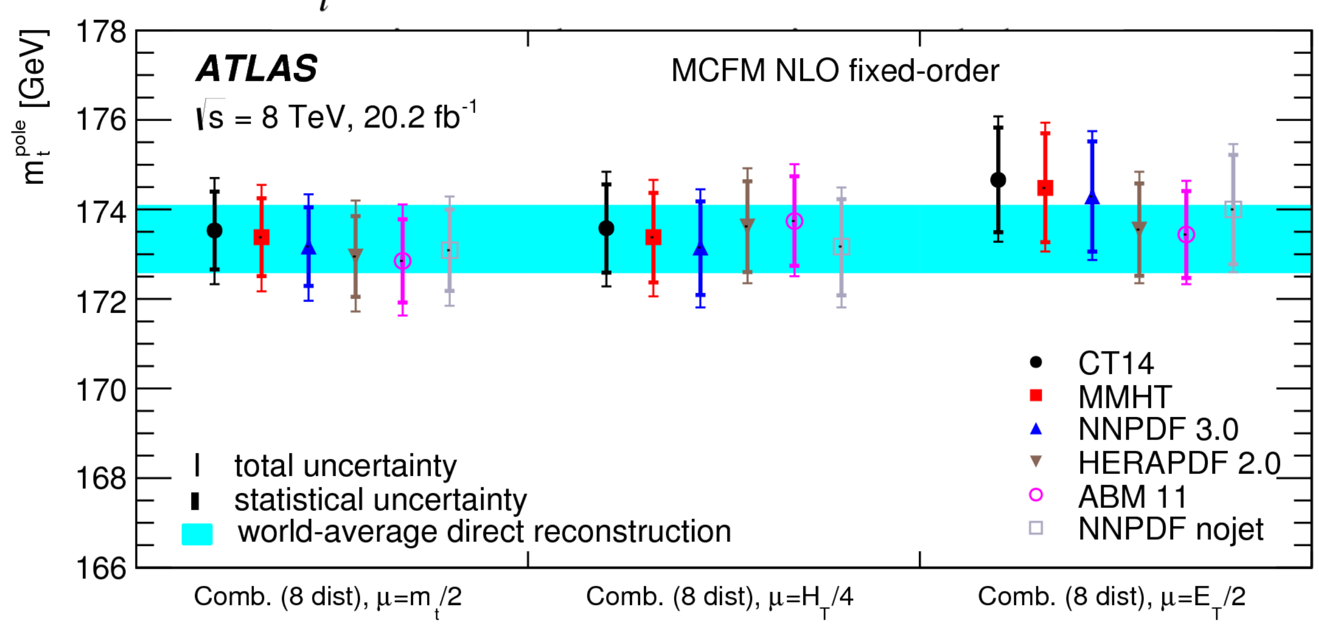
Top mass

- Mass dependence of predicted top-pair (differential) cross sections allows for indirect determinations of m_t
  - Reduced experimental uncertainties and less ambiguities on m_t, but large dependence on theory (scales + PDFs)
  - Typically use in-situ profiling of PDFs and QCD scales
- Still large uncertainties, but exp. precision will improve and NNLO theory is now available; PDFs will become the bottleneck

EPJC77 (2017) 804



$$m_t^{\text{pole}} = 173.2 \pm 0.9 \pm 0.8 \pm 1.2\,\text{GeV}$$



PDF PROFILING VS FITS

- PDF profiling/reweighting is not a replacement for a full fit
- Profiling is only strictly valid in the vicinity of the χ^2 minimum; it will fail if the impact of the new data is too large
  - Cannot account for methodological changes, such as modification in the PDFs parametrisation or to the theory
- Typically assume a $\Delta\chi^2 = 1$
  - The impact of the new data estimated with profiling will generally be different than including the same data in a new fit.
- The input PDF might have unknown correlations with the new data
  - e.g. CT18A includes ATLAS 7 TeV W-asymmetry which would statistically overlap with W-mass dataset

PDF+EW fits

- A simultaneous fit of PDFs and determination of EW parameters would overcome the issues with profiling
  - Biases from the choice of theory, parametrisation, input data and statistical treatment can be evaluated and exposed
- And allow a simultaneous determination of multiple EW parameters, properly accounting for all correlations
- A PDF+EW fit would also provide additional benefits:
  - Can further reduce PDF uncertainties by including data providing additional constraints
  - Provides a framework for combinations of measurements at different $\sqrt{s}$, from different experiments (LHCb, HERA, LHeC?)
  - Evaluate consistency/compatibility of different measurements
  - By explicitly considering only well understood/well predicted data, might hope for a statistically well-defined uncertainty (tolerance of 1)

HERA QCD+EW FITS

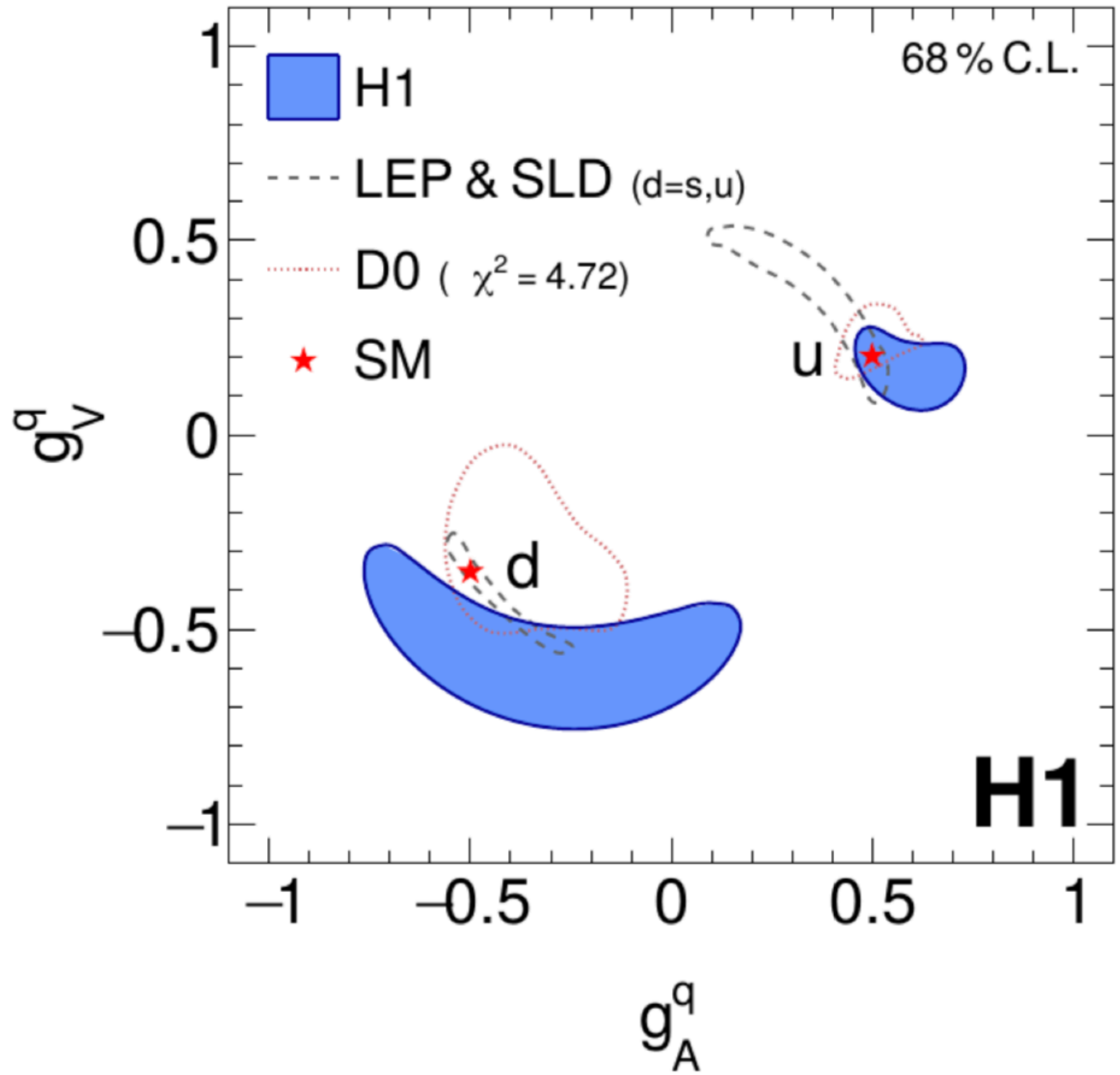
* Long tradition of PDF fits from HERA Collaborations
* These datasets are included in global PDFs, assuming EW parameters are SM-like. Standard PDFs cannot be used for EW parameter interpretations
* Full fit of PDF+EW parameters as the only way to determine EW parameters in an unbiased way
  * Performed at NNLO QCD + NLO EW (in the OS scheme)
  * Using the standard HERAPDF methodology and parametrisation
  * Several combinations of parameters are fitted

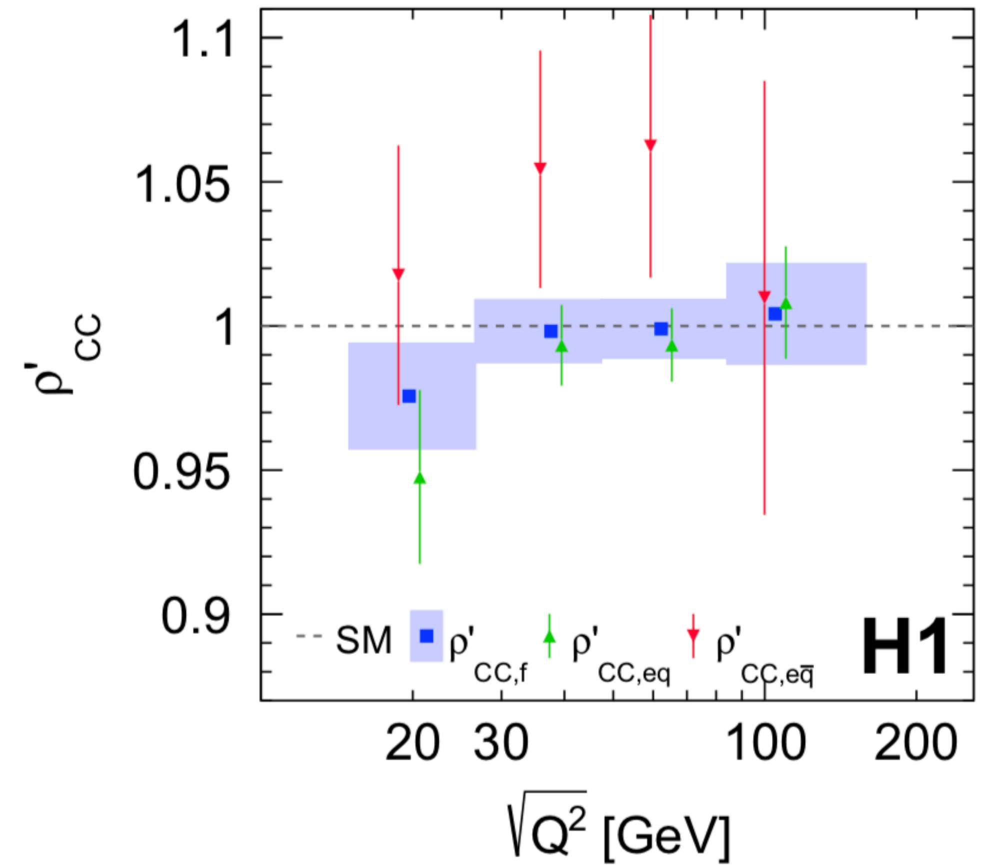
e.g. H1 EPJC(2018)78:777

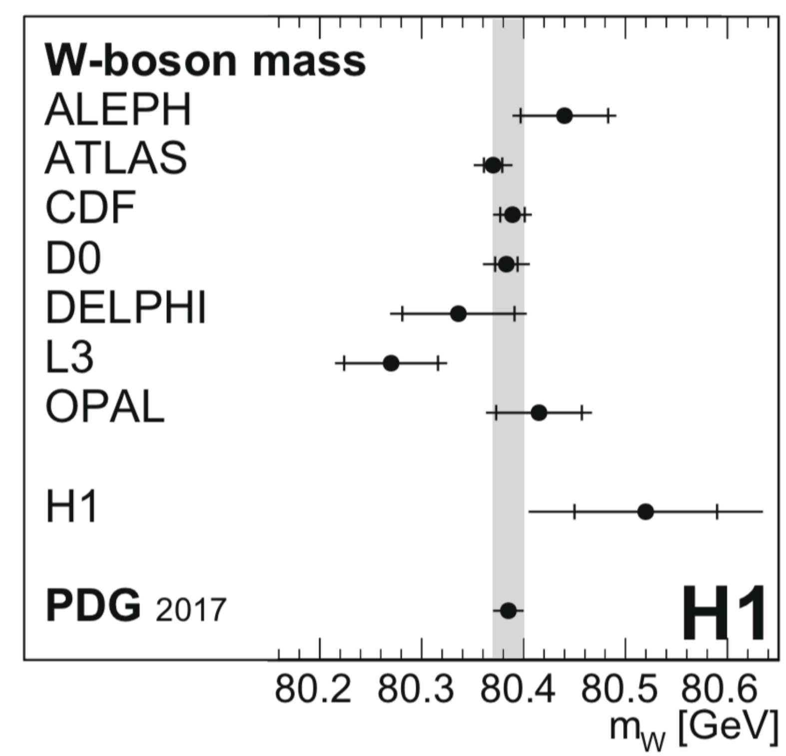
Fit parameters	Result		Independent input parameters
m_W+PDF	m_W	$= 80.520 \pm 0.070_{\mathrm{stat}} \pm 0.055_{\mathrm{syst}} \pm 0.073_{\mathrm{PDF}}$ GeV	α, m_Z, m_t, m_H, m_f
m_W^{prop}+PDF	m_W^{prop}	$= 80.62 \pm 0.67_{\mathrm{stat}} \pm 0.17_{\mathrm{syst}} \pm 0.38_{\mathrm{PDF}}$ GeV	α, m_W, m_Z, m_t, m_H, m_f
$m_W^{(G_F, m_W)}$+PDF	$m_W^{(G_F, m_W)}$	$= 82.05 \pm 0.51_{\mathrm{stat}} \pm 0.44_{\mathrm{syst}} \pm 0.37_{\mathrm{PDF}}$ GeV	α, G_F, m_t m_H, m_f
m_Z+PDF	m_Z	$= 91.084 \pm 0.064_{\mathrm{stat}} \pm 0.050_{\mathrm{syst}} \pm 0.070_{\mathrm{PDF}}$ GeV	α, m_W, m_t, m_H, m_f
m_t+PDF	m_t	$= 154 \pm 10_{\mathrm{stat}} \pm 12_{\mathrm{syst}} \pm 15_{\mathrm{PDF}} \pm 15_{m_W}$ GeV	α, m_W, m_Z, m_H, m_f

And interpretations

- With PDF uncertainties under control many possible interpretations can be derived







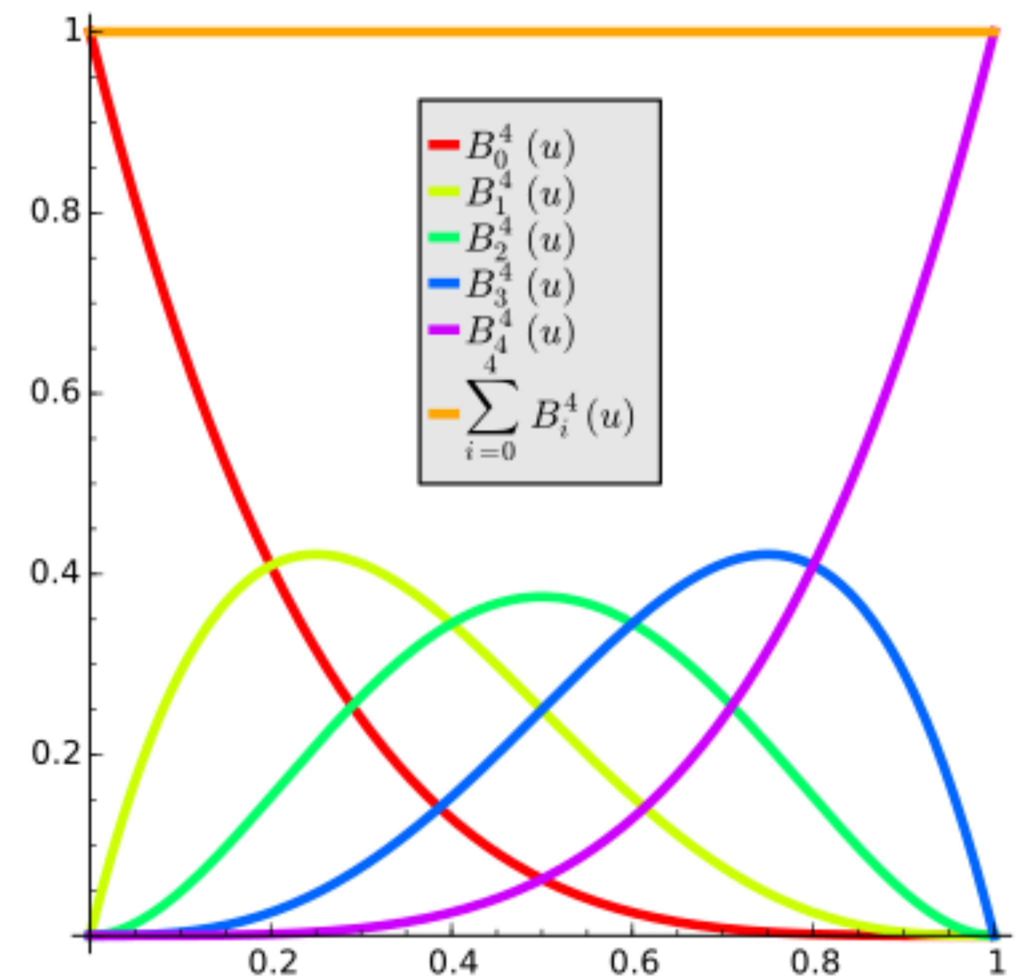
- Similar fits have been used to determine PDF+α_S, PDFs+m_b,m_c

PDF+EW fits in xFitter

* xFitter already established tool to perform PDF fits studies within the HERA and LHC communities
* Need to ensure a coherent treatment of QCD and EW theory corrections in the predictions used
    * Can a unified EW scheme be used for DIS and pp(ppbar)?
* Need to incorporate additional EW parameters obtained from dedicated predictions in the chi2
* The traditional HERAPDF parametrisation is likely too restrictive and would bias EW parameter extractions
* Using Minuit might also not be the best approach if trying to determine more than o(10) parameters

Parametrisation

- With the amount of data the standard HERA based PDF parametrisation in xFitter becomes a bottleneck.
- I have recently started implementing a parametrisation based on Bernstein polynomials, following what is done in the CT fits
  - Each parameter peaks at a given x
  - Low-correlations among parameters
  - Better uncertainty estimates (a data point at high-x won't constrain low-x PDFs)
- Available in an xFitter branch, pending some more validation

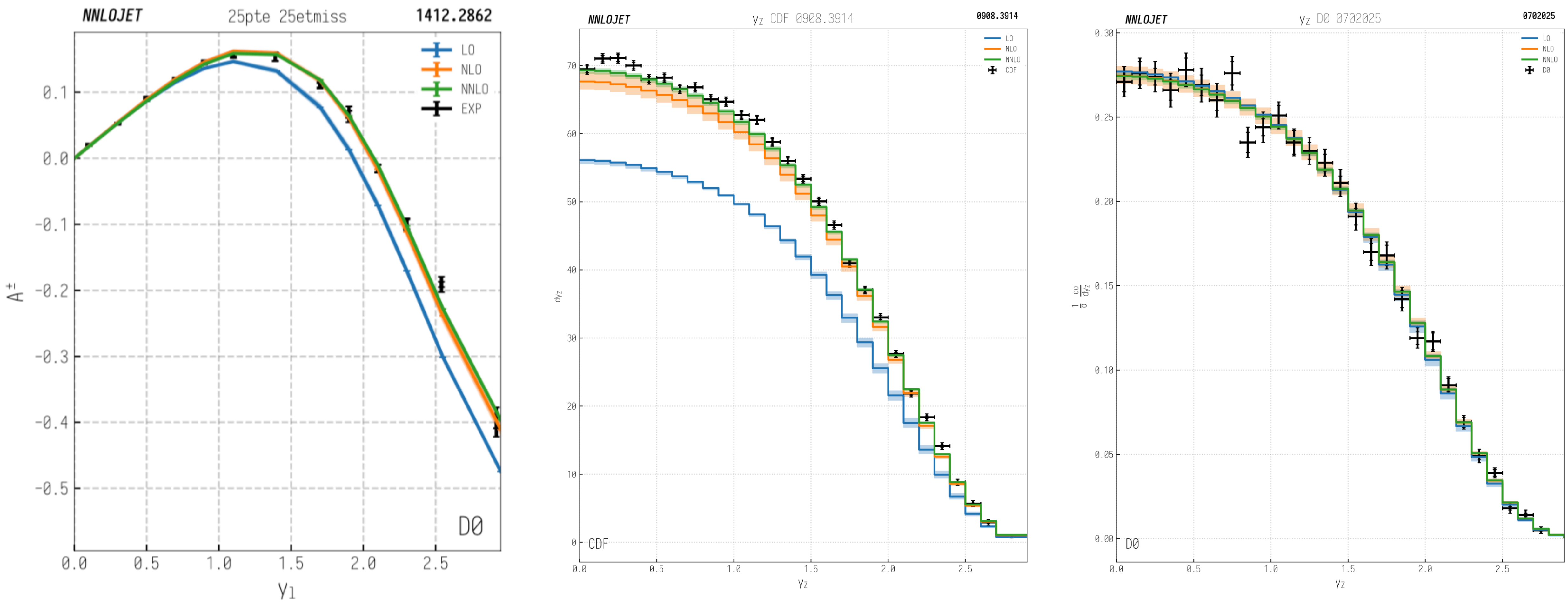


Datasets

exp.	obs	channel	$\sqrt{s}$	lumi	pts	ref
CDF	$d\sigma/dy$	$Z \to ee$	1.8 TeV	108 pb^{-1}	28	hep-ex/0006025
CDF	A_W	$e\nu$	1.96 TeV	1fb $^{-1}$	13	arXiv:0901.2169
CDF	$d\sigma/dy$	ee	1.96 TeV	2.1 fb^{-1}	28	arXiv:0908.3914
D0	$d\sigma/dy$	ee	1.96 TeV	0.4 fb^{-1}	28	hep-ex/0702025
D0	A_W	$\mu\nu$	1.96 TeV	7.3 fb^{-1}	12	arXiv:1309.2591
D0	A_W	$e\nu$	1.96 TeV	9.7 fb^{-1}	14	arXiv:1312.2895
D0	A_l	$e\nu$	1.96 TeV	9.7 fb^{-1}	13	arXiv:1412.2862
ATLAS	Z, W	-	7 TeV	4.7 fb^{-1}	61	arXiv:1612.03016
ATLAS	Z	$Z \to \mu\mu, ee$	8 TeV	20.2 fb^{-1}	504	arXiv:1710:05167
ATLAS	W	$W \to \mu\nu$	8 TeV	20.2 fb^{-1}	22	arXiv:1710:05167
ATLAS	W,Z	e,μ	5.02 TeV	25 pb^{-1}	5+22	arXiv:1810.08424
ATLAS	W,Z	e,μ	2.76 TeV	4 pb^{-1}	3	arXiv:1907.03567
CMS	Z	e,μ	7 TeV	4.5 fb^{-1}	144	arXiv:1310.7291
CMS	W	μ	7 TeV	4.7 fb^{-1}	11	arXiv:1312.6283
CMS	Z	e,μ	8 TeV	19.7 fb^{-1}	120	arXiv:1412.1115
CMS	W	$e\nu$,$\mu\nu$	8 TeV	18.8 fb^{-1}	22	arXiv::1603.01803
CMS	Z	e,μ	13 TeV	35.9 fb^{-1}	12	arXiv:1909.04133
LHCb	Z	$Z \to ee$	7 TeV	0.94 $^{-1}$	9	arXiv:1212.4620
LHCb	W	$W \to \mu\nu$	7 TeV	1 fb^{-1}	16	arXiv:1408.4354
LHCb	Z	$Z \to ee$	8 TeV	2 fb^{-1}	17	arXiv:1503.00963
LHCb	Z	$Z \to \mu\mu$	7 TeV	1 fb^{-1}	17	arXiv:1505.07024
LHCb	Z	$Z \to \mu\mu$	8 TeV	2 fb^{-1}	17	arXiv:1511.08039
LHCb	Z	$Z \to \mu\mu, ee$	13 TeV	294 pb^{-1}	35	arXiv:1607.06495
CDF	$A_{FB}(m)$	$Z \to ee$	1.8 TeV	108 pb^{-1}	11	hep-ex/0106047
CDF	$A_{FB}(m)$	$Z \to \mu\mu$	1.96 TeV	9.2 fb^{-1}	15	arXiv:1402.2239
CDF	$A_{FB}(m)$	$Z \to ee$	1.96 TeV	9.4 fb^{-1}	15	arXiv:1605.02719
D0	$A_{FB}(m)$	$Z \to ee$	1.96 TeV	5 fb^{-1}	15	arXiv:1104.4590
LHCb	$A_{FB}(m)$	$Z \to \mu\mu$	7,8 TeV	1-2 fb^{-1}	26	arXiv:1509.07645
CMS	$A_{FB}(m,y)$	$Z \to \mu\mu, ee$	7 TeV	5 fb^{-1}	40	arXiv:1207.3973
CMS	$A_{FB}(m,y)$	$Z \to \mu\mu, ee$	8 TeV	19.7 fb^{-1}	63	arXiv:1207.3973
ATLAS	$A_{FB}(m,y)$	$Z \to \mu\mu, ee$	7 TeV	4.7 fb^{-1}	48	arXiv:1503.03709

And theory predictions

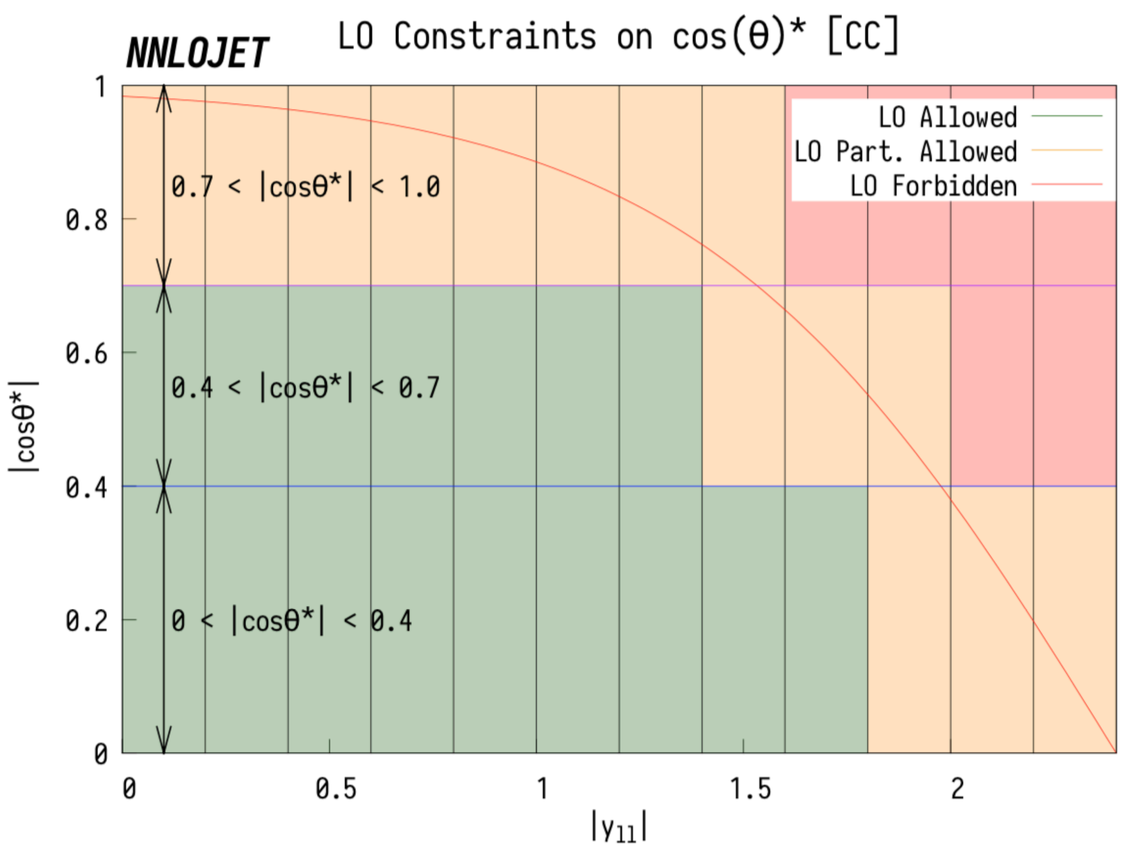
- Collaboration with NNLOJET to get NNLO QCD predictions for Tevatron and LHC Drell-Yan measurements

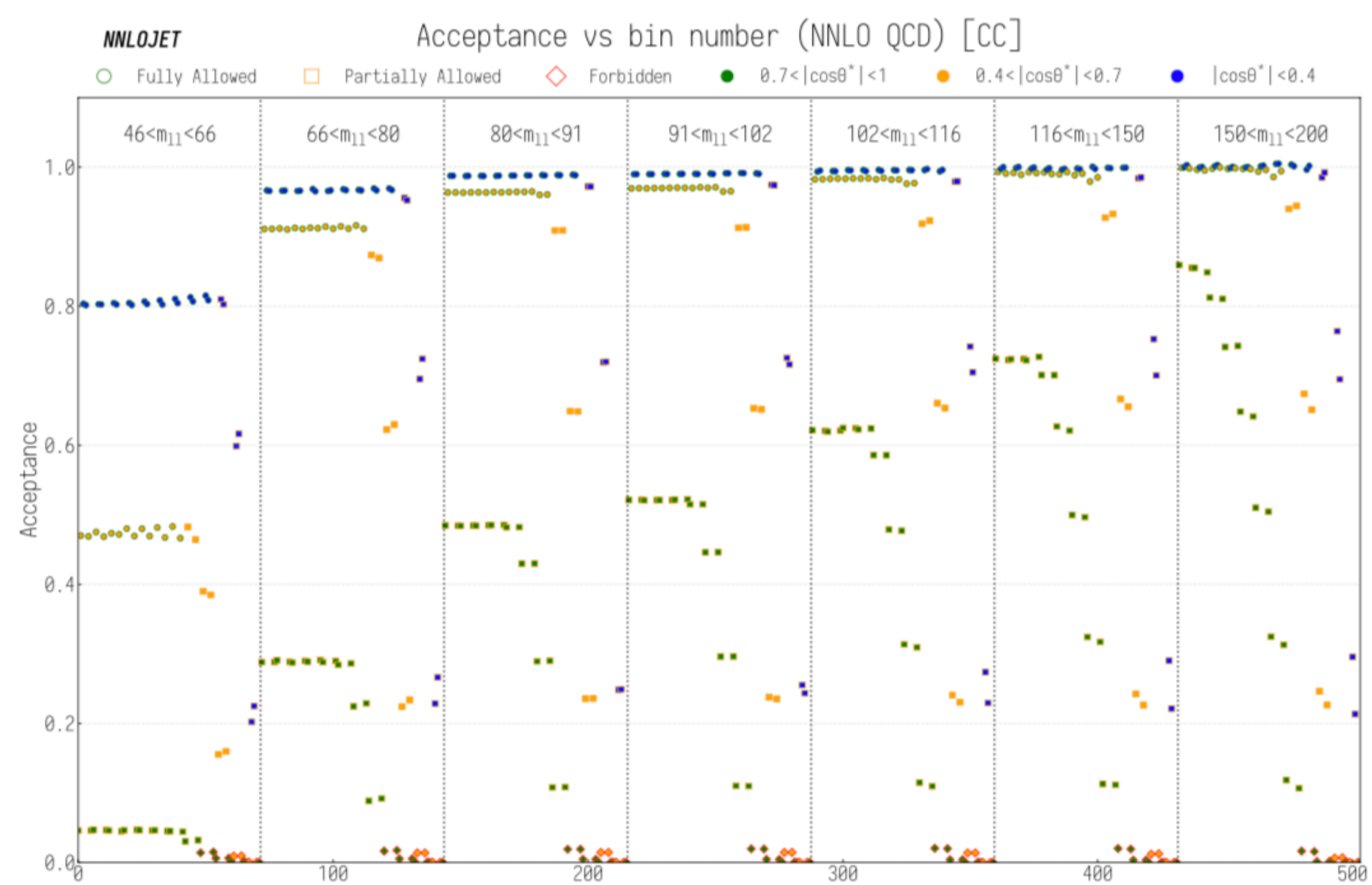


- Currently updating the xFitter study of Tevatron asymmetries to NNLO QCD and with more measurements

Some open issues

- At the current level of precision theory uncertainties are important and cannot be left out of a fit.
  - Restrict the fit to bins with high fiducial acceptance/small scale dependence? Apply q_T-resummation k-factors?
  - Include scales as NNPDF? Decorrelate the scale uncertainties until they are not constrained by the fit? Take an average of fits done with different scale choices/values?

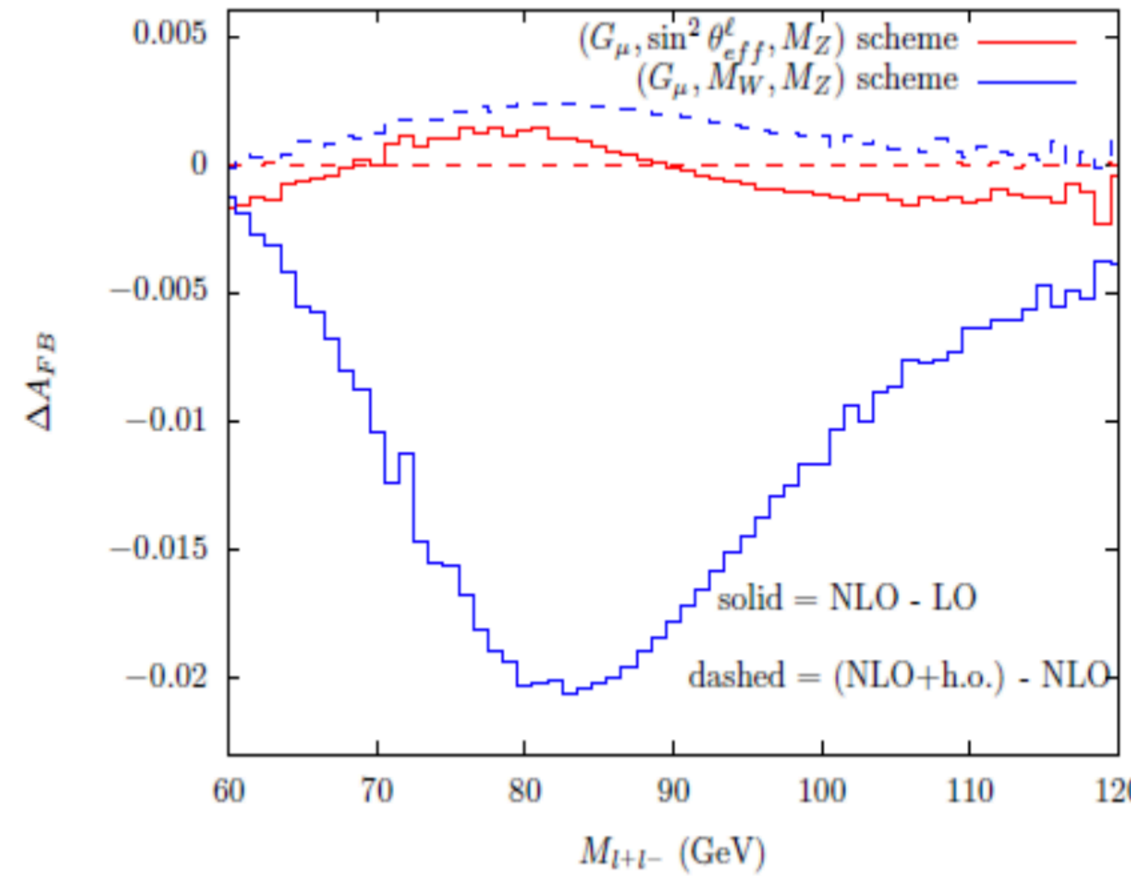




D. Walker, Ph.D. thesis

EW corrections

- At the level of precision we have reached NLO EW corrections are a must
  - Started to work with MCSANC to obtain NLO EW k-factor corrections on top of NLO QCD
  - At the same time NNLOJET is trying to include NLO EW



- Variations of EW parameters (m_W, $\sin^2\theta_{eff}$, g_V, g_A) can then be obtained through separate runs.
  - The dimensionality of this space can be large (~10D).
- How can we easily incorporate this in xFitter?

Professor+xFitter

- Common problem in Monte Carlo tuning, which attempts parameters optimisation for cases where evaluating the cost function at each iteration is expensive

- The Professor MC tuning code:
  - Computes an analytic interpolation of the change in each bin content due to a change in a parameter
  - Takes as input a random scan of the n-dimensional space
  - Interpolation obtained through SVD and saved in text files
  - The uncertainty on the prediction at each generated point is propagated to the interpolation coefficients
  - Alternative interpolation based on ML regression (SVM,NN) exists
- Ongoing work to write a reaction class to interface xFitter to the Professor interpolation files

An example

```
Summary:
DataDir: /Users/simoneamoroso/work/BfragmentationTunes/mc_A14_1p_rb
ProfVersion: 2.3.0beta
Date: 2018-06-13 23:15:19
DataFormat: binned 3
ParamNames: StringZ:rFactB
Dimension: 1
MinParamVals: 0.002286
MaxParamVals: 1.965571
DoParamScaling: 1
NumInputs: 60
Runs: 0000 0001 0002 0003 0004 0005 0006 0007 0008 0009 0010 0011 0012 0013 0014 0015 0016 0017 0018 0019 0020 0021 0
030 0031 0032 0033 0034 0035 0036 0037 0038 0039 0040 0041 0042 0043 0044 0045 0046 0047 0048 0049 0050 0051 0052 005
---
/ALEPH_1991_S2435284/d01-x01-y01#0 1.00000e+00 3.00000e+00
  val: 1 4 2.84593e-05 2.55076e-05 -6.54271e-05 8.56469e-05 -4.32299e-05  2.28564e-03 1.96557e+00
  err: 1 4 3.77197e-06 1.60237e-06 -4.00188e-06 5.05423e-06 -2.49733e-06  2.28564e-03 1.96557e+00
/ALEPH_1991_S2435284/d01-x01-y01#1 3.00000e+00 5.00000e+00
  val: 1 4 0.000491661 -9.54378e-05 0.000399962 -0.00070756 0.000396728  2.28564e-03 1.96557e+00
  err: 1 4 1.56779e-05 -1.52454e-06 6.39806e-06 -1.1346e-05 6.37009e-06  2.28564e-03 1.96557e+00
/ALEPH_1991_S2435284/d01-x01-y01#2 5.00000e+00 7.00000e+00
  val: 1 4 0.00363055 0.000180995 -0.00233696 0.00356659 -0.00159831  2.28564e-03 1.96557e+00
  err: 1 4 4.26037e-05 1.11832e-06 -1.40075e-05 2.13121e-05 -9.53633e-06  2.28564e-03 1.96557e+00
/ALEPH_1991_S2435284/d01-x01-y01#3 7.00000e+00 9.00000e+00
  val: 1 4 0.0156042 -0.00146718 -0.00218803 0.00279062 -0.000562467  2.28564e-03 1.96557e+00
  err: 1 4 8.83287e-05 -4.1255e-06 -6.45944e-06 7.90616e-06 -1.45605e-06  2.28564e-03 1.96557e+00
/ALEPH_1991_S2435284/d01-x01-y01#4 9.00000e+00 1.10000e+01
  val: 1 4 0.0428584 -0.00520551 -0.00170628 -0.00133849 0.00336442  2.28564e-03 1.96557e+00
  err: 1 4 0.000146387 -8.90091e-06 -3.00699e-06 -3.15481e-06 6.4654e-06  2.28564e-03 1.96557e+00
/ALEPH_1991_S2435284/d01-x01-y01#5 1.10000e+01 1.30000e+01
  val: 1 4 0.0823333 -0.00868661 -0.00508601 -0.00205666 0.00565062  2.28564e-03 1.96557e+00
  err: 1 4 0.0002029 -1.08337e-05 -5.76414e-06 -4.64623e-06 8.28669e-06  2.28564e-03 1.96557e+00
```

Dataset and bin number

Polynomial coefficients

Interpolation ranges

1 dimension
4 parameters

Parametrisation uncertainty

Possible Projects

- Update the xFitter NLO study of Tevatron asymmetries
  - Move to NNLO (+NLO EW?) and include QCD scales
  - Include more Tevatron measurements
- Perform a combination of the published unfolded AFB measurements, as a follow-up of the A_4 sensitivity paper
  - Impact on PDF fits and combination framework for $\sin^2\theta_{\mathrm{eff}}$
  - World best extraction of quark g_V, g_A couplings
- Perform a global fit to HERA+Tevatron+LHC DY measurements at NNLO QCD + NLO EW
  - Could be the fit of choice for precision EW measurements
- Any fit of PDF+MC parameters: non-perturbative corrections in jets, fragmentation function in W+charm,…
  - SM-EFT now available in MC generators (Powheg, aMC@NLO), interpolation would allow for PDF+EFT fits

Summary

* Already at the present level of precision, the prescriptions for dealing with PDF uncertainties and their in-situ constraints are insufficient
* A full PDF + EW fit would give significant benefits in understanding PDF uncertainties and reduce their impact
* This will come at the price of a significantly more complex analysis
* Can we start work in this direction with xFitter?
* In the long term this will be the only way to fully exploit the HL-LHC potential for precise EW measurements