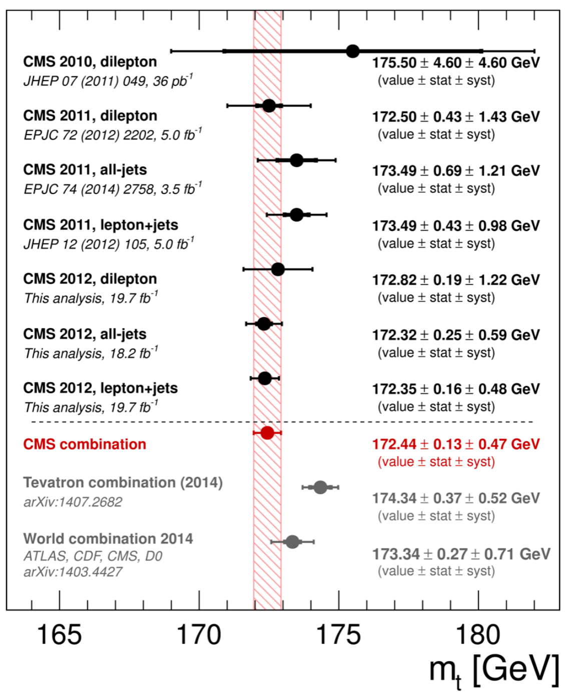
Precision MC for HL-LHC and beyond: non-perturbative issues

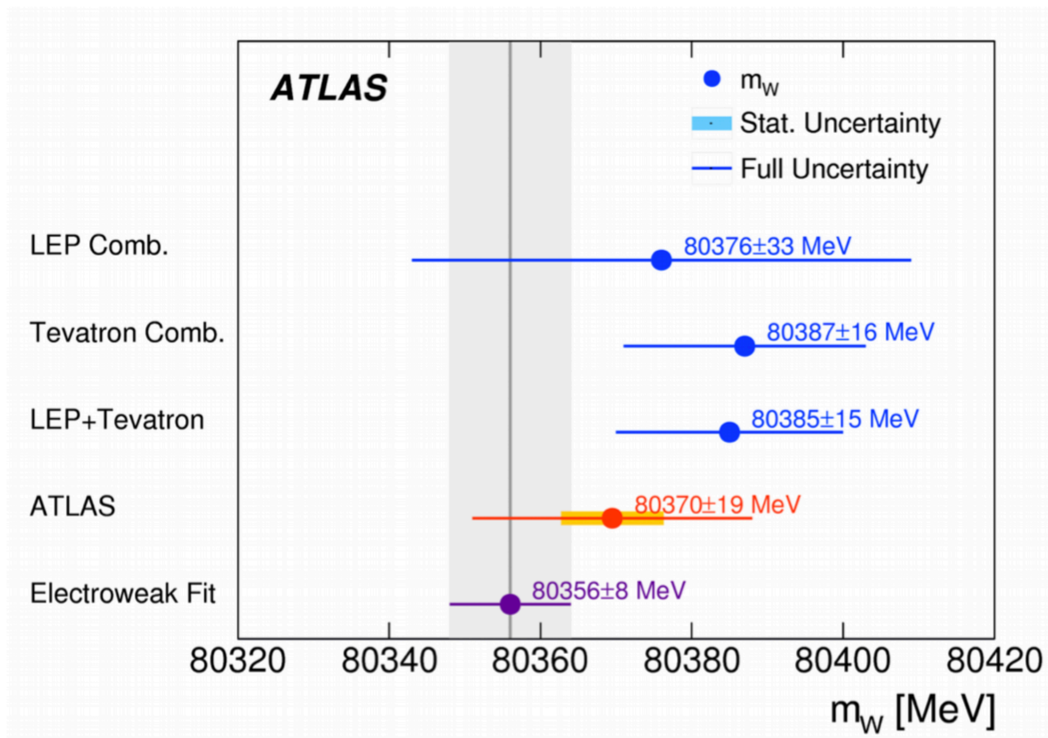
MCnet school

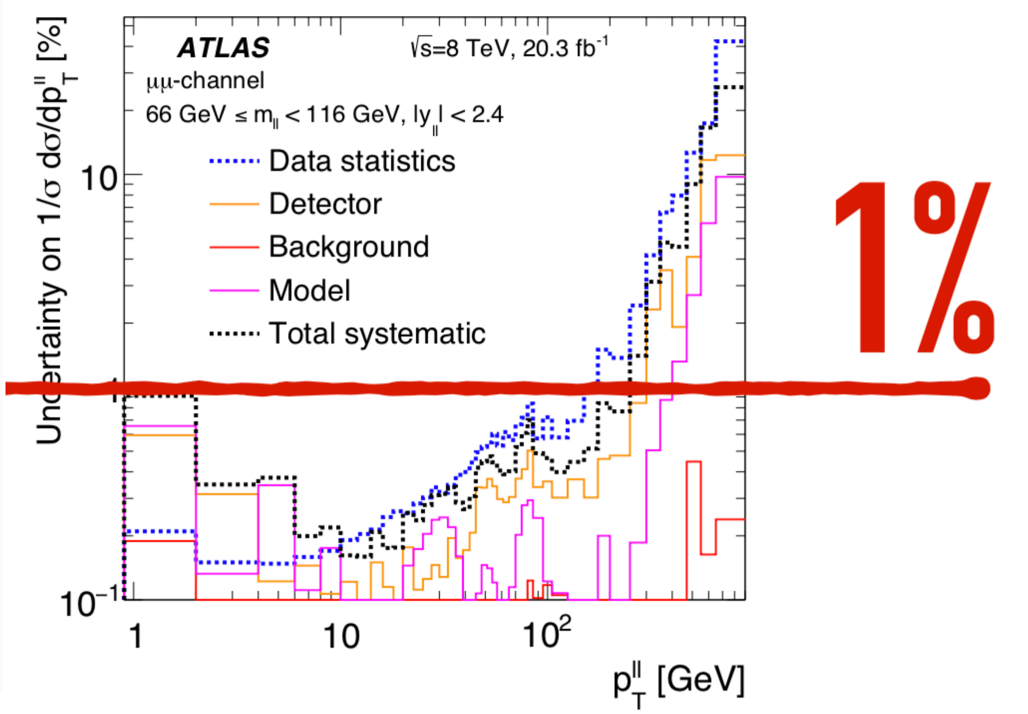
Aug. 02nd, 2019

Simone Amoroso

- The LHC measurements have now reached outstanding precision
  - Top mass now measured to below 500 MeV
  - First LHC W-mass measured by ATLAS with 19 MeV uncertainty
  - Z p_T measured to sub % precision





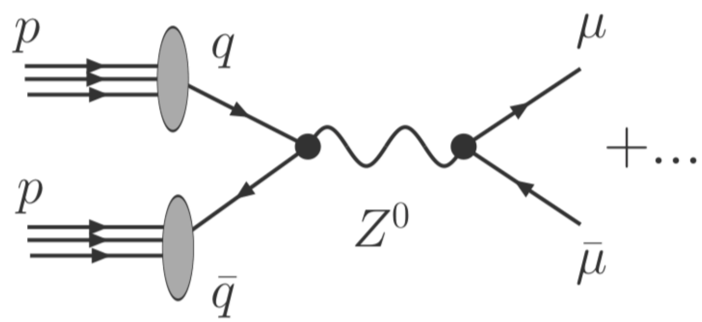


- At the HL-LHC a huge range of phase-space will reach sub-percent precision
- Is our theory up to it?
  - Need to demonstrate <1% control across a range of processes
  - For the perturbative part we know at least the route to achieve it
  - What about non perturbative physics ?

Outline

- Parton Distribution Functions
- How do we determine non-perturbative parameters in MC?
- Few important case studies
  - The W-boson mass and intrinsic-kT
  - Non-perturbative corrections to jets
  - Color reconnection and top mass
  - Bottom-quark fragmentation and top mass

Factorisation and PDFs



According to QCD factorization theorems, typical cross sections (e.g., for $p(k_1)p(k_2) \to \left[Z(q) \to \ell(k_3)\bar{\ell}(k_4)\right] X$) take the form

$$\sigma_{pp\to\ell\bar{\ell}X} = \sum_{a,b=q,\bar{q},g} \int_0^1 d\xi_1 \int_0^1 d\xi_2\, \widehat{\sigma}_{ab\to Z\to\ell\bar{\ell}} \left(\frac{x_1}{\xi_1}, \frac{x_2}{\xi_2}; \frac{Q}{\mu}\right) f_{a/p}(\xi_1,\mu) f_{b/p}(\xi_2,\mu)$$
$$+\, \mathcal{O}\left(\Lambda^2_{QCD}/Q^2\right)$$

$\widehat{\sigma}_{ab\to Z\to\ell\bar{\ell}}$ is the **hard-scattering cross section**

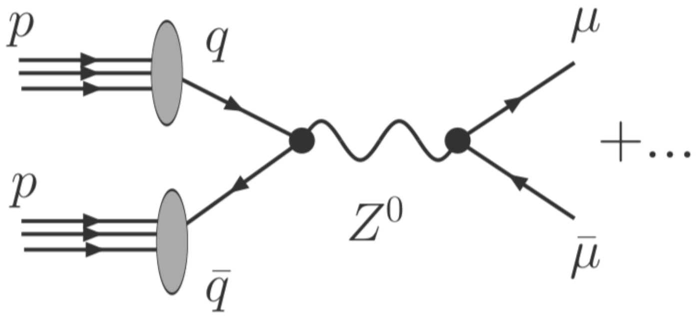
$f_{a/p}(\xi,\mu)$ are the **PDFs**

$Q^2 = (k_3+k_4)^2$, $x_{1,2} = (Q/\sqrt{s})\, e^{\pm y_V}$ — measurable quantities

ξ_1, ξ_2 are partonic momentum fractions (integrated over)

μ is a factorization scale (=renormalization scale from now on)

Factorisation and PDFs



According to QCD factorization theorems, typical cross sections (e.g., for $p(k_1)p(k_2) \to \left[Z(q) \to \ell(k_3)\bar{\ell}(k_4)\right] X$) take the form

$$\sigma_{pp\to\ell\bar{\ell}X} = \sum_{a,b=q,\bar{q},g} \int_0^1 d\xi_1 \int_0^1 d\xi_2\, \widehat{\sigma}_{ab\to Z\to\ell\bar{\ell}} \left(\frac{x_1}{\xi_1}, \frac{x_2}{\xi_2}; \frac{Q}{\mu}\right) f_{a/p}(\xi_1,\mu) f_{b/p}(\xi_2,\mu) + \mathcal{O}\left(\Lambda^2_{QCD}/Q^2\right)$$

μ is naturally set to be of order Q

Factorization holds up to terms of order Λ^2_{QCD}/Q^2

Subtract large collinear logarithms $\alpha_s^n \ln^k(Q^2/m_q^2)$ from $\widehat{\sigma}$

Resum them in $f_{a/p}(\xi,\mu)$ to all orders of α_s

PDFs evolution

The exact form of $f_{a/p}$ is not known; but its μ dependence is described by **D**okshitzer-**G**ribov-**L**ipatov-**A**ltarelli-**P**arisi (**DGLAP**) equations

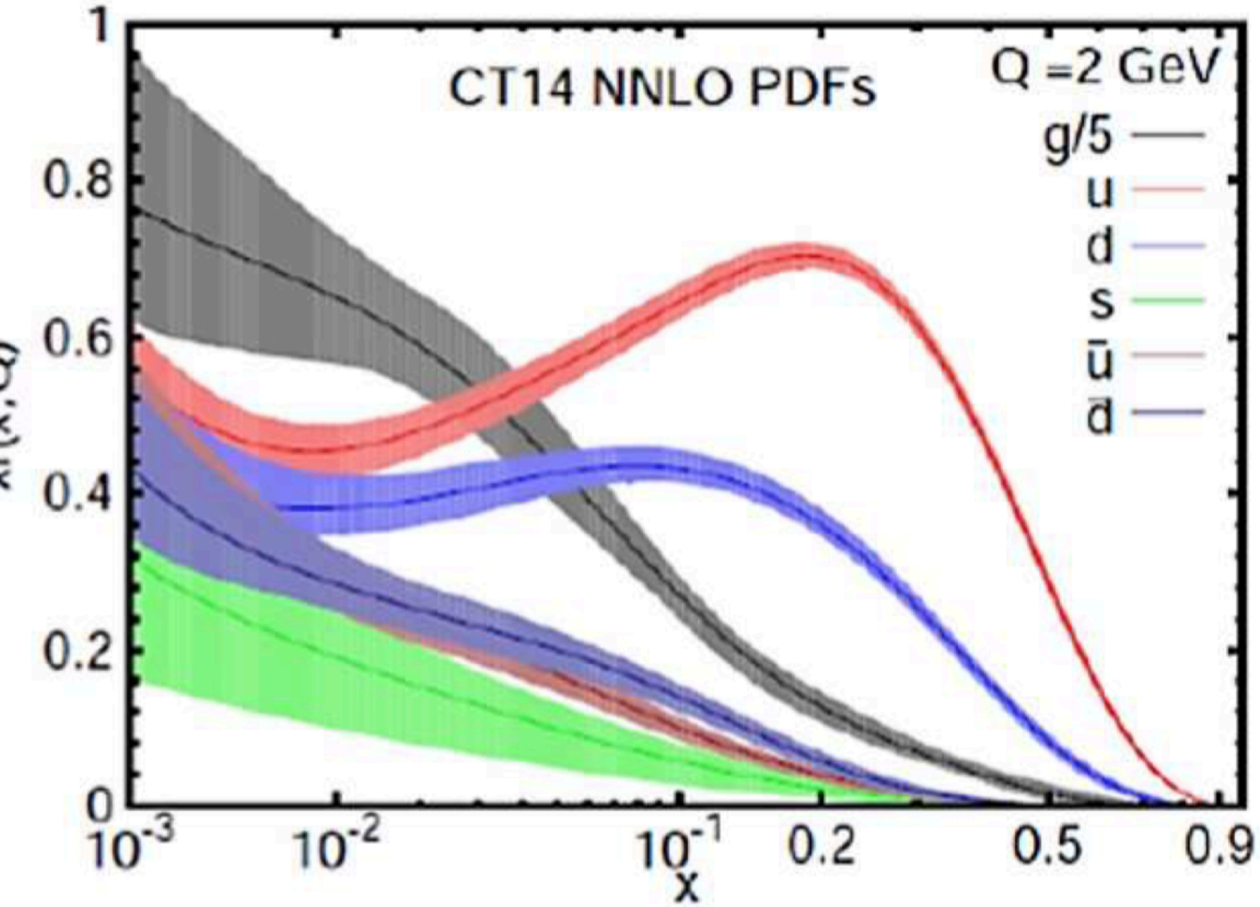
$$\mu \frac{df_{i/p}(x,\mu)}{d\mu} = \sum_{j=g,u,\bar{u},d,\bar{d},\ldots} \int_x^1 \frac{dy}{y} P_{i/j}\left(\frac{x}{y}, \alpha_s(\mu)\right) f_{j/p}(y,\mu)$$

$P_{i/j}$ are probabilities for $j \to ik$ collinear splittings;
are known to order α_s^3 (NNLO):

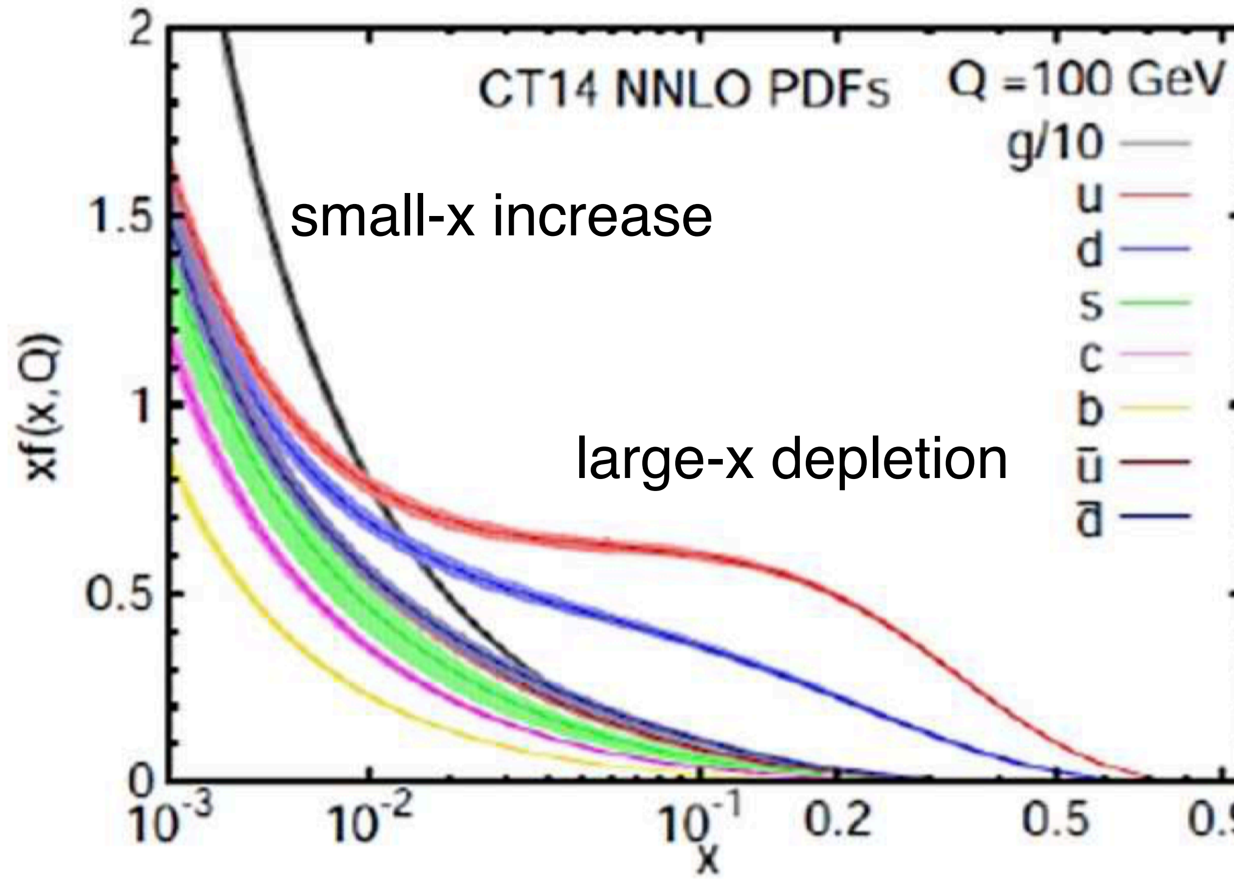
$$P_{i/j}\left(x, \alpha_s\right) = \alpha_s P_{i/j}^{(1)}(x) + \alpha_s^2 P_{i/j}^{(2)}(x)$$

PDFs are **universal** – depend only on the type of the hadron (p) and parton (q, $\bar{q}$, g)

Example of evolution



Q= 2 GeV

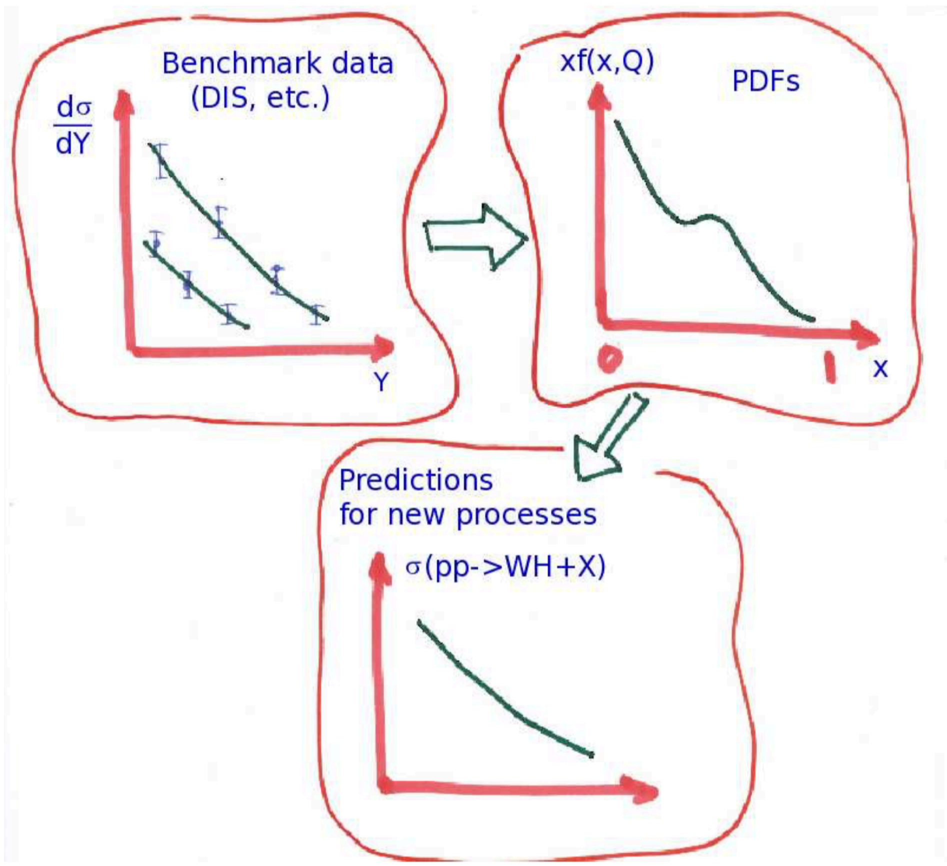


Q= 100 GeV

Evolution to higher scales produces more cartons at small momentum fraction (they loose energy by radiating)

Valence quarks slightly peaking at around ~1/3

How do we get PDFs?



- The DGLAP equations tell us how to evolve PDFs from a starting scale Q0 to any higher scale
- But we can't calculate them ab-initio (lattice?)
- We need to determine them by fitting to experimental data

- What do we need to obtain PDFs:
  - A value for Q_0, lower than any of the data used in the fit, ~ 1 GeV
  - A functional parametrisation

$$f_i(x, Q_0^2) = A_i x^{a_i}[1 + b_i\sqrt{x} + c_i x](1-x)^{d_i}$$

  - A scheme for the PDFs
  - Hard-scatter calculations at the same order of the fit considered
  - PDF evolution at the same order of the theory
  - Lots of data
  - A proper treatment of their experimental errors
  - A minimisation routine

PDFs at the starting scale

In practice, independent parametrizations $f_{a/p}(x, Q_0)$ are introduced for

$g, u, d, s, \bar{u}, \bar{d}, \bar{s}$ (always)
contribute $> 97\%$ of the proton's energy E_p at Q_0 ~ 1 GeV

- even in this case, the data are usually insufficient for constraining all PDF parameters; some of them can be fixed by hand
- e.g., $\bar{u} = \bar{d} = \bar{s}$ in outdated fits

c and or b (occasionally; in a model allowing nonperturbative "intrinsic heavy-quark production")

photons γ (in QCD+QED PDFs by CT, xFitter, LUX, MRST, NNPDF... groups)

Which data are used?

S. Dulat et al., arXiv:1506.07443

ID#	Experimental data set		$N_{pt,n}$	χ^2_n
101	BCDMS F_2^p	[24]	337	384
102	BCDMS F_2^d	[25]	250	294
104	NMC F_2^d/F_2^p	[26]	123	133
106	NMC σ^p_{red}	[26]	201	372
108	CDHSW F_2^p	[27]	85	72
109	CDHSW F_3^p	[27]	96	80
110	CCFR F_2^p	[28]	69	70
111	CCFR xF_3^p	[29]	86	31
124	NuTeV $\nu\mu\mu$ SIDIS	[30]	38	24
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS	[30]	33	39
126	CCFR $\nu\mu\mu$ SIDIS	[31]	40	29
127	CCFR $\bar{\nu}\mu\mu$ SIDIS	[31]	38	20
145	H1 σ_r^b	[32]	10	6.8
147	Combined HERA charm production	[33]	47	59
159	HERA1 Combined NC and CC DIS	[34]	579	591
169	H1 F_L	[35]	9	17

ID#	Experimental data set		$N_{pt,n}$	χ^2_n
201	E605 Drell-Yan process	[37]	119	116
203	E866 Drell-Yan process, $\sigma_{pd}/(2\sigma_{pp})$	[38]	15	13
204	E866 Drell-Yan process, $Q^3 d^2\sigma_{pp}/(dQdx_F)$	[39]	184	252
225	CDF Run-1 electron A_{ch}, $p_{T\ell} > 25$ GeV	[40]	11	8.9
227	CDF Run-2 electron A_{ch}, $p_{T\ell} > 25$ GeV	[41]	11	14
234	DØ Run-2 muon A_{ch}, $p_{T\ell} > 20$ GeV	[42]	9	8.3
240	LHCb 7 TeV 35 pb^{-1} W/Z $d\sigma/dy_\ell$	[43]	14	9.9
241	LHCb 7 TeV 35 pb^{-1} A_{ch}, $p_{T\ell} > 20$ GeV	[43]	5	5.3
260	DØ Run-2 Z rapidity	[44]	28	17
261	CDF Run-2 Z rapidity	[45]	29	48
266	CMS 7 TeV 4.7 fb^{-1}, muon A_{ch}, $p_{T\ell} > 35$ GeV	[46]	11	12.1
267	CMS 7 TeV 840 pb^{-1}, electron A_{ch}, $p_{T\ell} > 35$ GeV	[47]	11	10.1
268	ATLAS 7 TeV 35 pb^{-1} W/Z cross sec., A_{ch}	[48]	41	51
281	DØ Run-2 9.7 fb^{-1} electron A_{ch}, $p_{T\ell} > 25$ GeV	[14]	13	35
504	CDF Run-2 inclusive jet production	[49]	72	105
514	DØ Run-2 inclusive jet production	[50]	110	120
535	ATLAS 7 TeV 35 pb^{-1} incl. jet production	[51]	90	50
538	CMS 7 TeV 5 fb^{-1} incl. jet production	[52]	133	177

Modern fits involve up to 40 experiments,
>3000 data points and 100+ free parameters

- Typically estimated from the Hessian method

Hessian matrix

After diagonalization

$$\chi^2(a) = \chi_0^2 + \frac{1}{2}\frac{\partial\chi^2}{\partial a_i \partial a_j}(a-a_0)_i(a-a_0)_j + \cdots \to \chi_0^2 + \sum_i z_i^2$$

$$y_i = a_i - a_i^0$$

$$\chi^2 = \chi_0^2 + \sum_{i,j} H_{ij}\, y_i\, y_j\,, \qquad H_{ij} = \frac{1}{2}\left(\frac{\partial^2\chi^2}{\partial y_i\,\partial y_j}\right)_0\,,$$

$$\sum_j H_{ij}\, v_{jk} = \epsilon_k\, v_{ik}$$ Eigenvalue equation

$$\sum_i v_{ij}\, v_{ik} = \delta_{jk}\,.$$ Orthonormality

$$z_i = \sqrt{\epsilon_i}\sum_j y_j\, v_{ji}$$ Change of basis in terms of the eigenvalues

$$\Delta\chi^2 = \chi^2 - \chi_0^2 = \sum_i z_i^2$$

the surfaces of constant χ2 are spheres in z_i space, with Δχ2 the squared distance from the minimum.

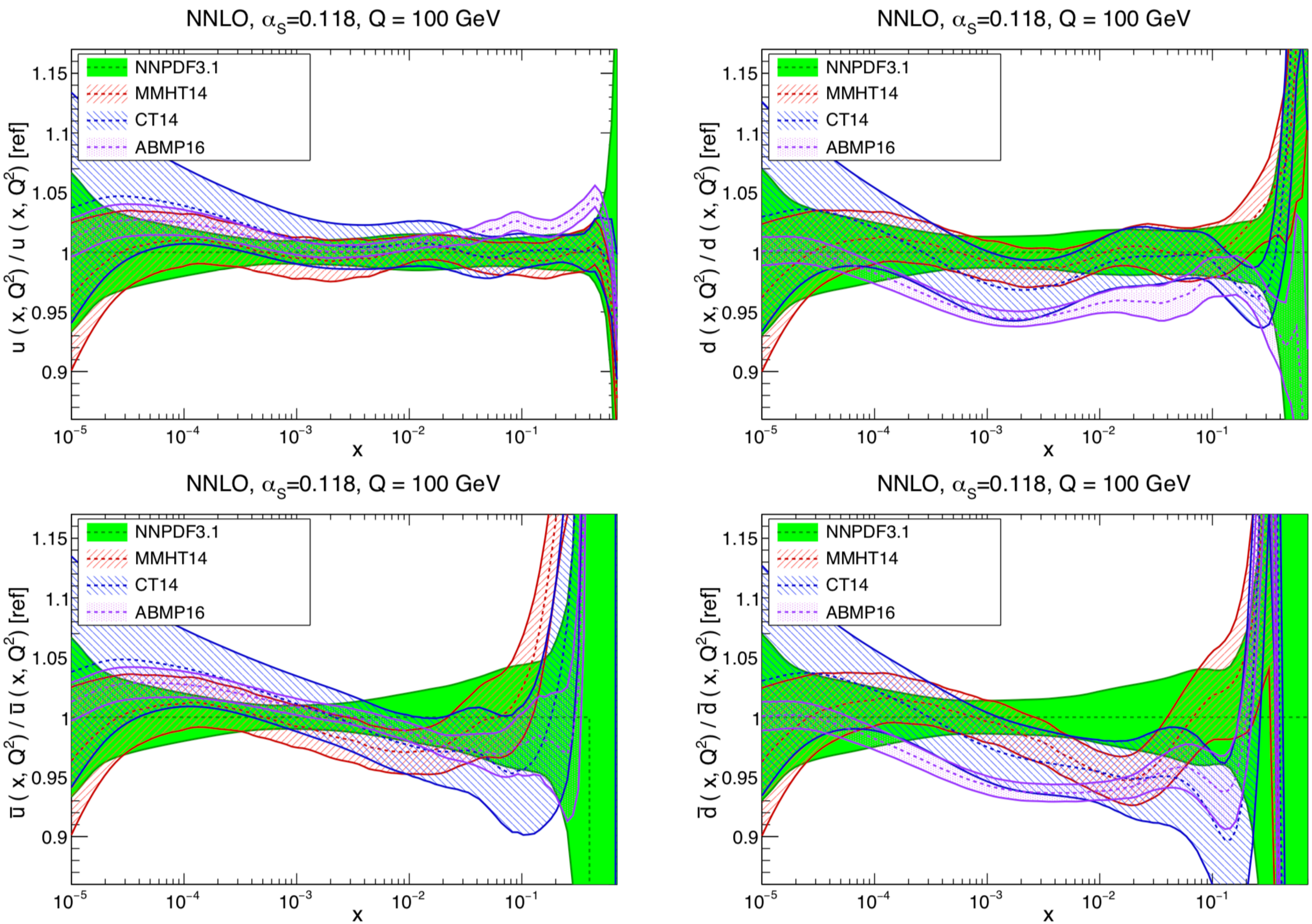
- Alternatively, Monte Carlo replicas can be used, as in NNPDF

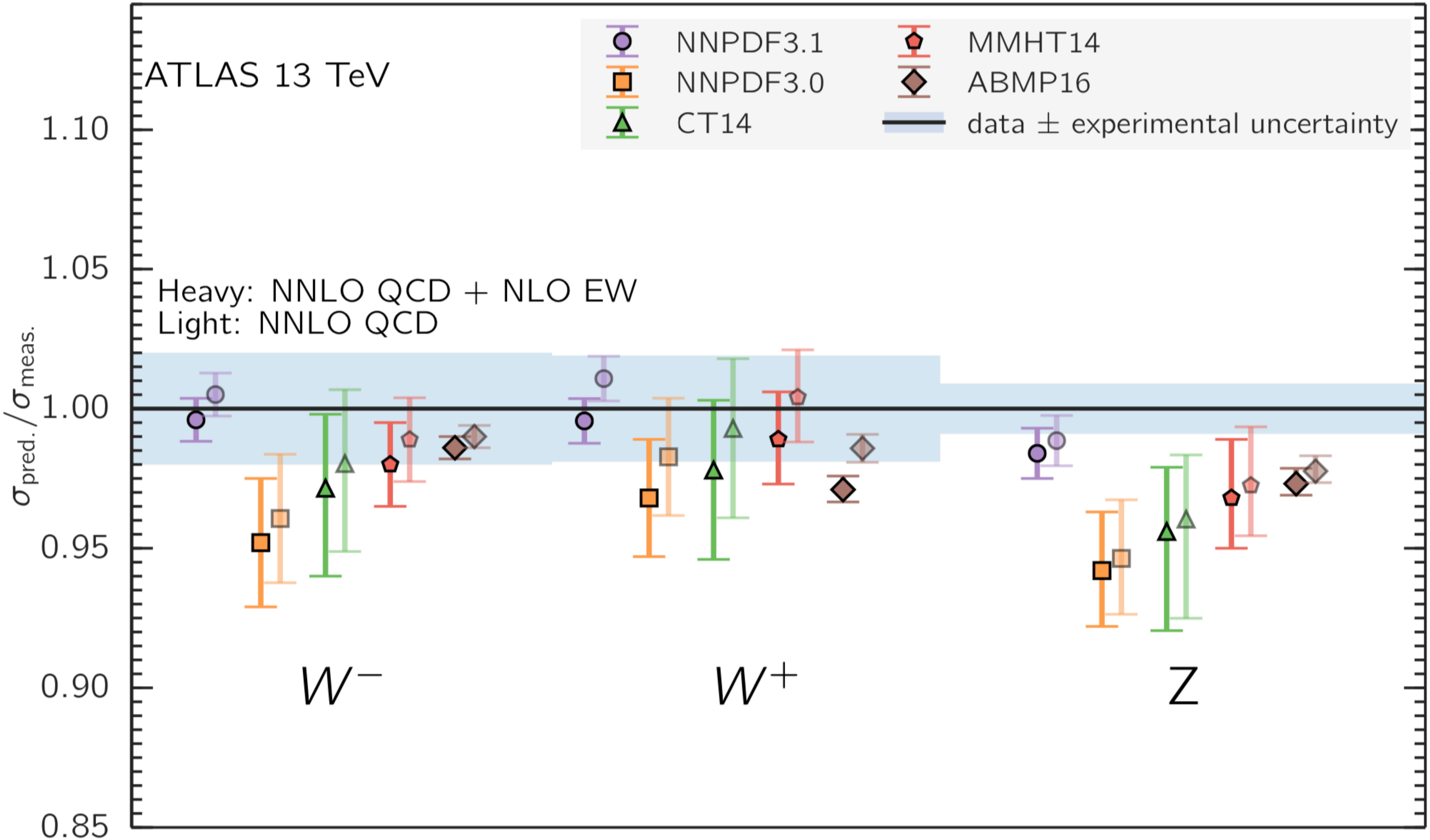
PDF FITS

- Some PDFs are specialised, make use of restricted datasets
  - **HERAPDF** use only HERA DIS data
  - **ABMP**: use Fixed target DIS, HERA, Drell-Yan and top data
  - **ATLASep:** HERA DIS data plus ATLAS DY measurements

- Other are general purpose global fit, including most of the available data
  - **CTEQ-TEA/MMHT/NNPDF**: HERA DIS, Fixed target DIS and Drell-Yan, Vector Boson, top and Inclusive jet production at colliders
  - **PDF4LHC** combinations, provide a "weighted average” of the three global sets and their uncertainties

And a comparison





NNPDF3.1, 1706.00428 (data from 1603.09222)

Where can you get PDFs from

LHAPDF 6.2.3

Main Page | Related Pages | Namespaces ▾ | Classes ▾ | Files ▾ | Examples

Search

LHAPDF Documentation

Introduction

LHAPDF is a general purpose C++ interpolator, used for evaluating PDFs from discretised data files. Previous versions of **LHAPDF** were written in Fortran 77/90 and are documented at http://lhapdf.hepforge.org/lhapdf5/.

LHAPDF6 vastly reduces the memory overhead of the Fortran **LHAPDF** (from gigabytes to megabytes!), entirely removes restrictions on numbers of concurrent PDFs, allows access to single PDF members without needing to load whole sets, and separates a new standardised PDF data format from the code library so that new PDF sets may be created and released easier and faster. The C++ LHAPDF6 also permits arbitrary parton contents via the standard PDG ID code scheme, is computationally more efficient (particularly if only one or two flavours are required at each phase space point, as in PDF reweighting), and uses a flexible metadata system which fixes many fundamental metadata and concurrency bugs in LHAPDF5.

Compatibility routines are provided as standard for existing C++ and Fortran codes using the LHAPDF5 and PDFLIB legacy interfaces, so you can keep using your existing codes. But the new interface is much more powerful and pleasant to work with, so we think you'll want to switch once you've used it!

LHAPDF6 is documented in more detail in http://arxiv.org/abs/1412.7420

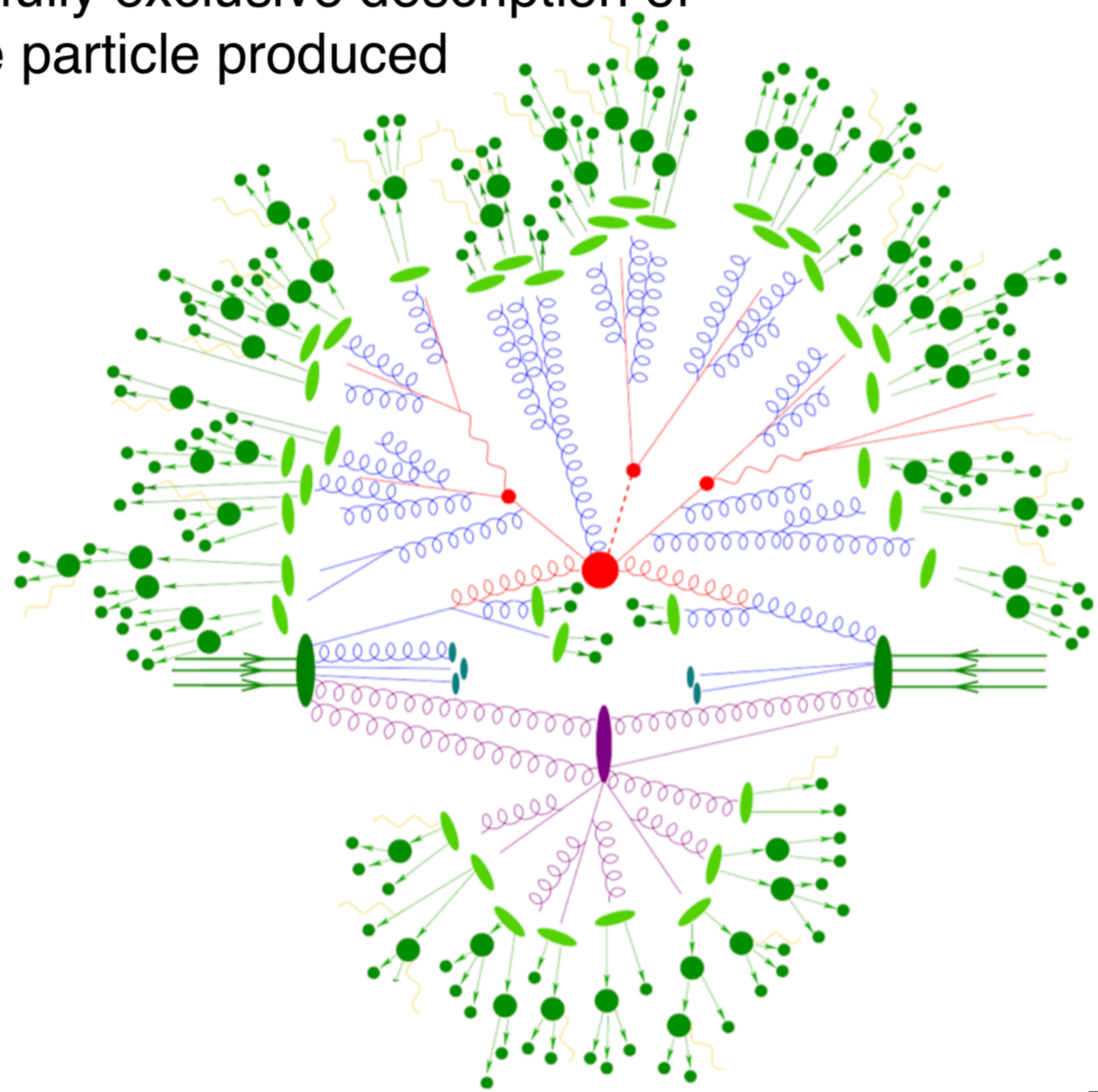
Table of Contents

https://lhapdf.hepforge.org

MC event generators

- Monte Carlo event generators provide a fully exclusive description of a collision event in terms of the final state particle produced

  - Matrix Element (ME) generators simulate the central part of the event
  - Parton Showers (PS) produce additional soft and collinear QCD radiation
  - Multiple Interaction (MPI) models produce secondary hard interactions
  - Fragmentation models the transition from QCD partons to the visible hadrons
  - Hadrons can further decay into other detector stable particles
  - Photon Emission generators simulate additional QED radiation



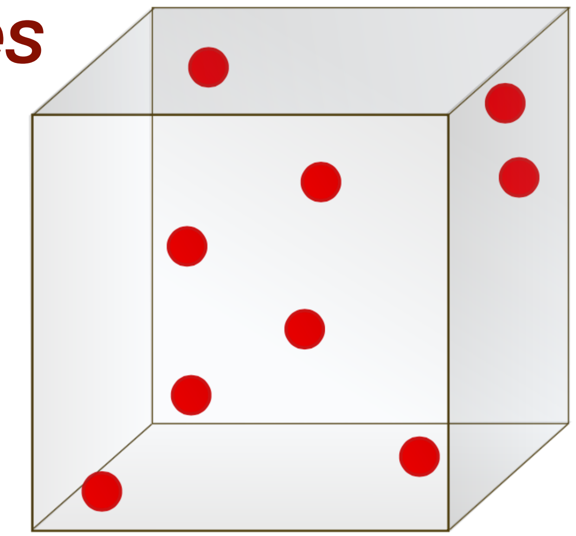
Non-perturbative effects

- With to the tremendous theory progress in the last decade we have available parton-level predictions up to N^2LO, complex matching/merging to higher multiplicity ME and showers
- Unfortunately we cannot compare them to our data, as they are expressed in terms of unphysical "partons"

- In MC generator the description of non-perturbative physics below the shower cut-off is done by phenomenological models
- These models carry parameters that cannot be determined from first principles and need to be constrained, "*tuned*", to data
- The procedure, and the complications that arise, are similar to the case of PDF determinations
  - For a long-time tuning done "by-eye"
  - More recently automated approaches, e.g. ***professor***

Tuning in professor

- Choose relevant generator ***parameters and ranges***
- Random sample the *D*-dimensional hypercube
- ***Generate samples*** for n anchor points (using **RIVET**)
- ***Interpolate*** the generator response to parameter changes with an analytic approximation (polynomials)
- Construct the goodness of fit measure



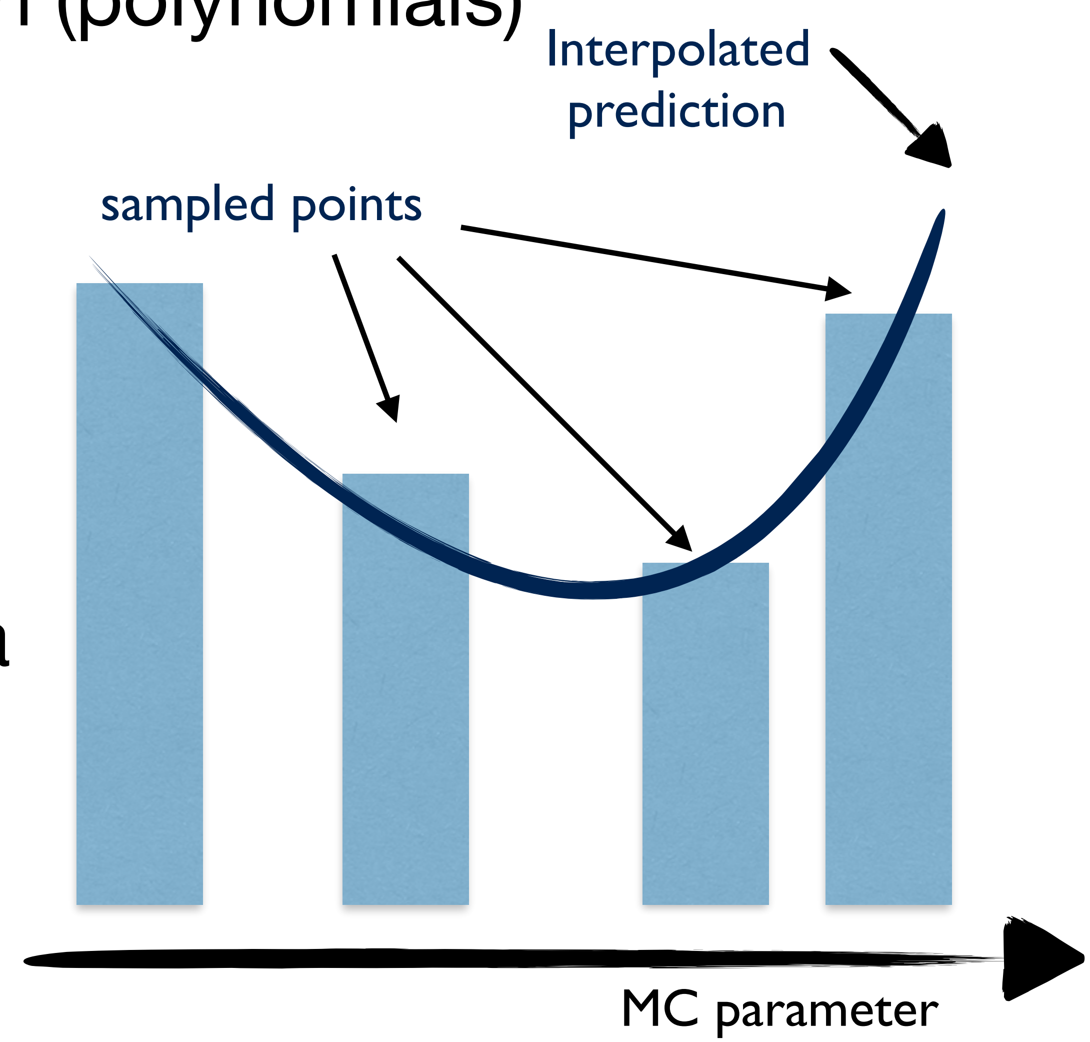
$$\chi^2(m) = \sum_{i,k}(m_i - \mu_i)C^{-1}_{ik}(m_k - \mu_k),$$

prediction

data

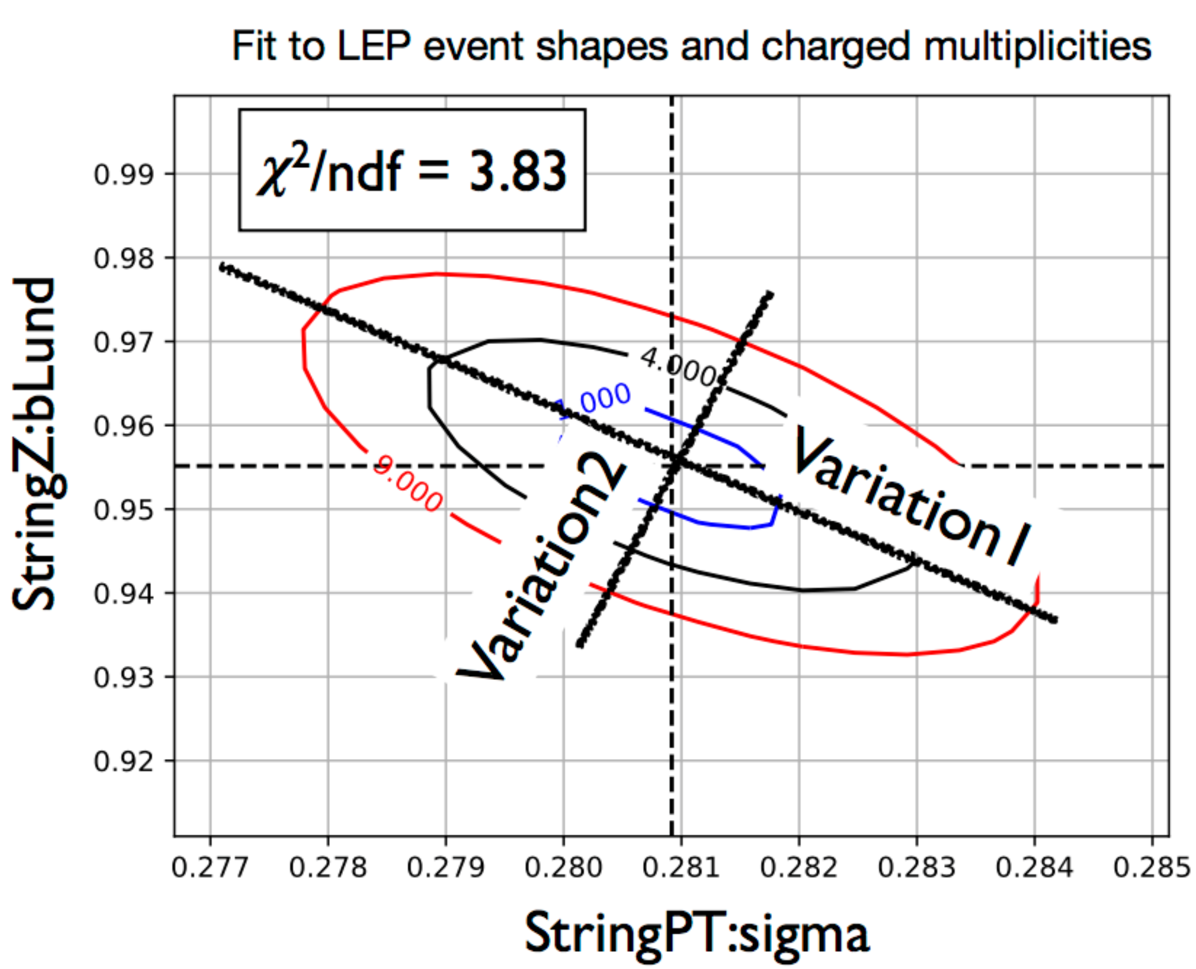
cov. matrix

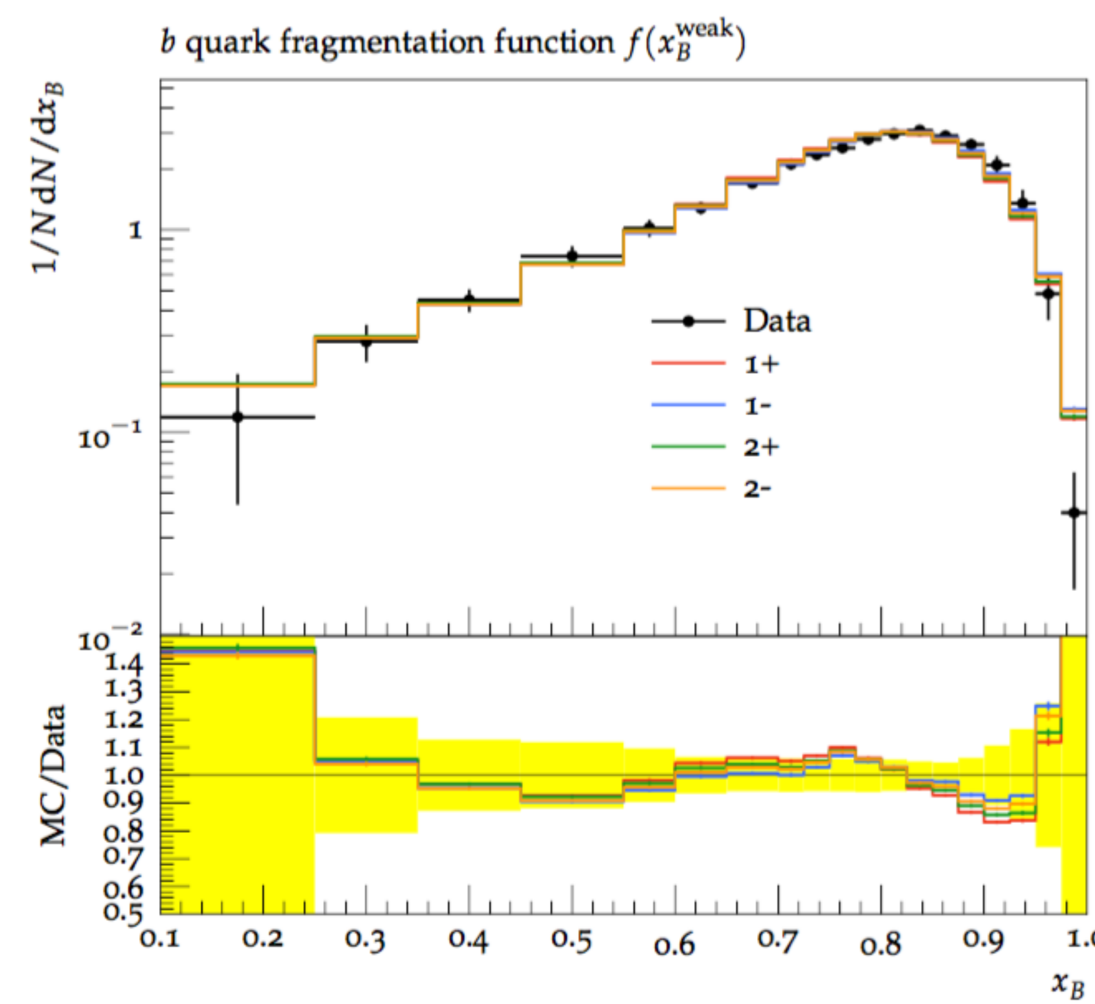
- ***Minimize*** the interpolation against data



Parameter uncertainties

- In addition to the best fit point, we need to derive *sensible uncertainties* on the fitted parameters
- We can obtain with the Hessian method from $\Delta\chi^2$ contours in the directions of maximum variations around the fit minimum
  - But most of the time the MC will NOT be able to describe all of the data properly; and an ad-hoc tolerance will be needed



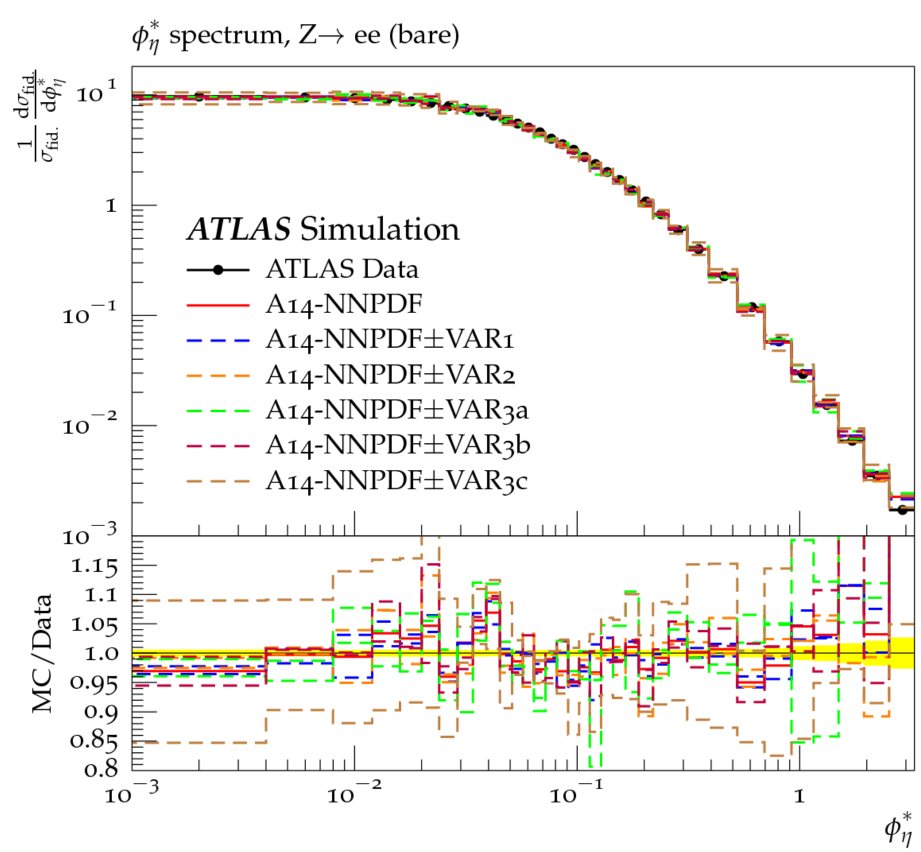


What do we tune?

- ***Intrinsic K_T:*** important to describe the first ~5 GeV of boson p_T
- ***FSR shower:*** α_S, IR cut-off, starting scale fudge factors; different shower evolutions (p_T,Q^2, …) -> different tunes. Tuned to event shape measurements in e^+e^- assuming universality
- ***ISR shower:*** similar to FSR, but tuned to hadron collider data (dijet angular correlations and Z p_T)
- ***Hadronisation:*** String or cluster parameters, separate for heavy-quark fragmentation. Tuned to identified particle spectra and multiplicities in e^+e^-
- ***Underlying event:*** matter distribution in the proton, cut-off for multiple parton interactions (MPI), color reconnection. Tuned to hadron collider data, sensitive to the PDF choice.
- ***Matching/merging parameters:*** matched and merged calculations carry parameters that can significantly affect the predictions (h_{damp}, frac_upp/low, …). Sometimes these are optimised using data.

An ATLAS example

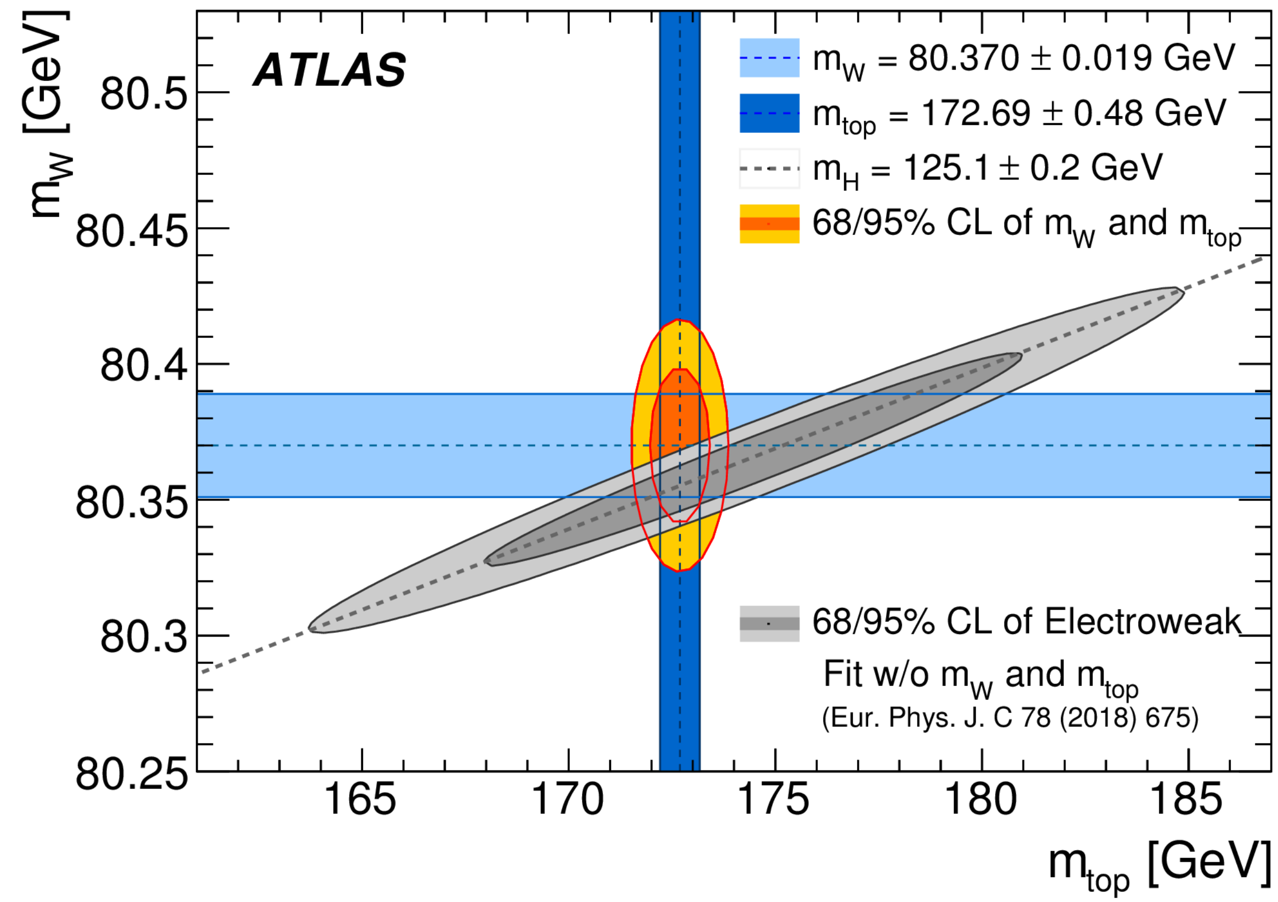
- A14, an ATLAS professor tune of Pythia8 PS and MPI
  - A total of 10 parameters are considered
  - Using ATLAS 7 TeV measurements of jets, top and DY production
  - 10 different measurements, ~80000 different bins
- Uncertainties are obtained using Eigentunes
  - Ad-hoc determination of their size, reduction in their number



Param	+ variation	− variation
VAR1: MPI+CR (UE activity and incl jet shapes)		
`BeamRemnants:reconnectRange`	1.73	1.69
`MultipartonInteractions:alphaSvalue`	0.131	0.121
VAR2: ISR/FSR (jet shapes and substructure)		
`SpaceShower:pT0Ref`	1.60	1.50
`SpaceShower:pTdampFudge`	1.04	1.08
`TimeShower:alphaSvalue`	0.139	0.111
VAR3a: ISR/FSR ($t\bar{t}$ gap)		
`MultipartonInteractions:alphaSvalue`	0.125	0.127
`SpaceShower:pT0Ref`	1.67	1.51
`SpaceShower:pTdampFudge`	1.36	0.93
`SpaceShower:pTmaxFudge`	0.98	0.88
`TimeShower:alphaSvalue`	0.136	0.124
VAR3b: ISR/FSR (jet 3/2 ratio)		
`SpaceShower:alphaSvalue`	0.129	0.126
`SpaceShower:pTdampFudge`	1.04	1.07
`SpaceShower:pTmaxFudge`	1.00	0.83
`TimeShower:alphaSvalue`	0.114	0.138
VAR3c: ISR ($t\bar{t}$ gap, dijet decorrelation and Z-boson p_T)		
`SpaceShower:alphaSvalue`	0.140	0.115

NP effects and the W-mass

- W-mass is a crucial measurement to improve the sensitivity of the global EW fit
- ATLAS recently released the first LHC measurement of the W-boson mass at 7 TeV
- Equalling in precision the best measurement from CDF

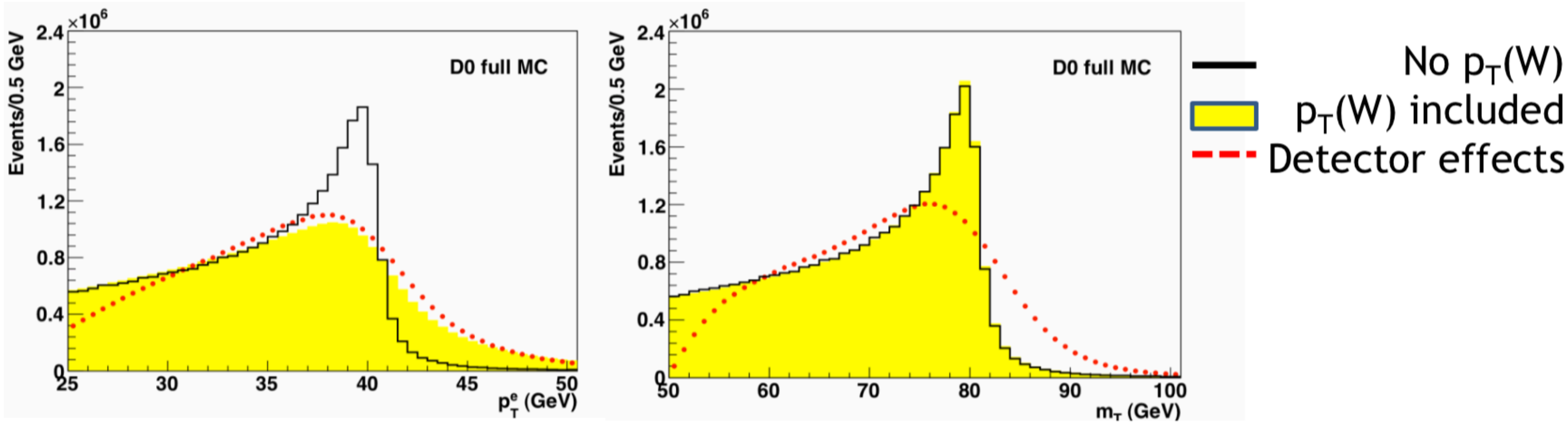


M_W = 80369.5 ± 6.8 (stat) ± 10.6 (exp.syst.) ± 13.6 (model.syst.) MeV

The dominant uncertainty comes from physics modelling

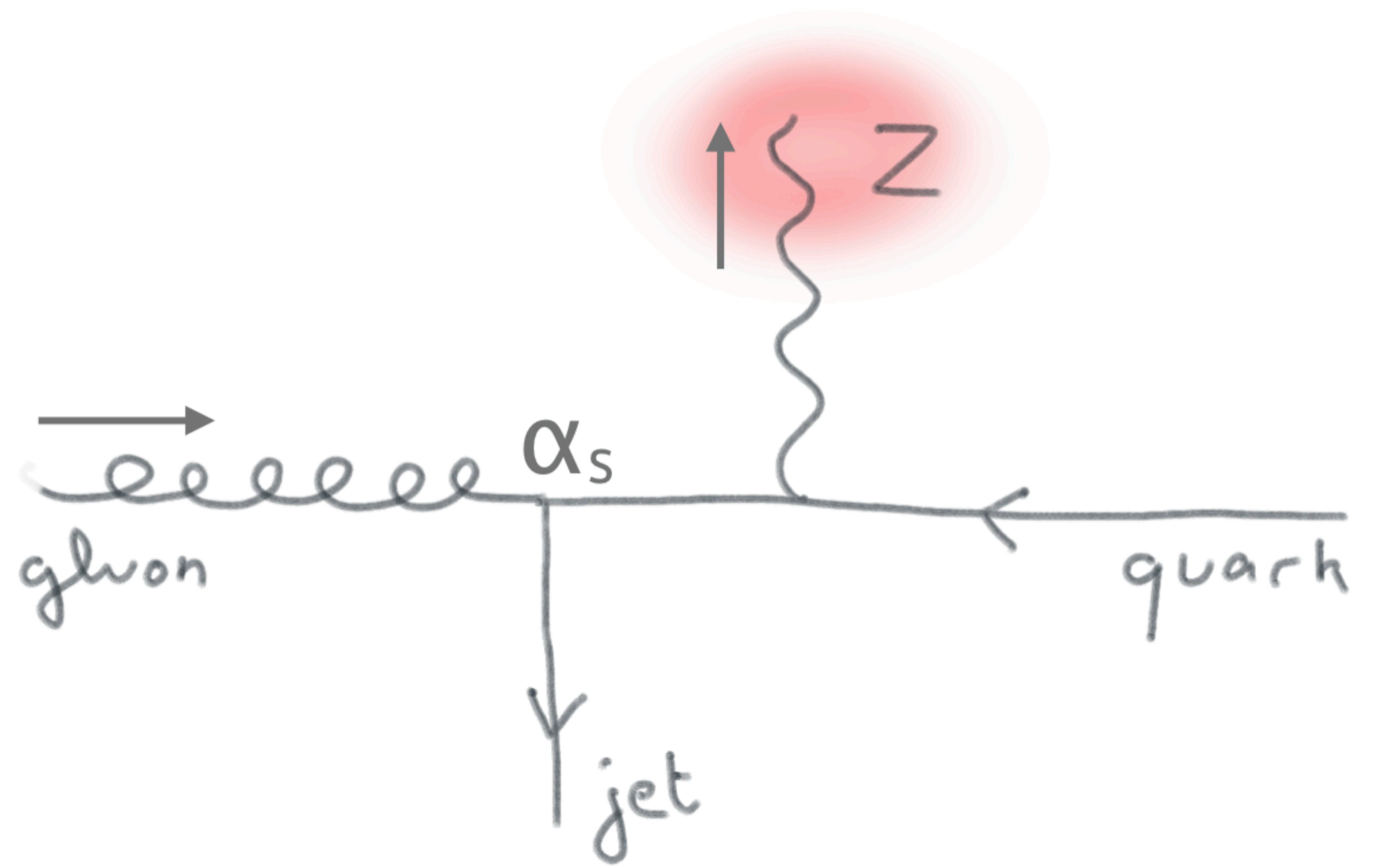
Combined categories	Value [MeV]	Stat. Unc.	Muon Unc.	Elec. Unc.	Recoil Unc.	Bckg. Unc.	QCD Unc.	EW Unc.	PDF Unc.	Total Unc.	χ^2/dof of Comb.
m_T-p_T^ℓ, $W^\pm$, e-μ	80369.5	6.8	6.6	6.4	2.9	4.5	8.3	5.5	9.2	18.5	29/27

Extracting the W-mass



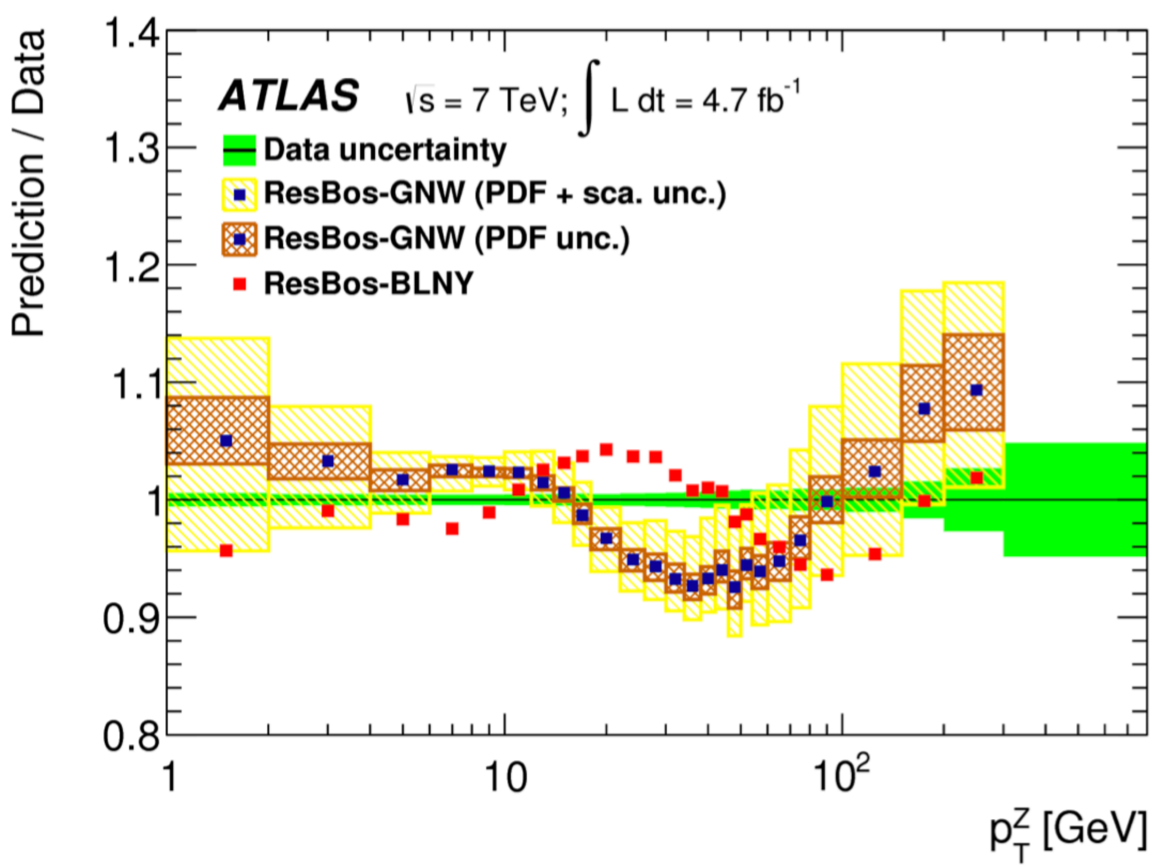
- The W-mass is extracted from template fits to two variables
- The transverse mass $m_T(W)$
  - Theoretically stable against radiative corrections
  - Challenging due to the incomplete event reconstruction from the escaping neutrino
- The lepton transverse momentum, $p_T(l)$
  - Experimentally clean
  - Extremely sensitive to the modelling of $p_T(W)$
- For a W-mass uncertainty of 10-20 MeV need control on the shapes of these distributions to below percent level !

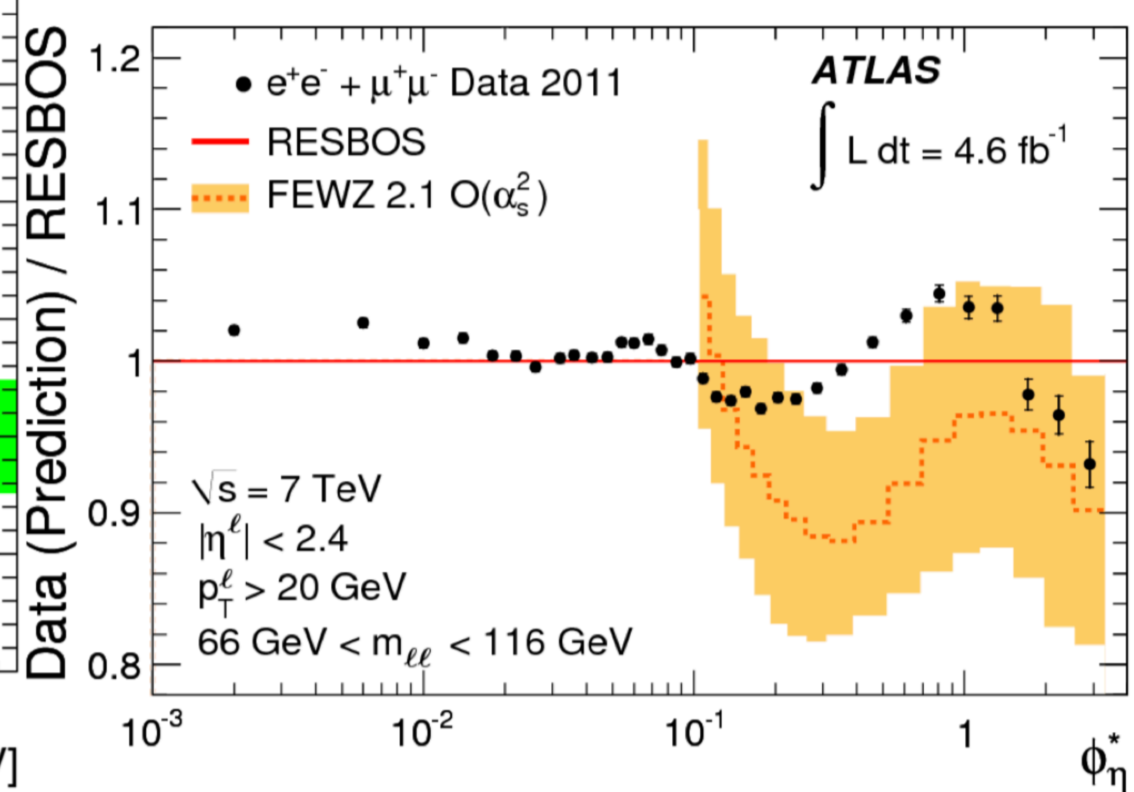
- Large p_T (> 20 GeV), where perturbation theory is viable
  - State of the art is now NNLO
- Small p_T (< 20 GeV), which makes 90% of the cross section
  - Large logs of M/q_T, can be resumed analytically or with a parton shower
  - Non-perturbative effects become important: heavy-quark masses, intrinsic-kT, ISR shower cut-off



- Intrinsic-k_T is used to parametrise non-perturbative transverse momentum of the partons inside the proton
  - From Fermi motion would expect an average of order 0.3 - 0.5 GeV
  - But fits to data give larger values, of about 1 - 2 GeV

- A direct measurement of $p_T(W)$ is experimentally challenging
- But very precise measurements of Z pT and ϕ^* have been performed
  - Typical accuracy for pT < 30 GeV: ~ 0.5 %
- The measurements are used to validate the theory predictions
- Then propagate $p_T(Z)$ to $p_T(W)$ using a Monte Carlo model
  - Non-perturbative form-factors, or intrinsic-kT, largely correlated between W and Z. Differences of ~3% expected (see1309.1393).



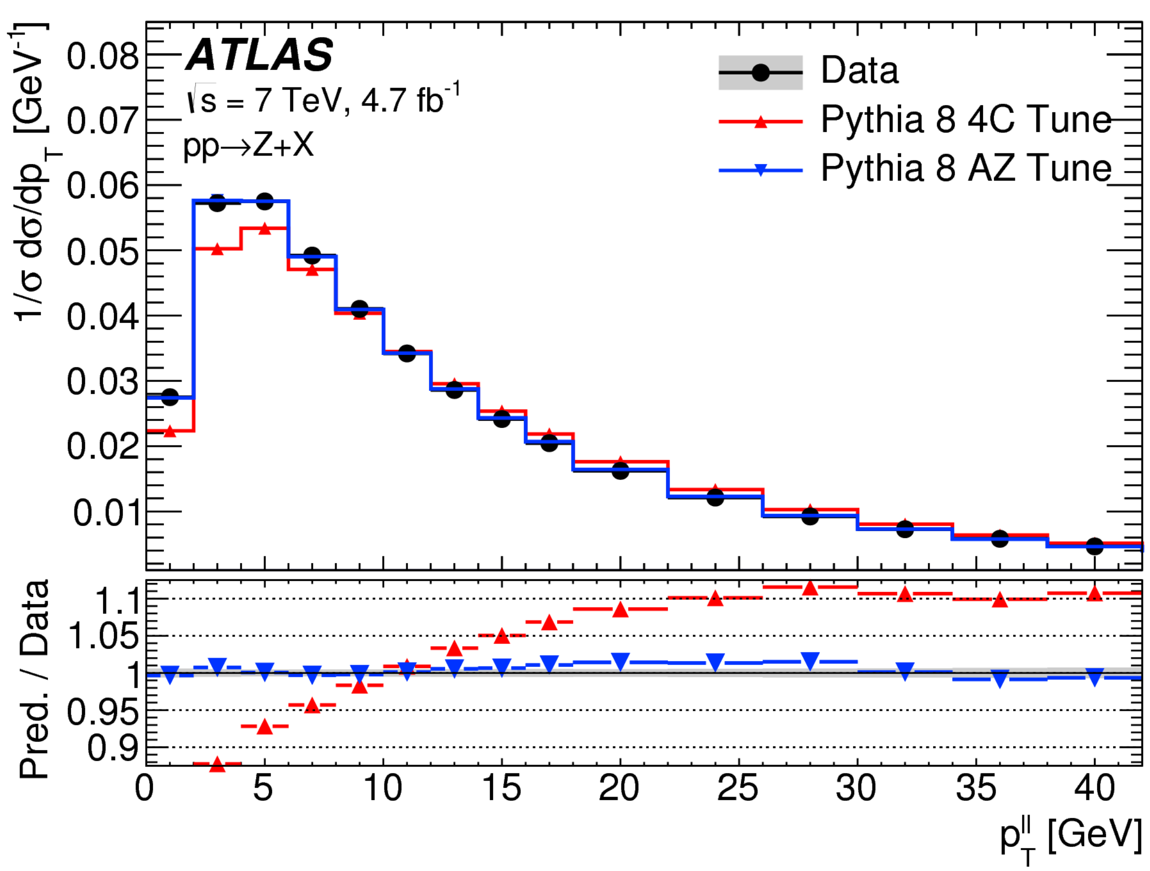


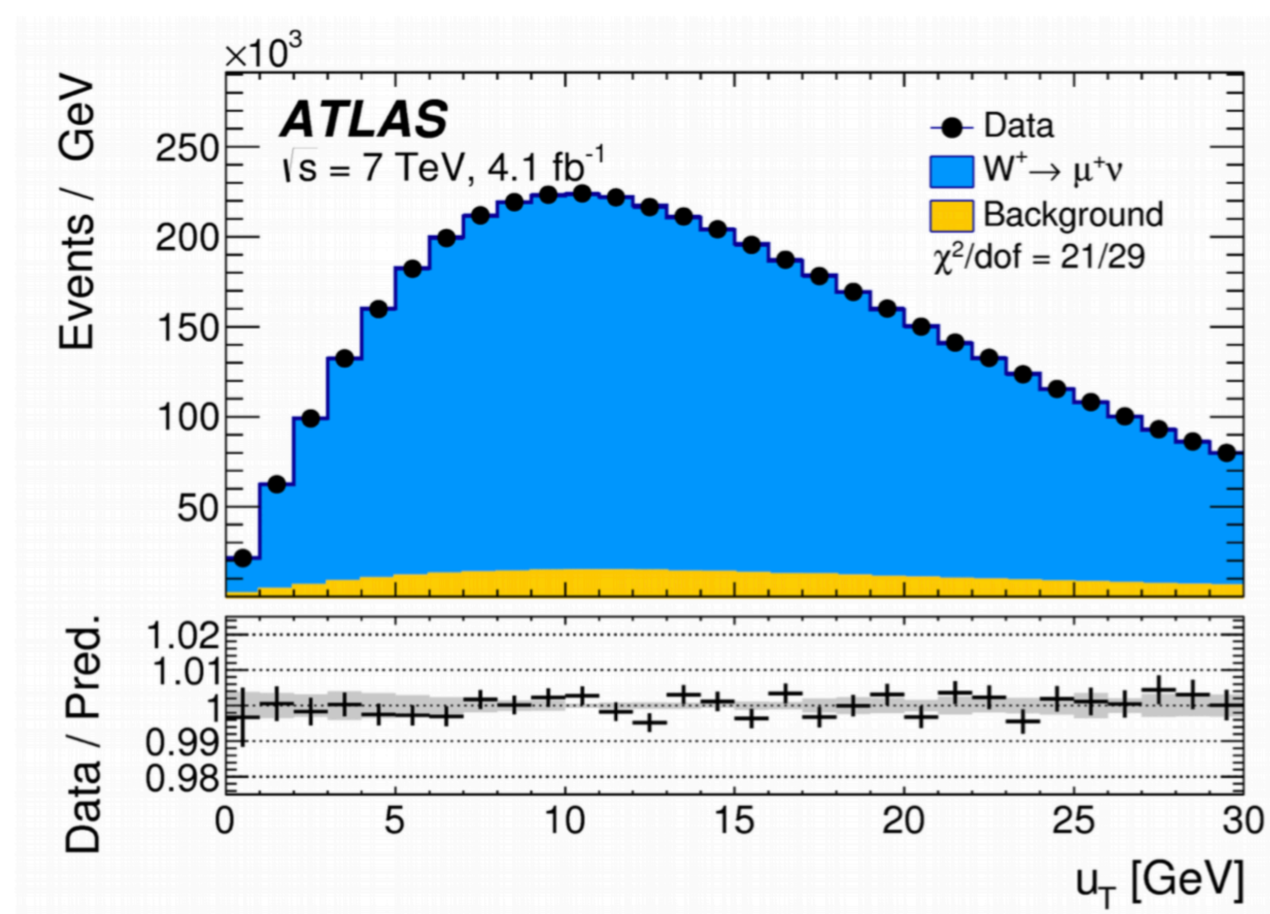
AZ tune

- Pythia8 AZ is a fit to the Z p_T and ϕ^* distributions at 7 TeV
  - Fitting intrinsic-kT, the ISR strong coupling and the ISR cut off
  - The tune describes the pT(Z) data within 2%

	Pythia8
Tune Name	AZ
Primordial k_T [GeV]	1.71 ± 0.03
ISR $\alpha_S^{ISR}(m_Z)$	0.1237 ± 0.0002
ISR cut-off [GeV]	0.59 ± 0.08
$\chi^2_{min}/\mathrm{dof}$	$45.4/32$

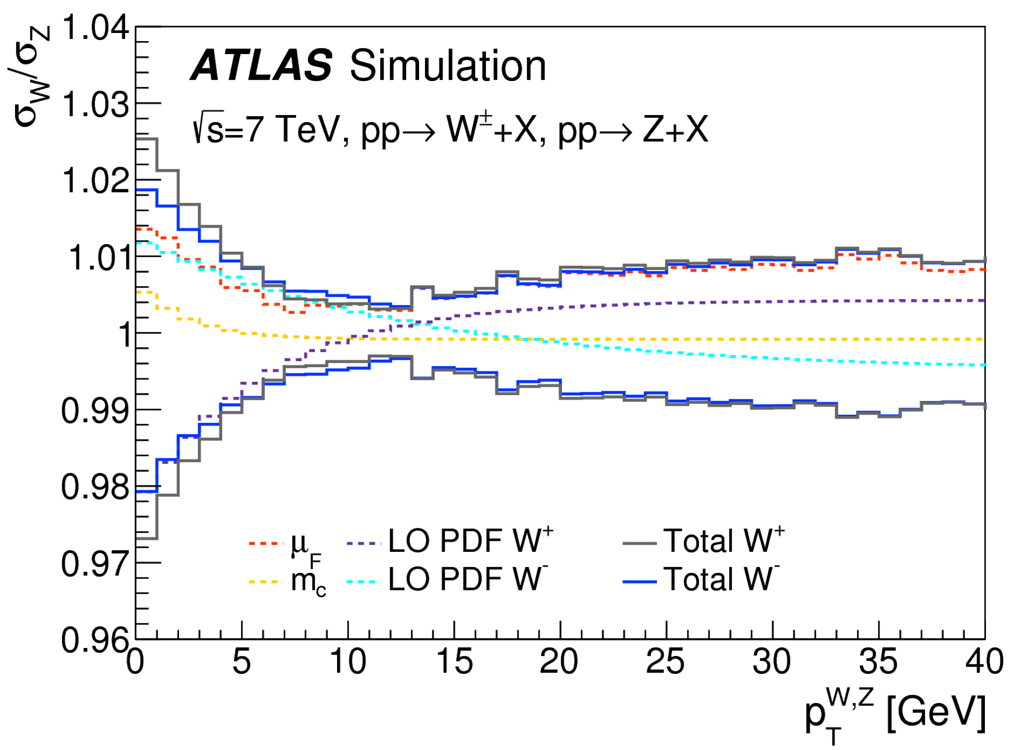
- It is used to predict the p_T W distribution and its uncertainties

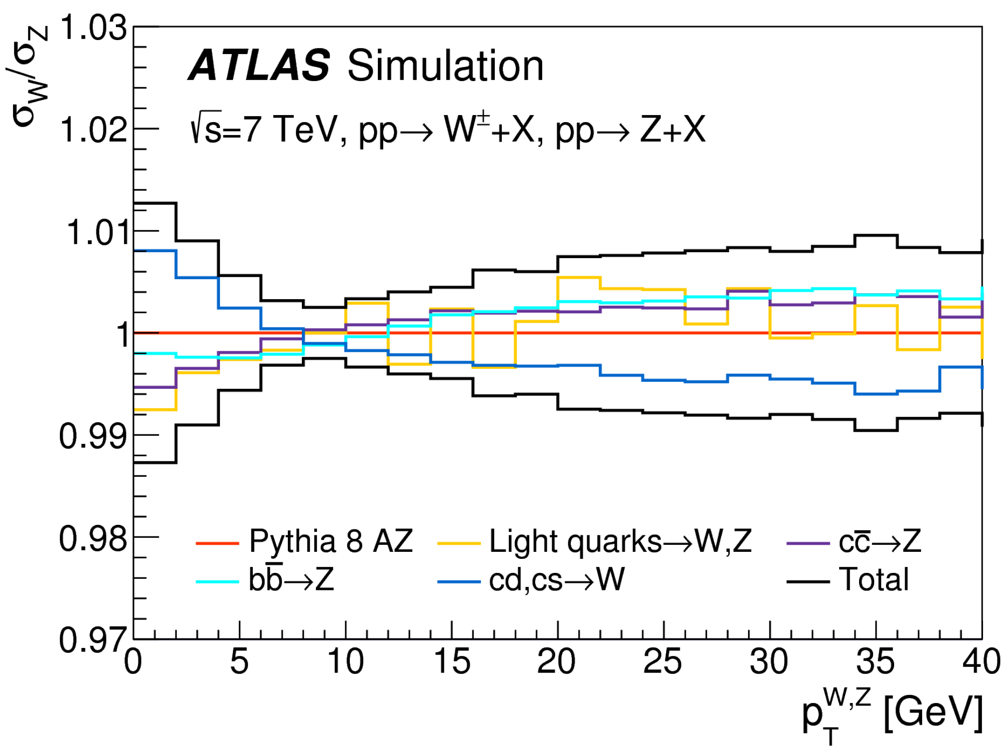




Uncertainties on pT W

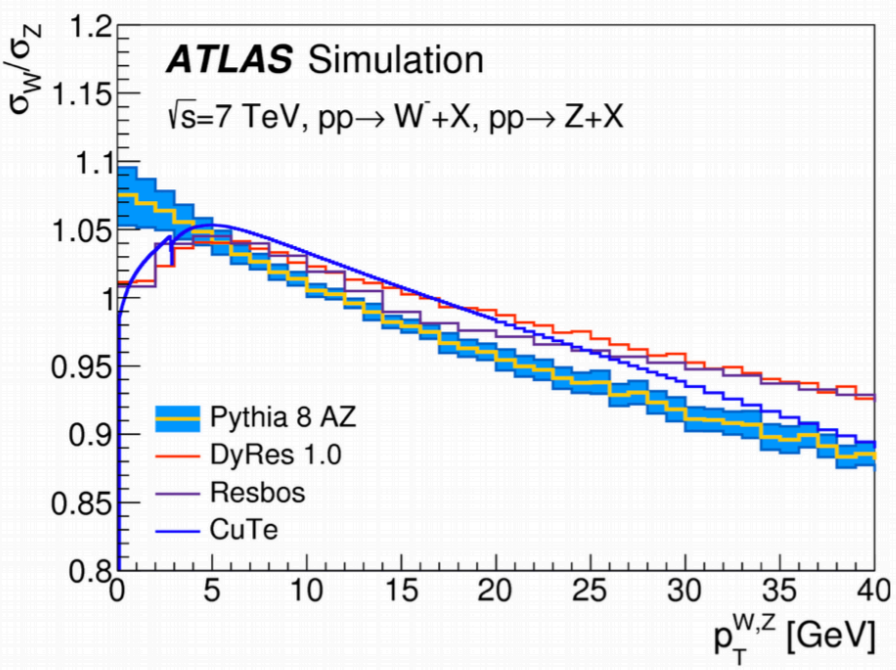
- One large source of differences between Z and W production are the heavy-flavour initiated processes
  - cc→Z and bb→Z are 6% and 3% of Z production, cs→W is ~20% of W production
- As a proxy to vary HF matching scales in the PDFs two variations are performed in the shower
  - Variations of the charm mass
  - Decorrelating μF between light and heavy-flavour processes

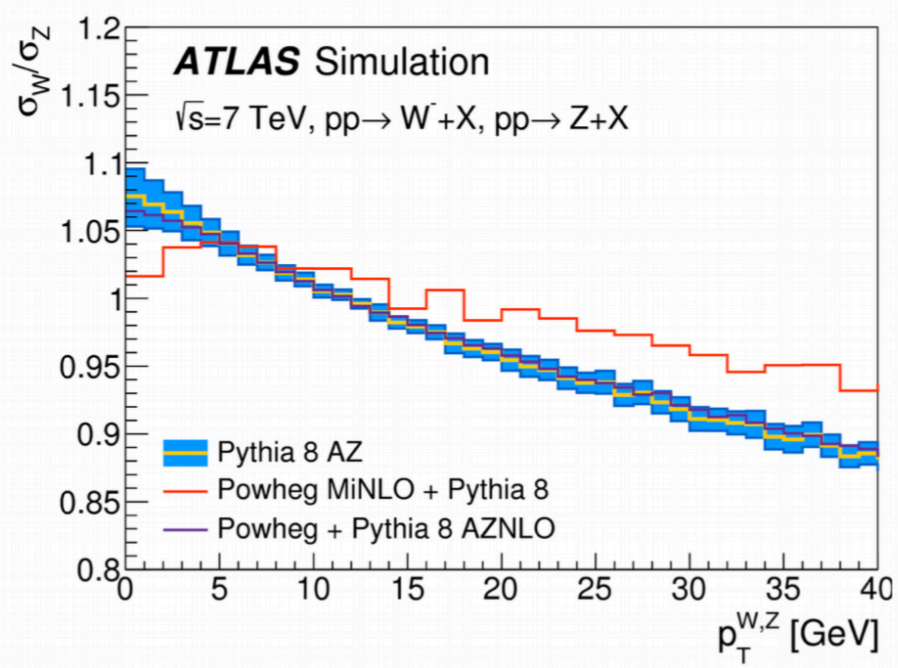


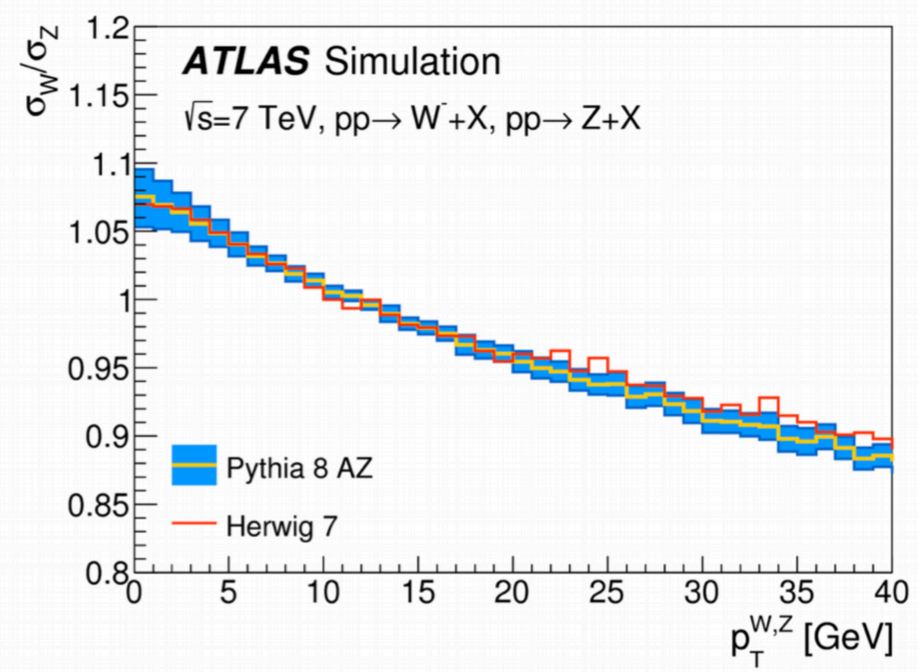


Other predictions

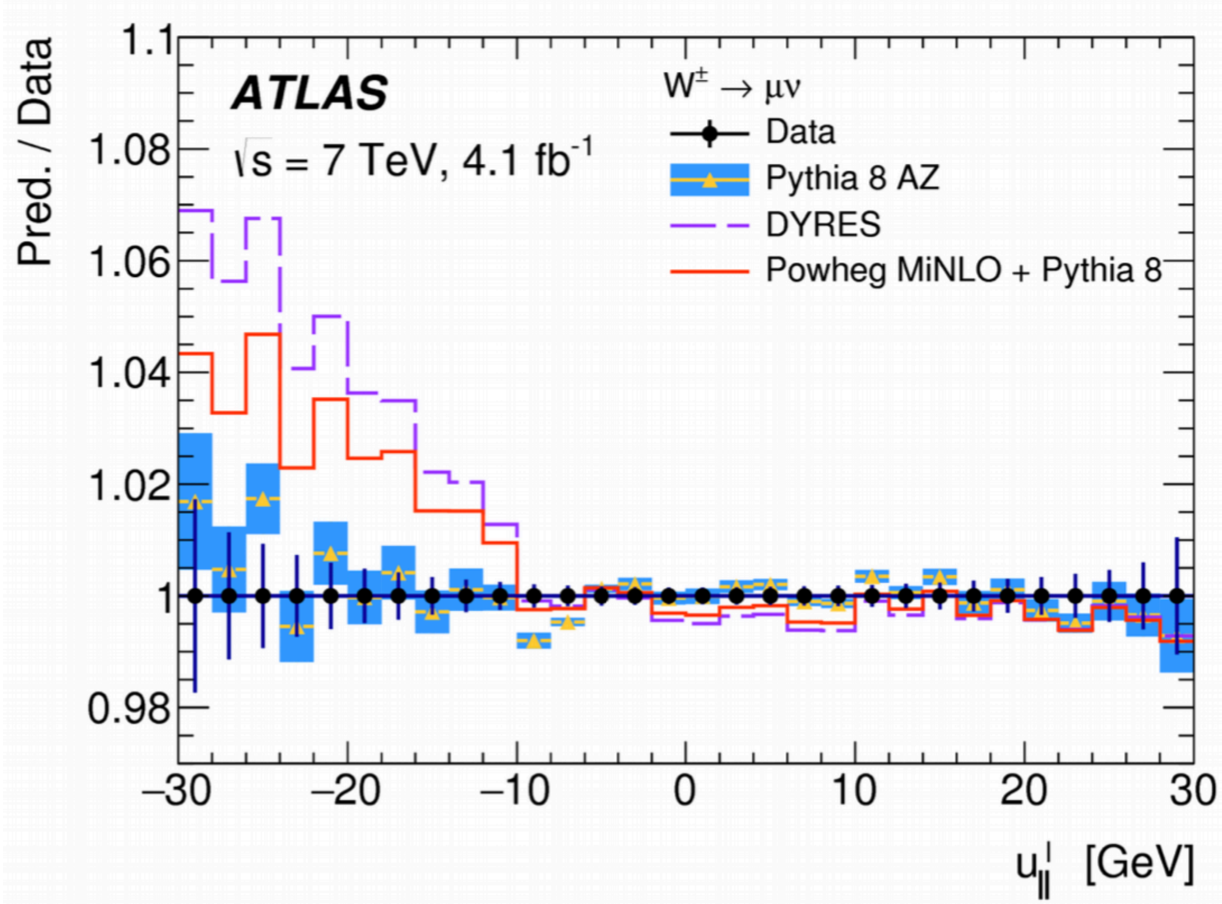
- MINLO and NNLL resummed predictions as are strongly disfavoured by the $u_{||}$ distribution in data







- Effect of heavy-flavour treatment?
- Corrections to the Sudakov due to multi-parton-interactions?
- Poor convergence of the resummation?
- Something else?



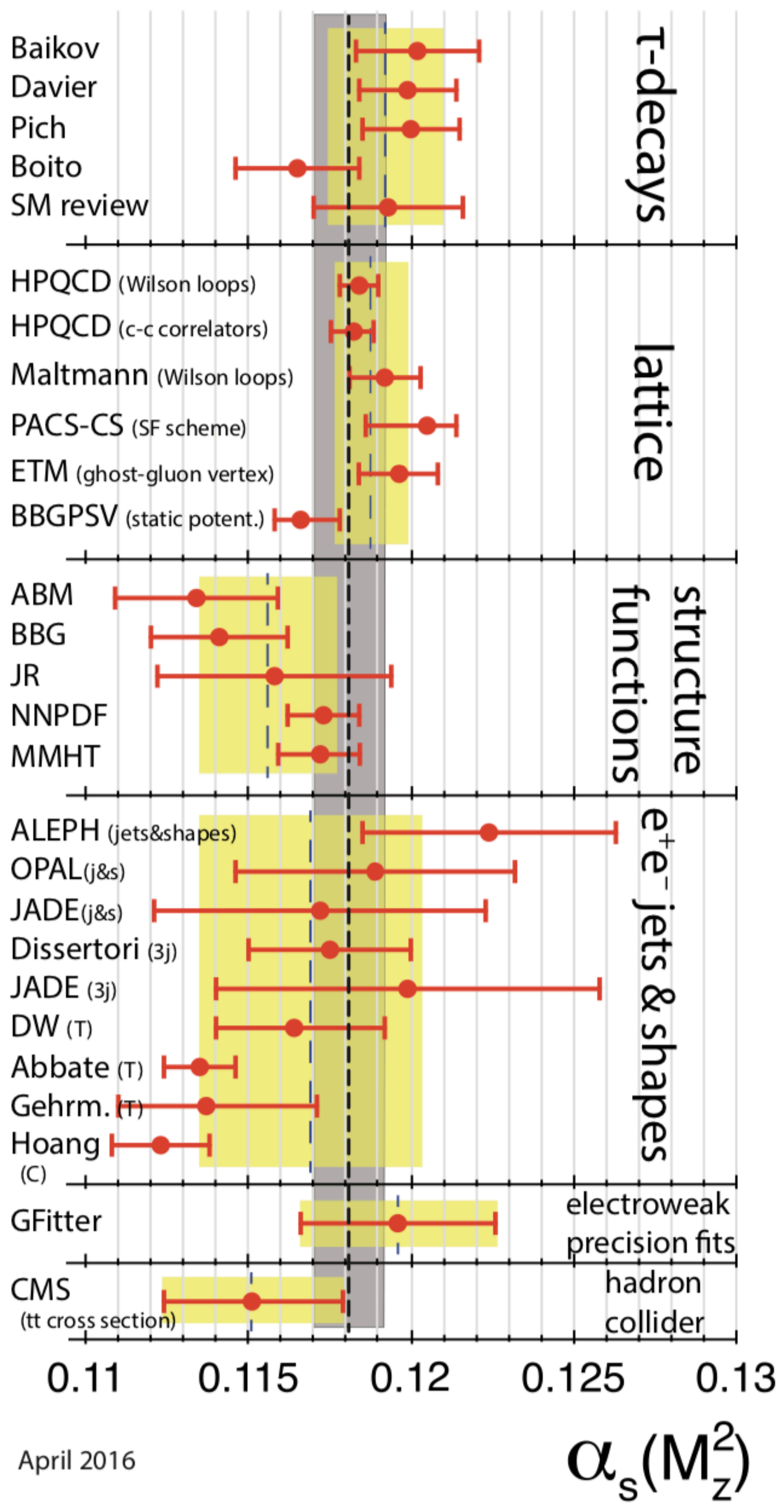
QCD Uncertainties on pT W

- The PDF and from the $p_T(W)$ modelling dominate the QCD modelling uncertainties

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
AZ tune	3.0	3.4	3.0	3.4	3.0	3.4
Charm-quark mass	1.2	1.5	1.2	1.5	1.2	1.5
Parton shower μ_F with heavy-flavour decorrelation	5.0	6.9	5.0	6.9	5.0	6.9
Parton shower PDF uncertainty	3.6	4.0	2.6	2.4	1.0	1.6
Angular coefficients	5.8	5.3	5.8	5.3	5.8	5.3
Total	15.9	18.1	14.8	17.2	11.6	12.9

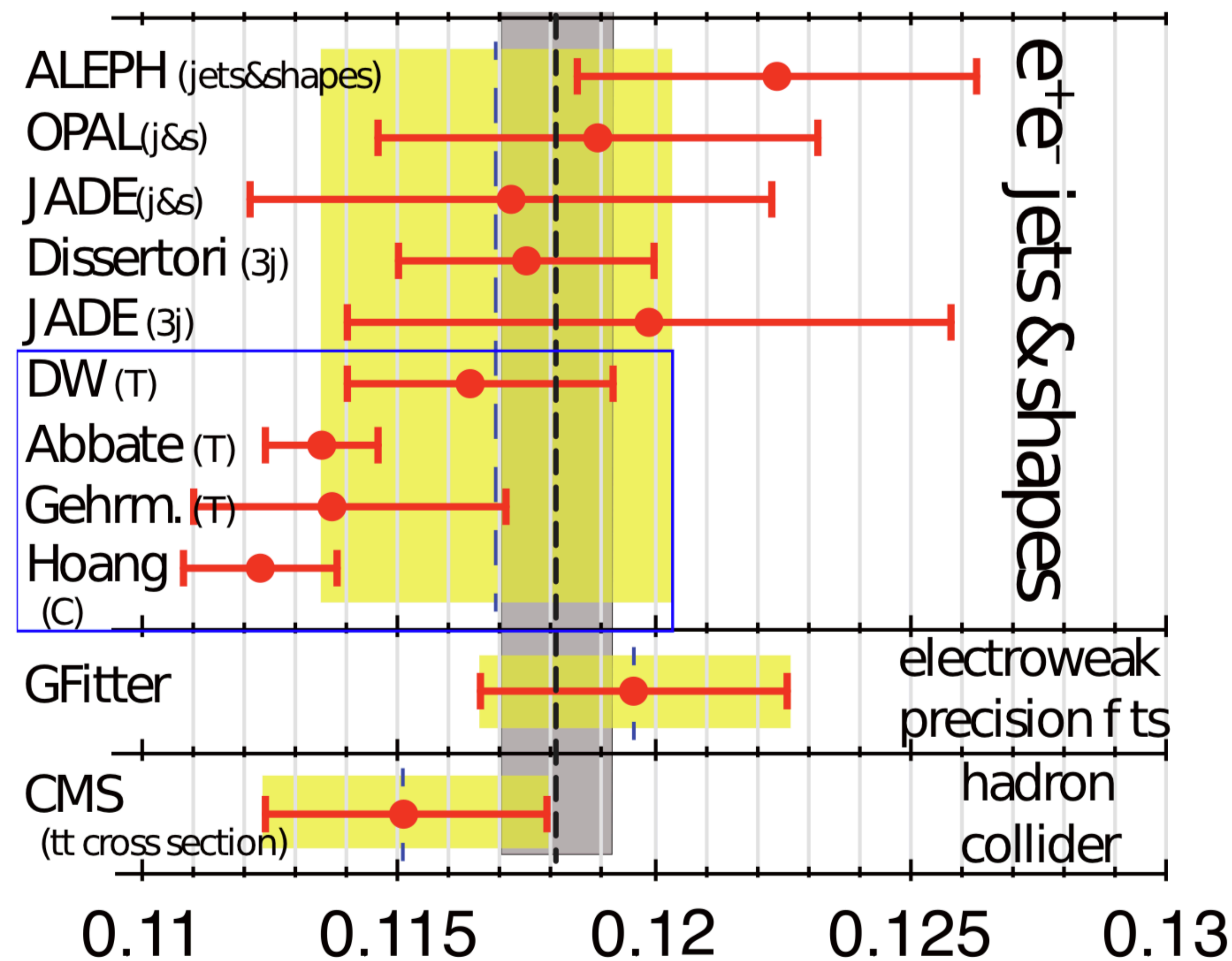
- To improve over the existing W-mass precision it will be crucial to reduce QCD systematics
  - And particularly understand the issues we see with higher-order perturbative predictions

NON-PERTURBATIVE CORRECTIONS TO EVENT-SHAPES



[Particle Data Group; '16]

- Precise measurement of α_s critical, uncertainties feed back into all other theoretical predictions

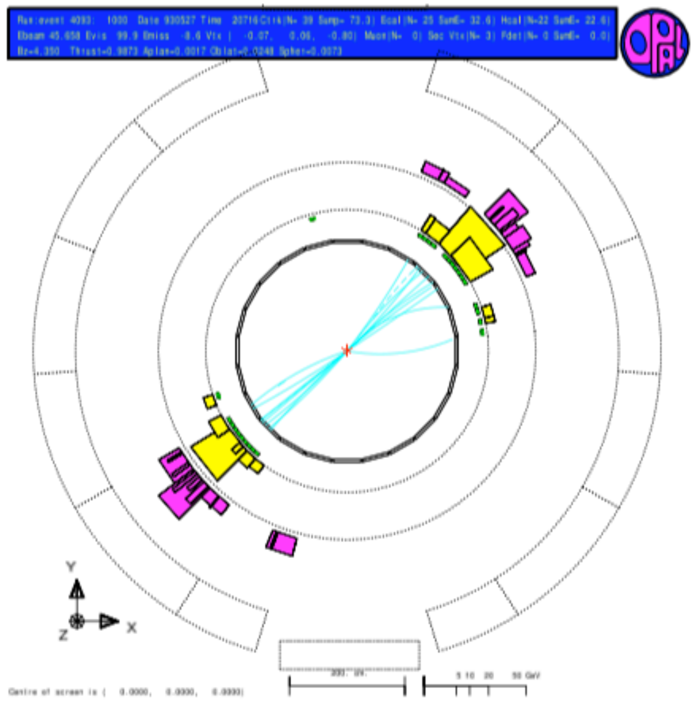


- Event shapes mean geometrical properties of energy flow
  - Sensitive to QCD radiation, used to tune MC, measure α_s
- Many measured at LEP, will focus on the thrust distribution

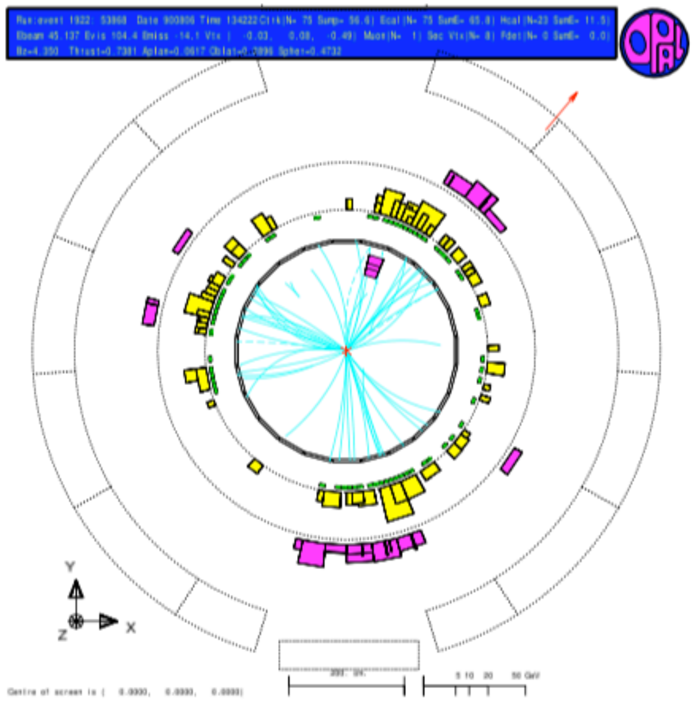
$$\tau \equiv 1 - T = \min_{\vec{n}} \left(1 - \frac{\sum_{i \in \mathcal{E}} |\vec{n} \cdot \vec{p}_i|}{\sum_{i \in \mathcal{E}} |\vec{p}_i|} \right)$$

Longitudinal charged particle alignment

Pencil-like event: $T \lesssim 1$

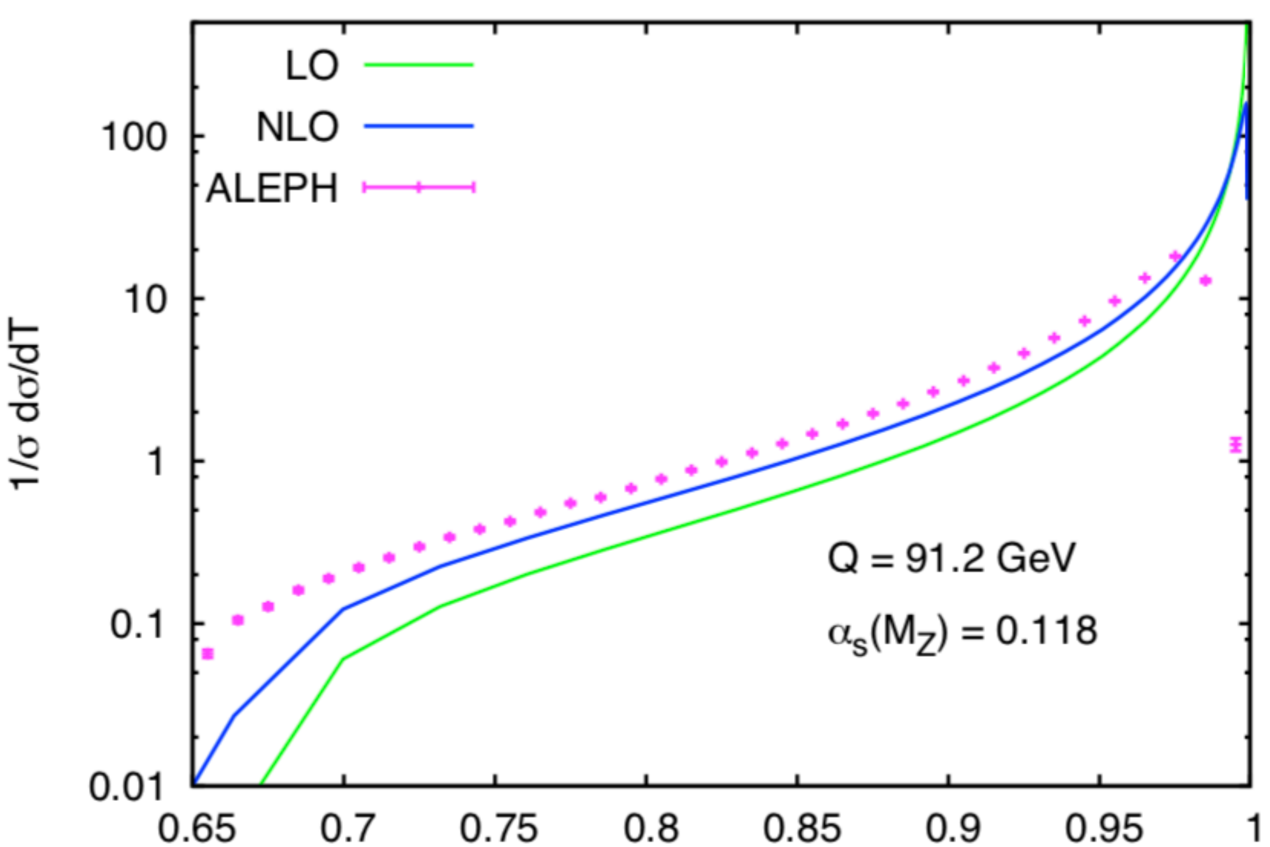


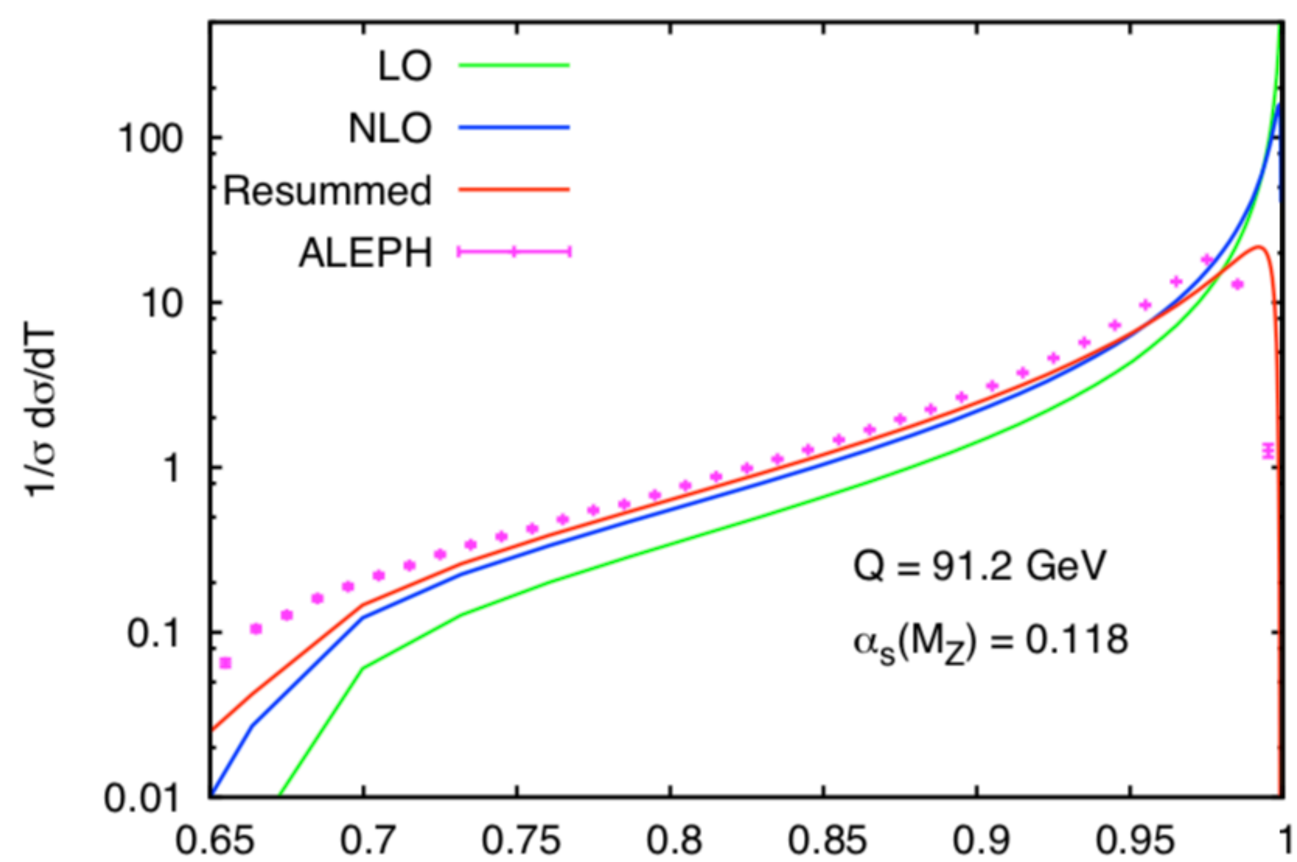
Planar event: $T \gtrsim 2/3$



The Thrust distribution

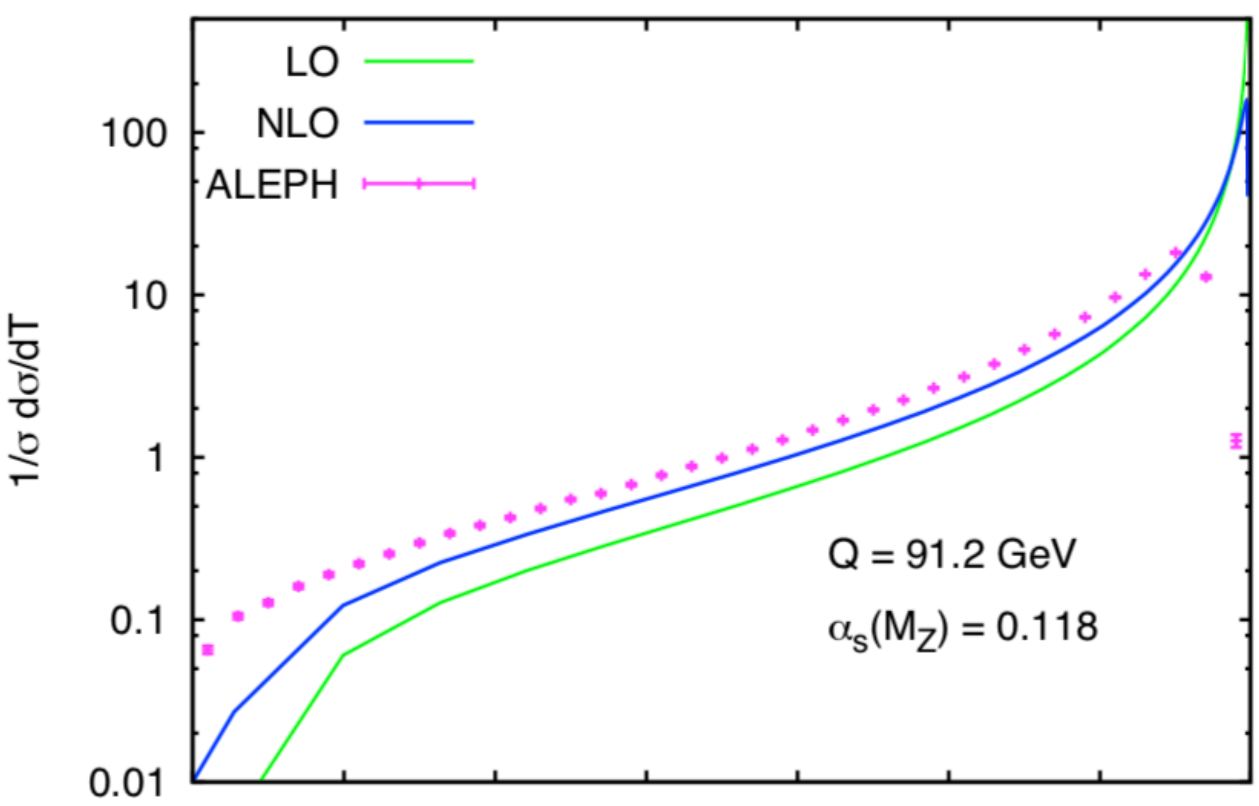
- Fixed order prediction do not reproduce the thrust shape
  - NNLO helps, but not enough
- Only resummed predictions give the correct behaviour in the two-jet limit
- Still a mismatch between perturbative prediction and data
- Agreement can be restored by shifting the distribution



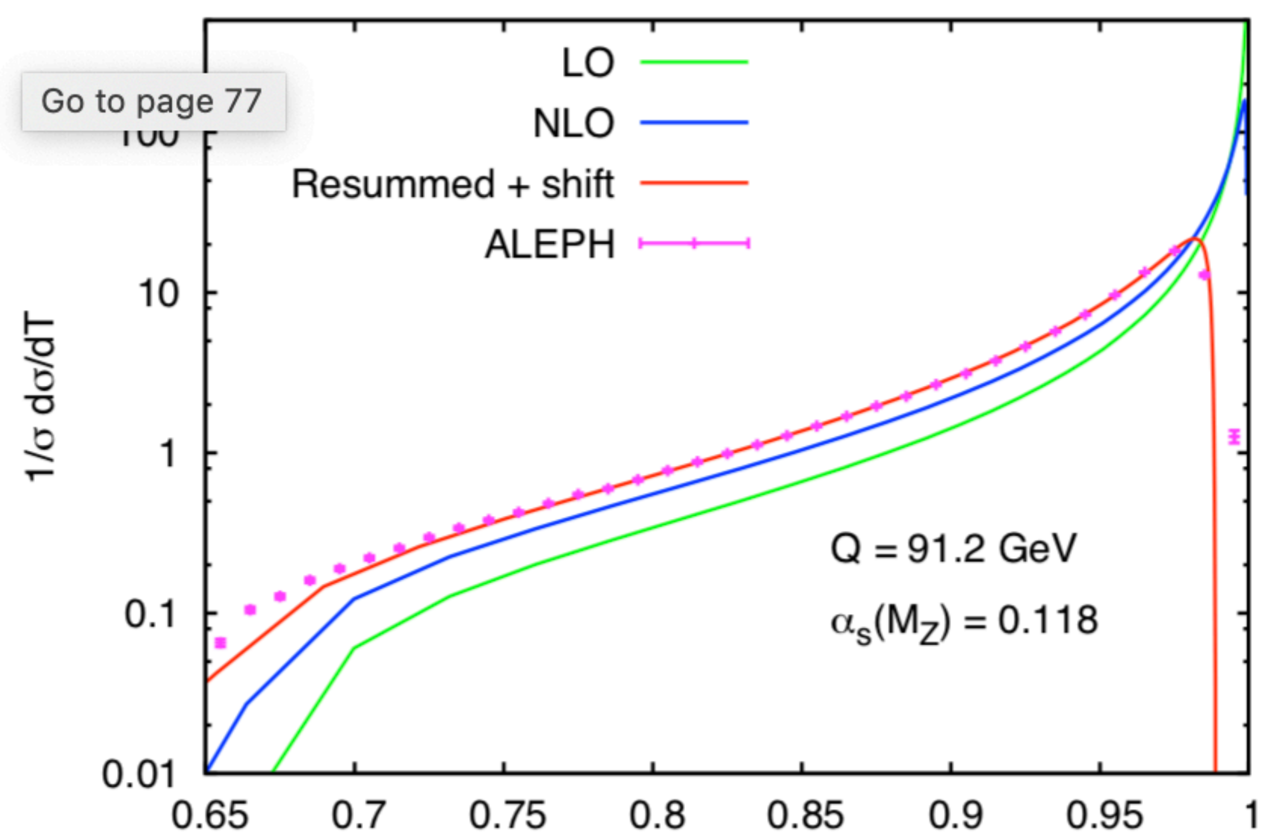


The Thrust distribution

- Fixed order prediction do not reproduce the thrust shape
  - NNLO helps, but not enough
- Only resummed predictions give the correct behaviour in the two-jet limit
- Still a mismatch between perturbative prediction and data
- Agreement can be restored by shifting the distribution
  - A way to incorporate leading hadronization corrections, which scale as 1/Q



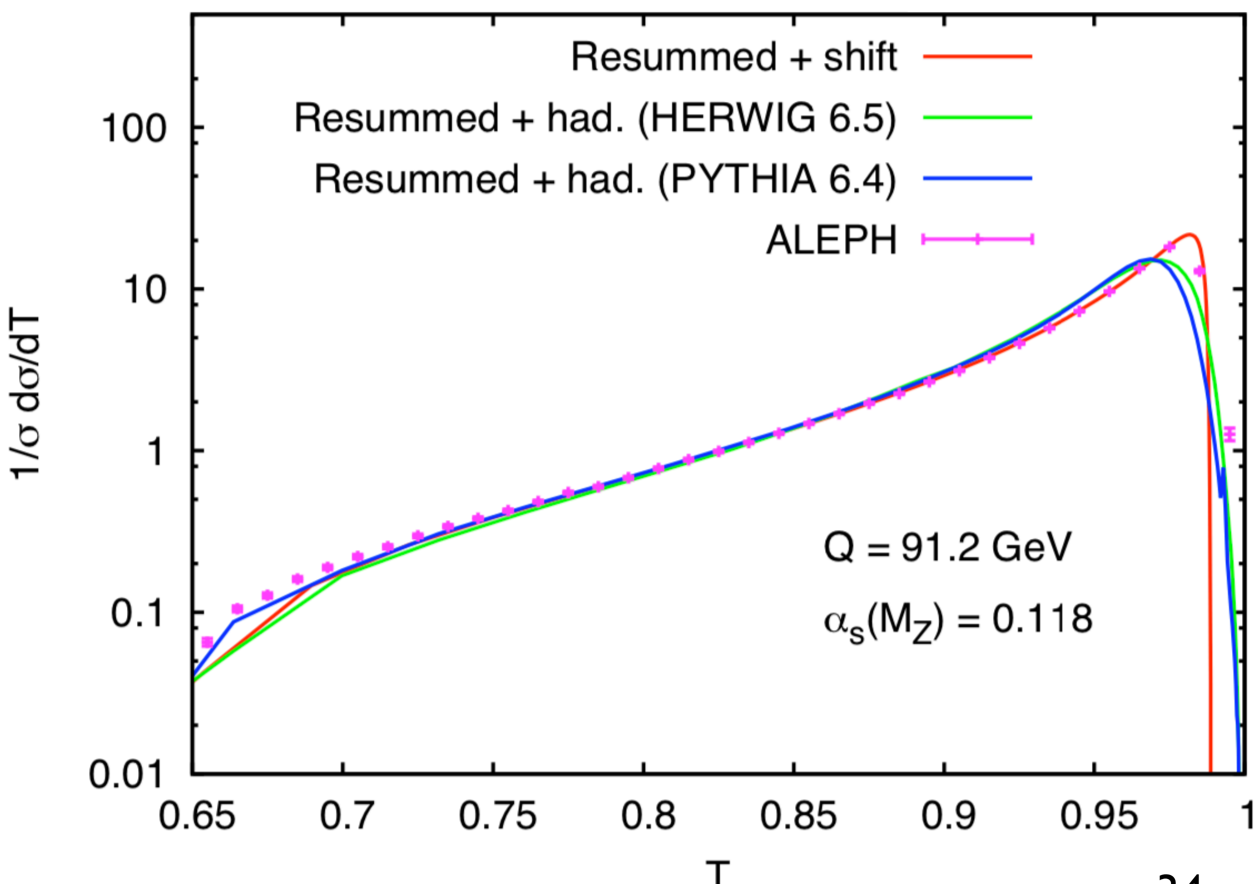
Go to page 77



The Thrust distribution

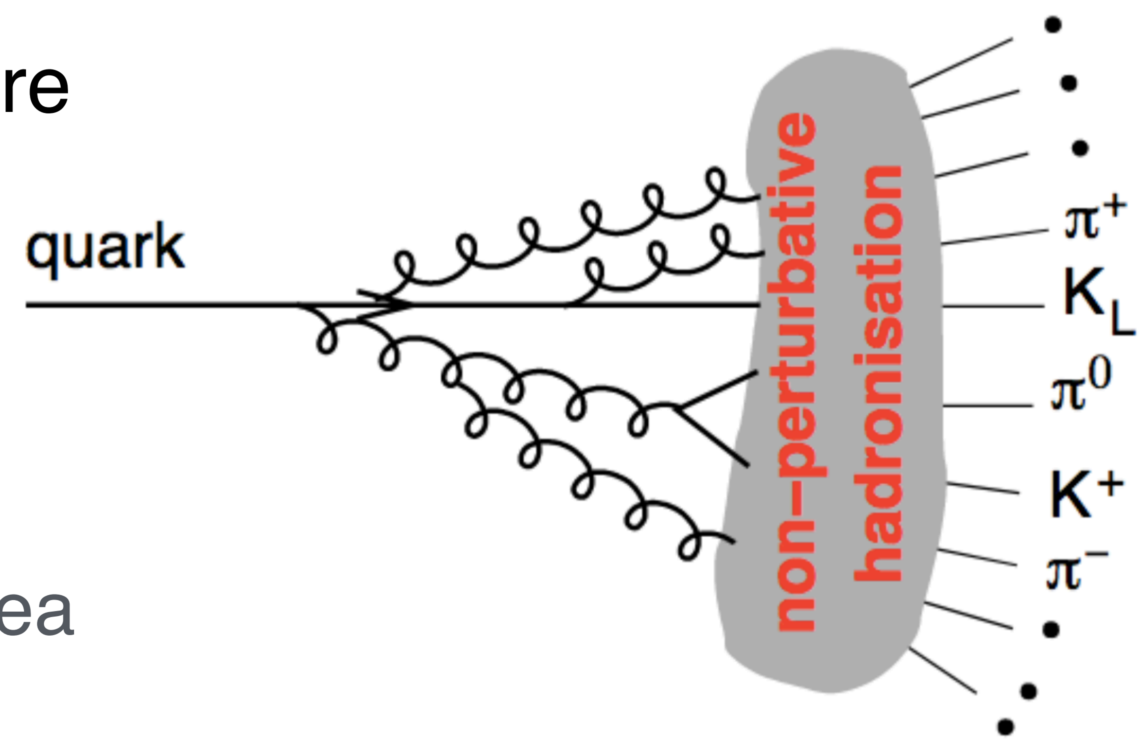
- Mismatch between perturbative predictions and the data ➡ hadronization
- MC generators describe event shapes very well
- Can incorporate hadronization corrections into a perturbative calculation as

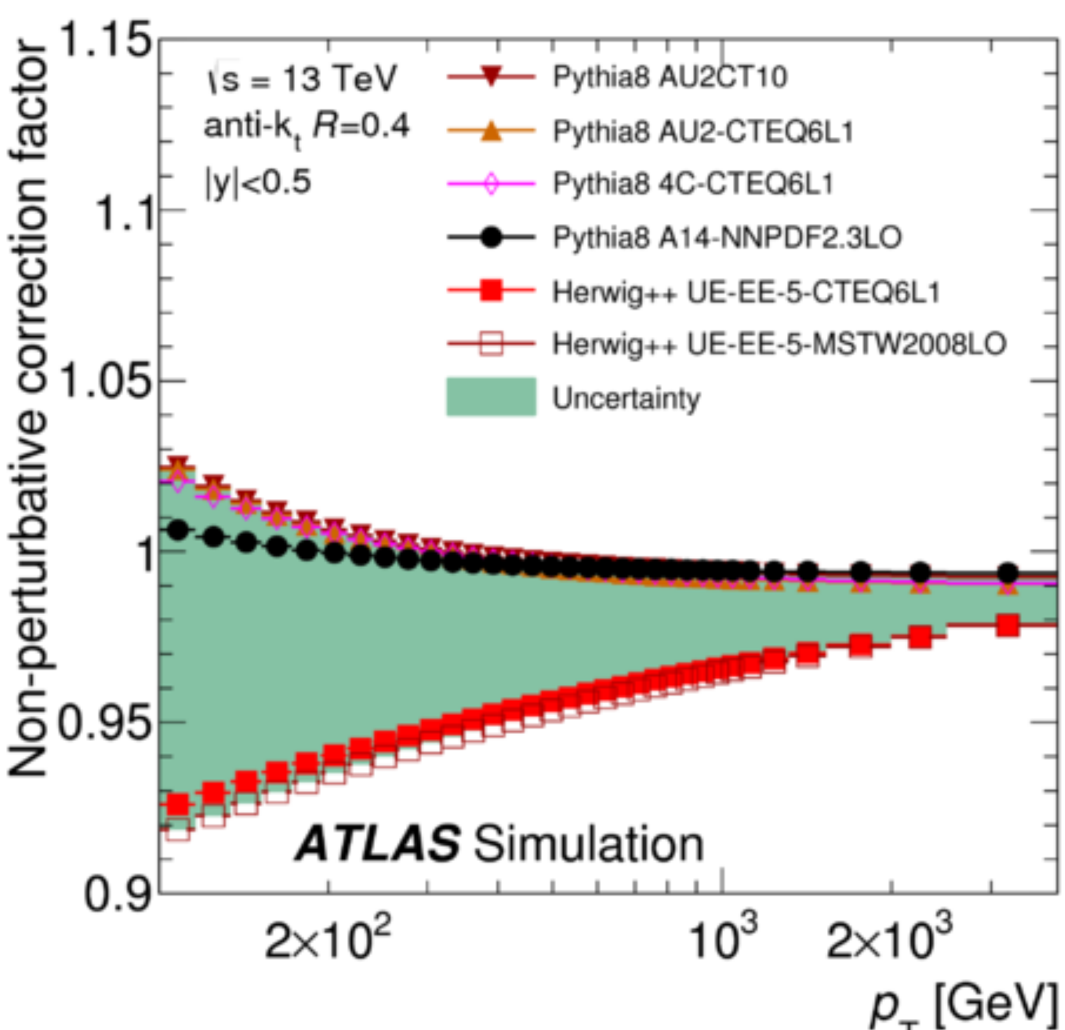
$$\left[\frac{1}{\sigma}\frac{d\sigma}{dT}\right]_{\mathbf{hadron}} = \mathrm{hadcor} \otimes \left[\frac{1}{\sigma}\frac{d\sigma}{dT}\right]_{\mathbf{parton}}$$



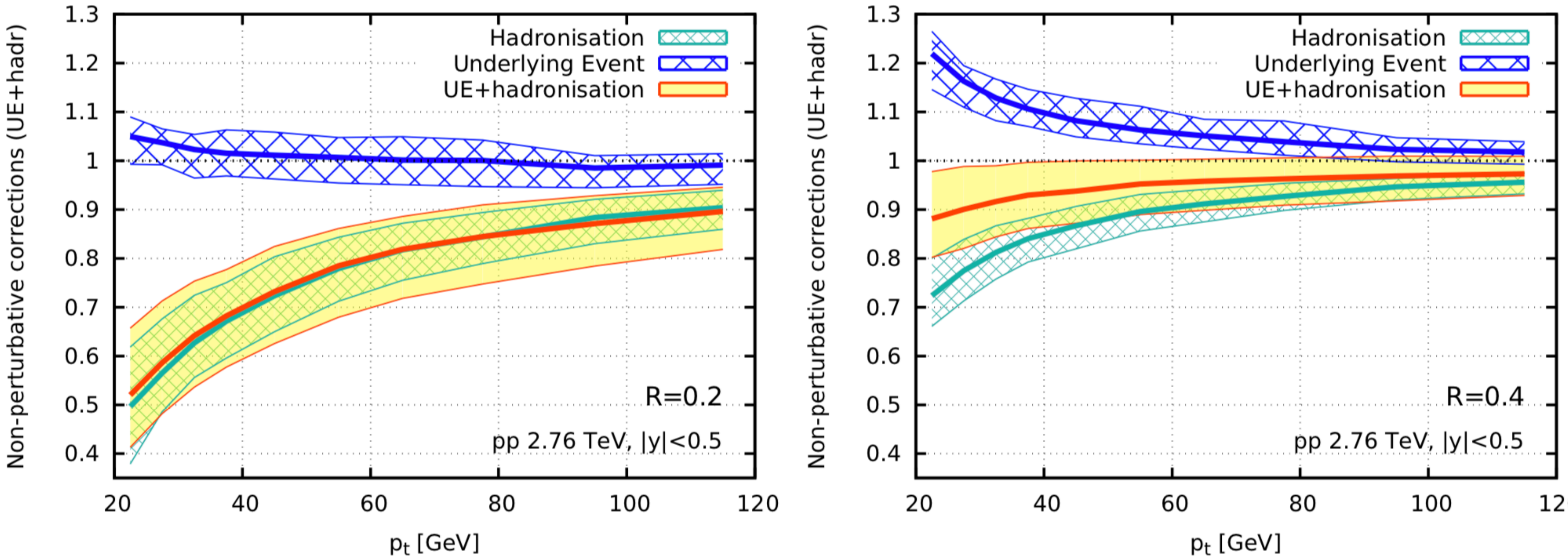
- After tuning the hadronization parameters of the MC we can use it to extract the strong coupling
- Striking success $\alpha_s(M_Z) = 0.1202 \pm 0.0050$
- Hadronization corrections will be of the same size as the perturbative one at high energy lepton colliders

- Jet cross-sections and observables are affected by non-perturbative effects
  - Particles might get just enough momentum from hadronization to go outside of the jet cone
  - Particles from MPI will enter the jet area





- Usually this effect is evaluated with Monte Carlos
  - By taking the difference between NP effects on and off
  - But large spread between NP corrections evaluated with different codes



- NP corrections affect differently jets of different radii
  - Hadronization negative and decreasing with radius
  - Underlying event positive and growing with radius
- Measurement of the jet shapes are able to probe this structure and constrain MC parameters

Which jet size ?

Parton $p_t \longrightarrow$ jet p_t

Ill-defined: MC "parton"

PT radiation:

$$q: \quad \Delta p_t \simeq \frac{\alpha_s C_F}{\pi} p_t \ln R$$

$$g: \quad \Delta p_t \simeq \frac{\alpha_s C_A}{\pi} p_t \ln R$$

Hadronisation:

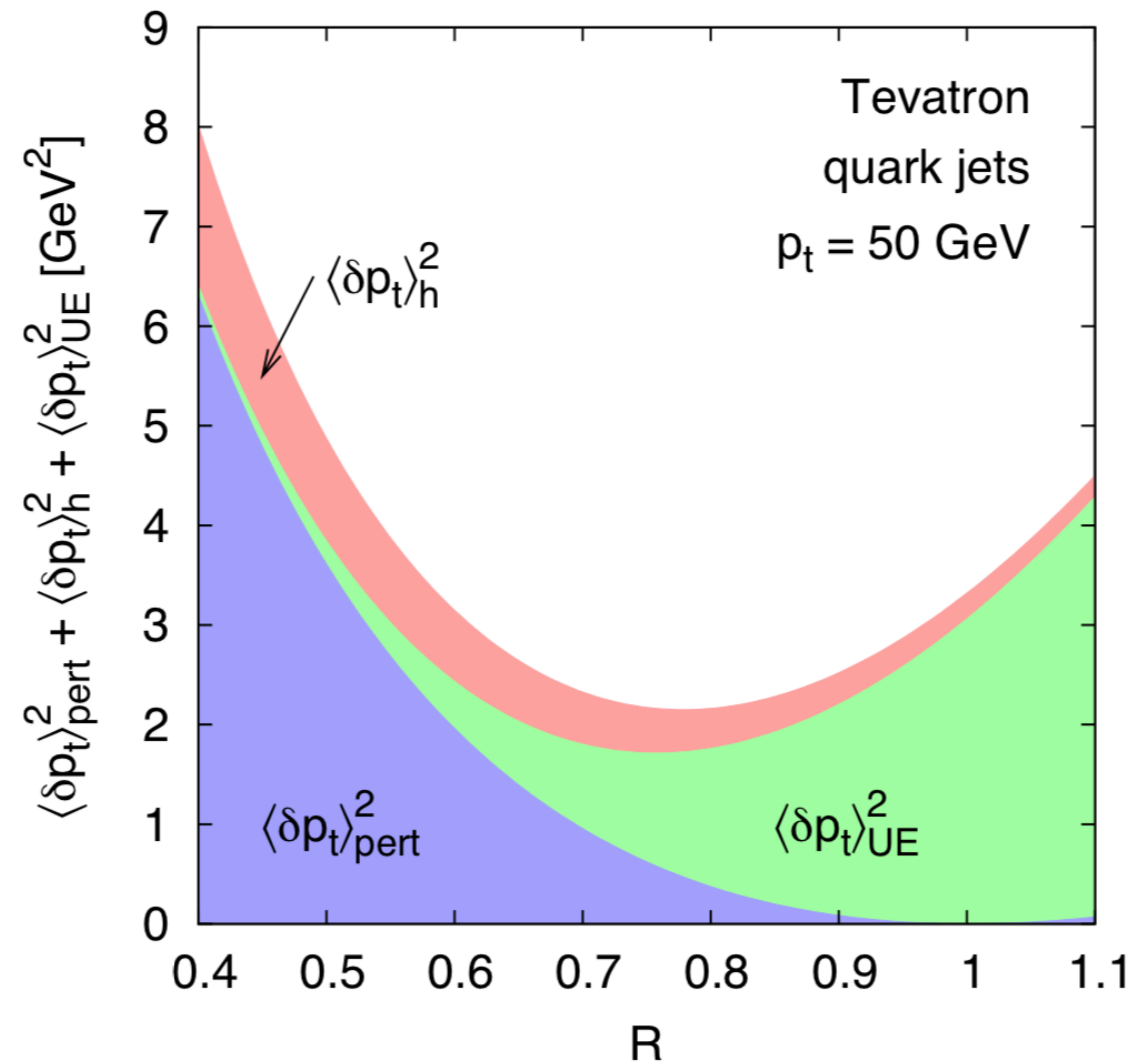
$$q: \quad \Delta p_t \simeq \frac{C_F}{R} \cdot 0.4 \text{ GeV}$$

$$g: \quad \Delta p_t \simeq \frac{C_A}{R} \cdot 0.4 \text{ GeV}$$

Underlying event:

$$q, g: \quad \Delta p_t \simeq \frac{R^2}{2} \cdot 2.5\text{–}15 \text{ GeV}$$

[1602.01110]



Parton $p_t \longrightarrow$ jet p_t

Ill-defined: MC "parton"

[1602.01110]

PT radiation:

$$q: \quad \Delta p_t \simeq \frac{\alpha_s C_F}{\pi} p_t \ln R$$

$$g: \quad \Delta p_t \simeq \frac{\alpha_s C_A}{\pi} p_t \ln R$$

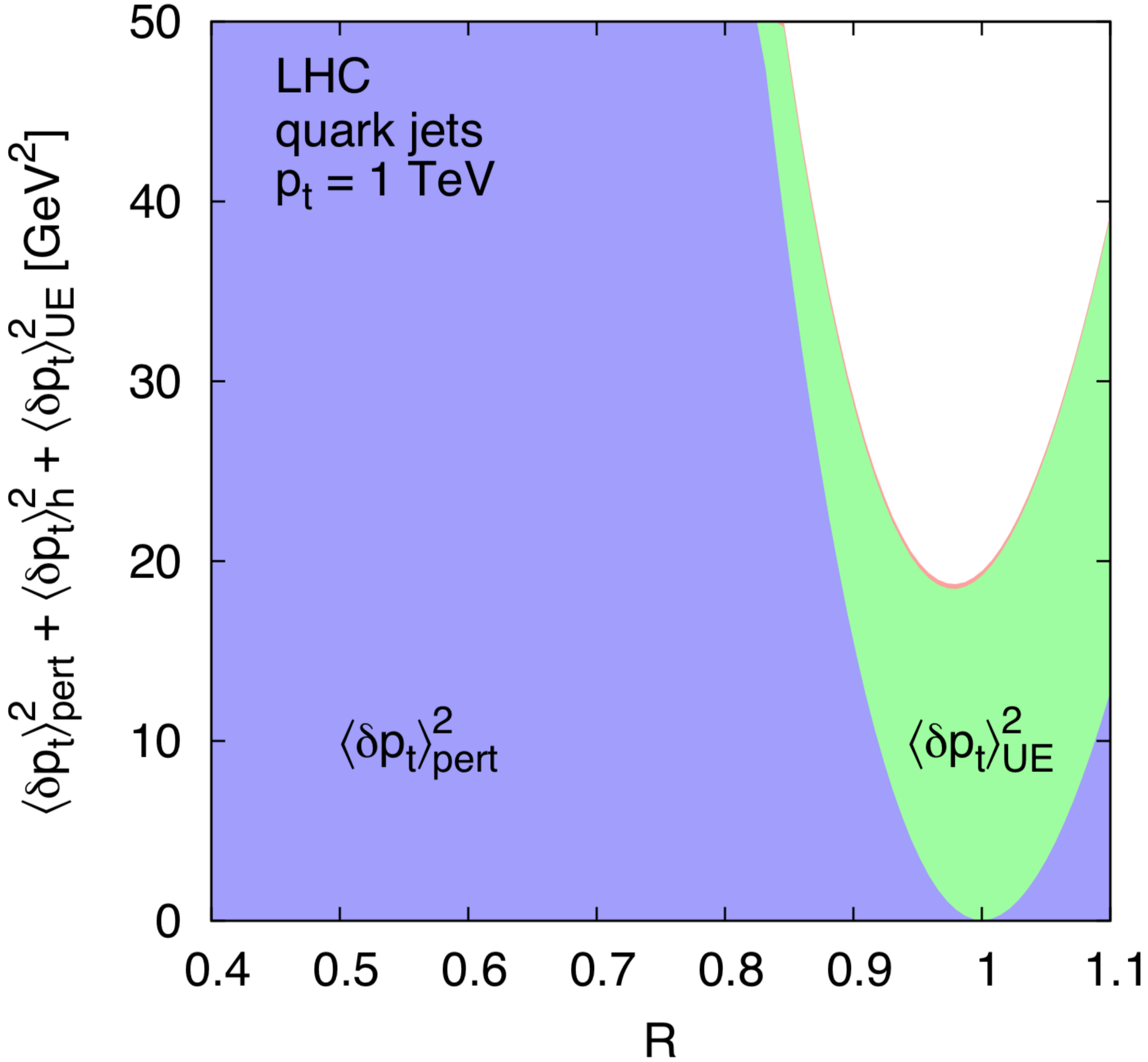
Hadronisation:

$$q: \quad \Delta p_t \simeq \frac{C_F}{R} \cdot 0.4 \text{ GeV}$$

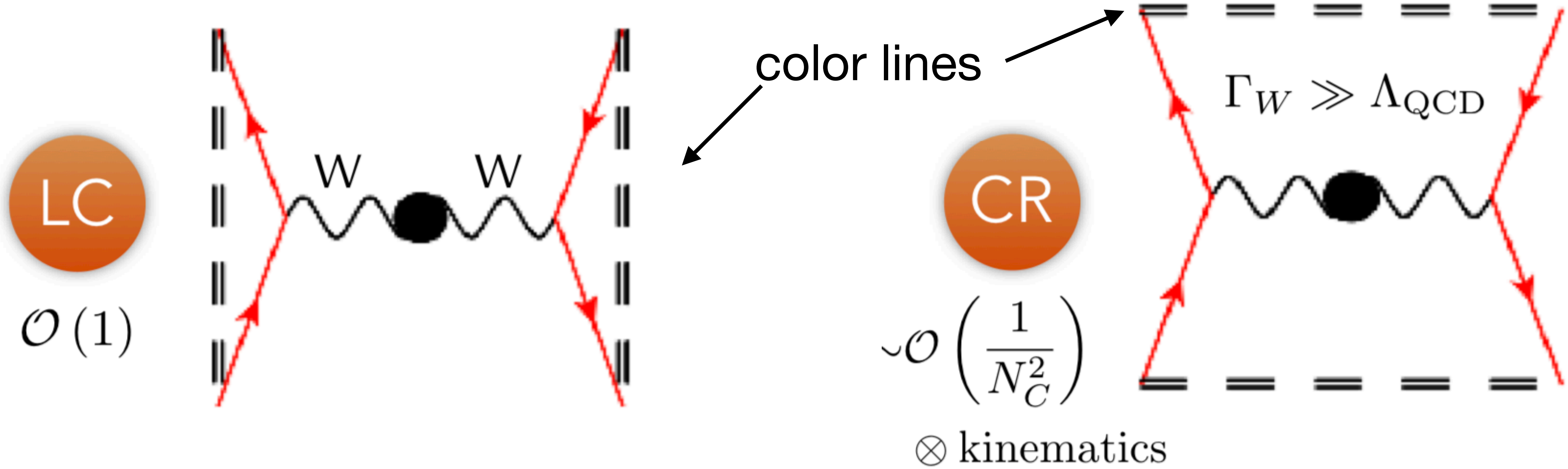
$$g: \quad \Delta p_t \simeq \frac{C_A}{R} \cdot 0.4 \text{ GeV}$$

Underlying event:

$$q, g: \quad \Delta p_t \simeq \frac{R^2}{2} \cdot 2.5\text{–}15 \text{ GeV}$$



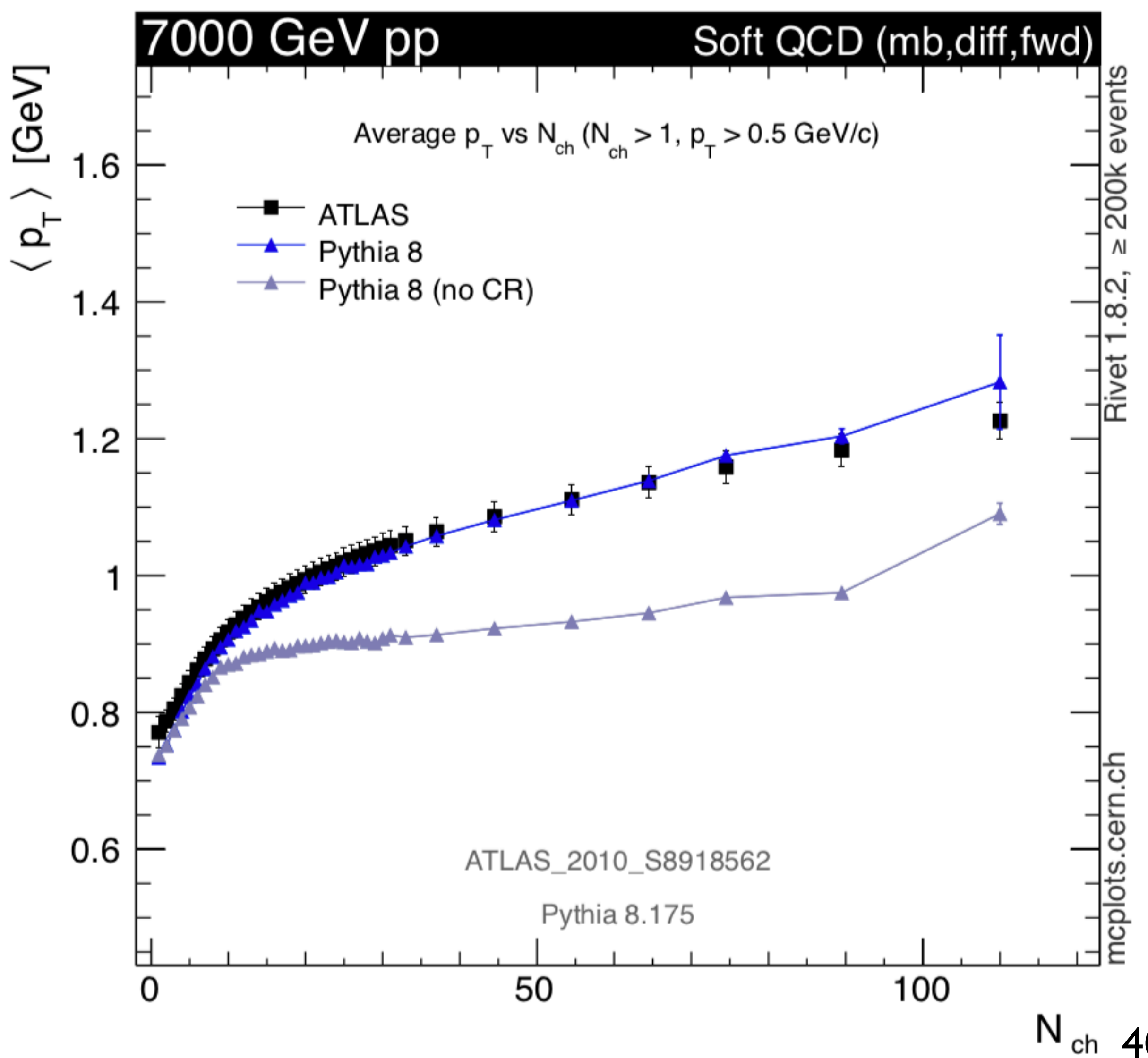
- First studied at LEP2 in e+e- → WW → 4 jets
  - Color reconnection implied a non-perturbative uncertainty on the W mass measurement, $\Delta M_W \sim 40$ MeV
  - Best fit of CR strength ~ 10% ~ $1/N_C^2$



- In pp collisions many more coloured particles involved, we can expect color reconnection effect to be more important
  - Coloured particles in the initial state
  - Multiple partonic interactions

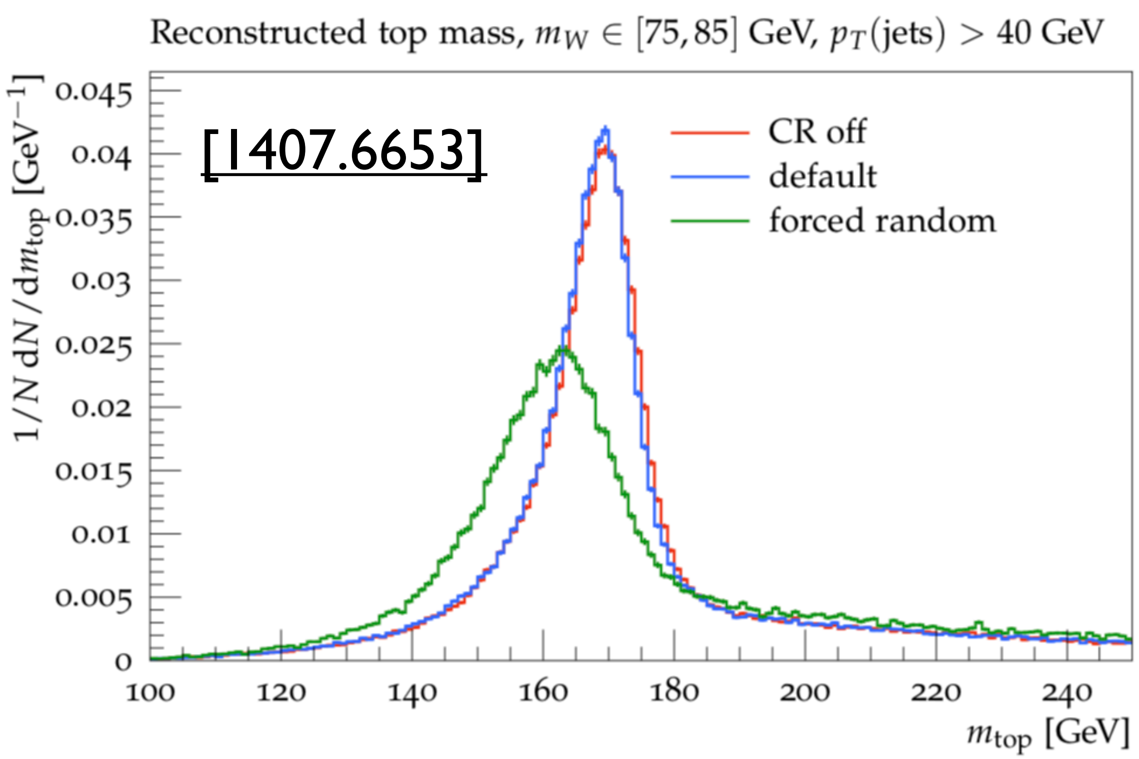
Color reconnection and top mass

- In pp collisions CR is essential to describe $<p_T>(n_{ch})$
- CR acts before hadrons are generated. The ends of the colour lines are reconnected, resulting in a different colour configuration.



- In MC generators CR models typically attempt to minimise string lengths or cluster masses
  - With color-reconnection the MPIs hadronize collectively
  - Trade particle multiplicity for an increased p_T

- Top mass measurements have reached unprecedented precision
  - CMS m_t = 172.35±0.16(stat)±0.48 (sys) GeV (PRD93 (2016) 072004)
  - Effect from CR estimated to be ~ 100 MeV by switching off CR in Pythia6



Δm_{top} relative to no CR:

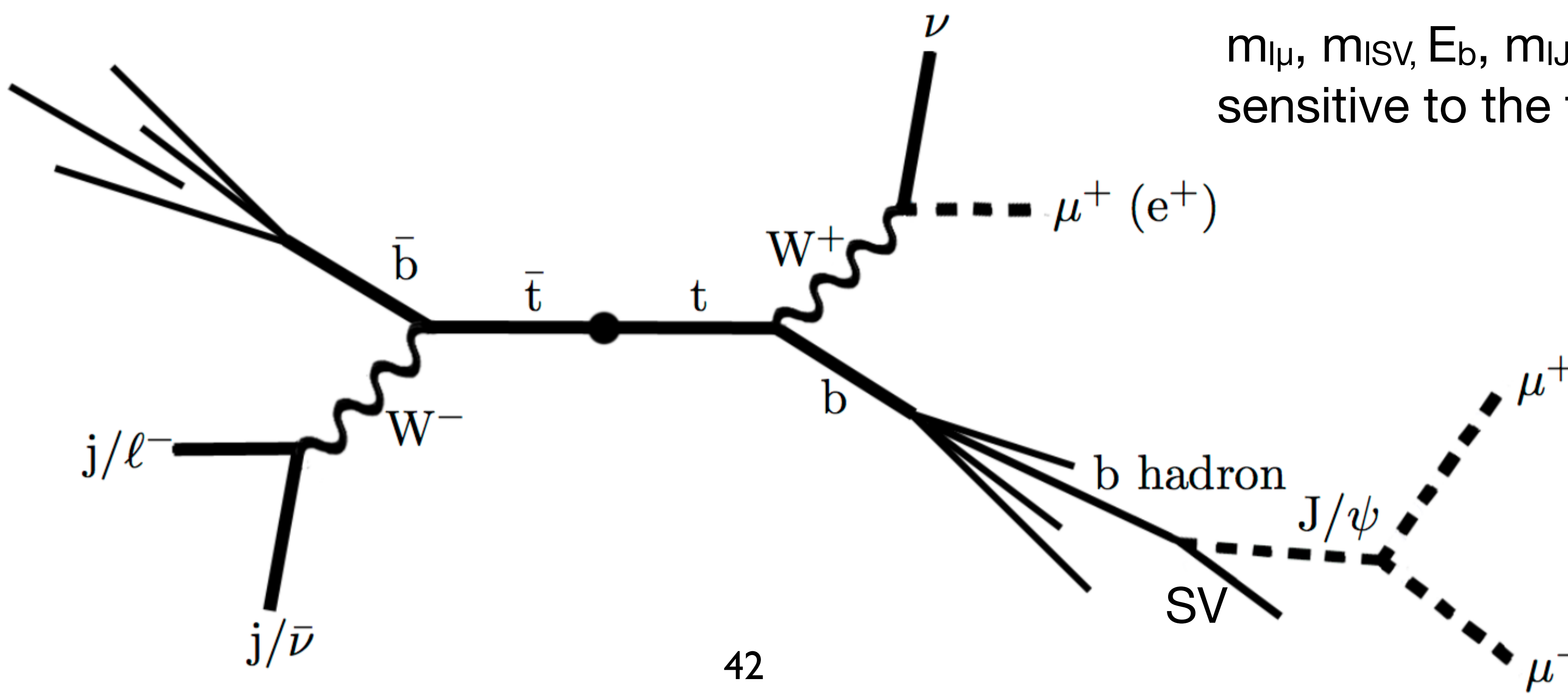
model	Δm_{top} [GeV]	Δm_{top} rescaled
default (late)	−0.415	+0.209
default early	+0.381	+0.285
forced random	−6.970	−6.508

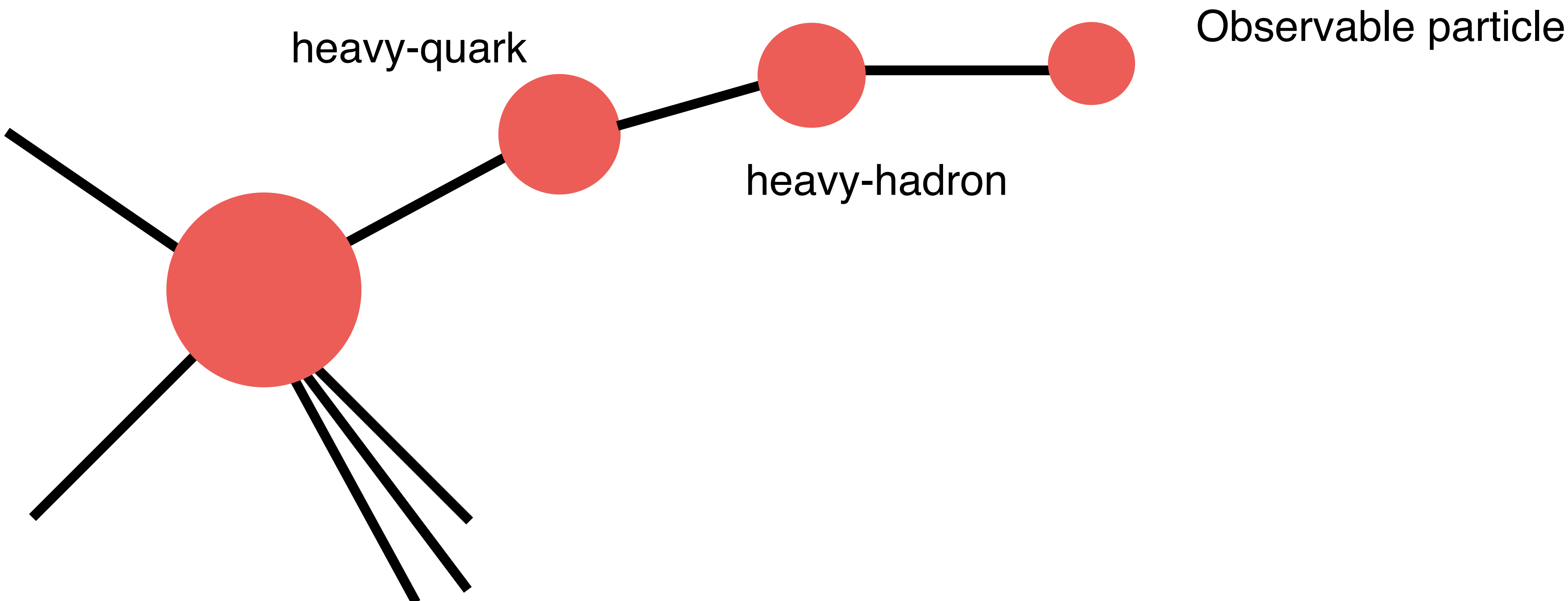
- Top Width ~ 1.5 GeV close to hadronization scale
  - Top decay product also reconnect?
- Several extreme CR models are implemented and studied in Pythia8
  - Comparing them CR systematic on top mass estimated to be of order 0.50 GeV
  - Larger than what one gets varying the parameters within any given model
  - For more precise measurement crucial to further constrain these models using jet profiles and underlying event in ttbar

TOP MASS FROM B-HADRON DECAYS

- Can look at properties of the fragmentation of the b-quark as proxies to the top mass:
  - First paper from CDF used the *B-hadron decay length* [***0612061***]
  - CMS published 8 TeV measurements using samples of *ttbar+J/ψ* [***1603.06536***] and *ttbar+charged particles* [***1608.03560***]
- Little sensitivity to the jet energy scale (which dominates standard methods)
- Smaller dataset and a significant dependence on the modelling of fragmentation

$m_{l\mu}$, m_{lSV}, E_b, $m_{lJ/\psi}$, … are sensitive to the top mass





* A factorisation structure, similar as what we saw for PDFs, is true for fragmentation functions

Hard process
(computed in pQCD)

Weak decay

$$\frac{d\sigma(b \to B \to J/\psi)}{dp_T} = \frac{\sigma(b)}{d\hat{p}_T} \otimes f(b \to B) \otimes g(B \to J/\psi)$$

Universal Non-perturbative
fragmentation function

Fragmentation Functions

- Fragmentation functions cannot be calculated perturbatively, but their *s*-dependence (scaling violations) is given by the DGLAP equations:

$$s\frac{\partial}{\partial s}D_i^h(x,s) = \sum_j \int_x^1 \frac{dz}{z} P_{ji}(z,\alpha_S(s))D_j^h(x/z,s) \; .$$

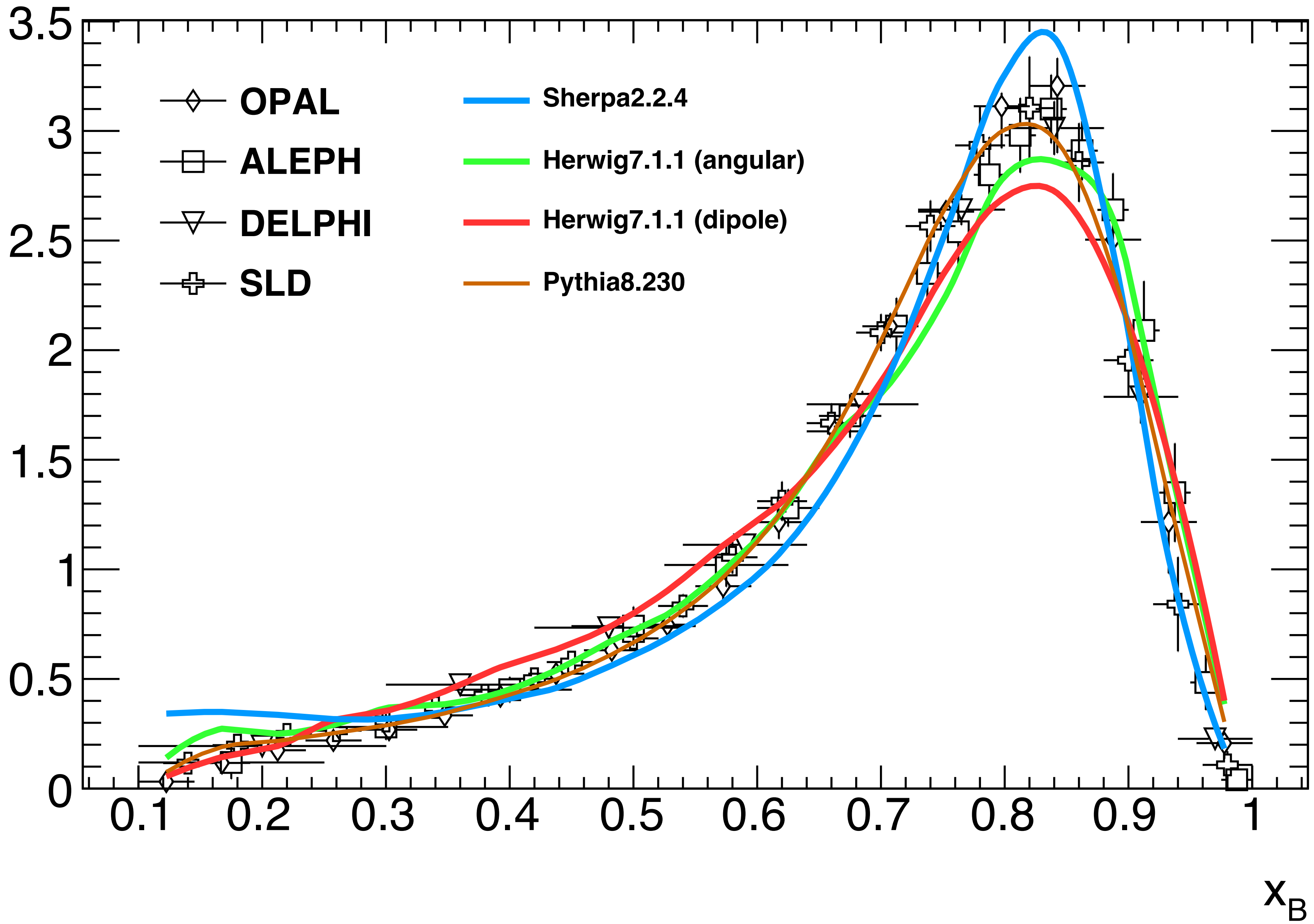
- They can be parametrised as some fixed scale s_0 and then predicted at other energies

- Until now I have not mentioned MC event generators at all
  - One could take an NLL calculation; fit the non perturbative fragmentation function to some data, then use it to predict a B-hadron spectrum
  - But this approach would not give us a fully exclusive description of the event and would not include effects from underlying-event or colour reconnection

COMPARED TO EVENT GENERATORS

- Fragmentation function for bottom-quarks has been measured by LEP/SLD experiments

$$s = q^2 , \qquad x = 2p_h \cdot q / q^2 = 2E_h / E_{cm}$$



Fragmentation models

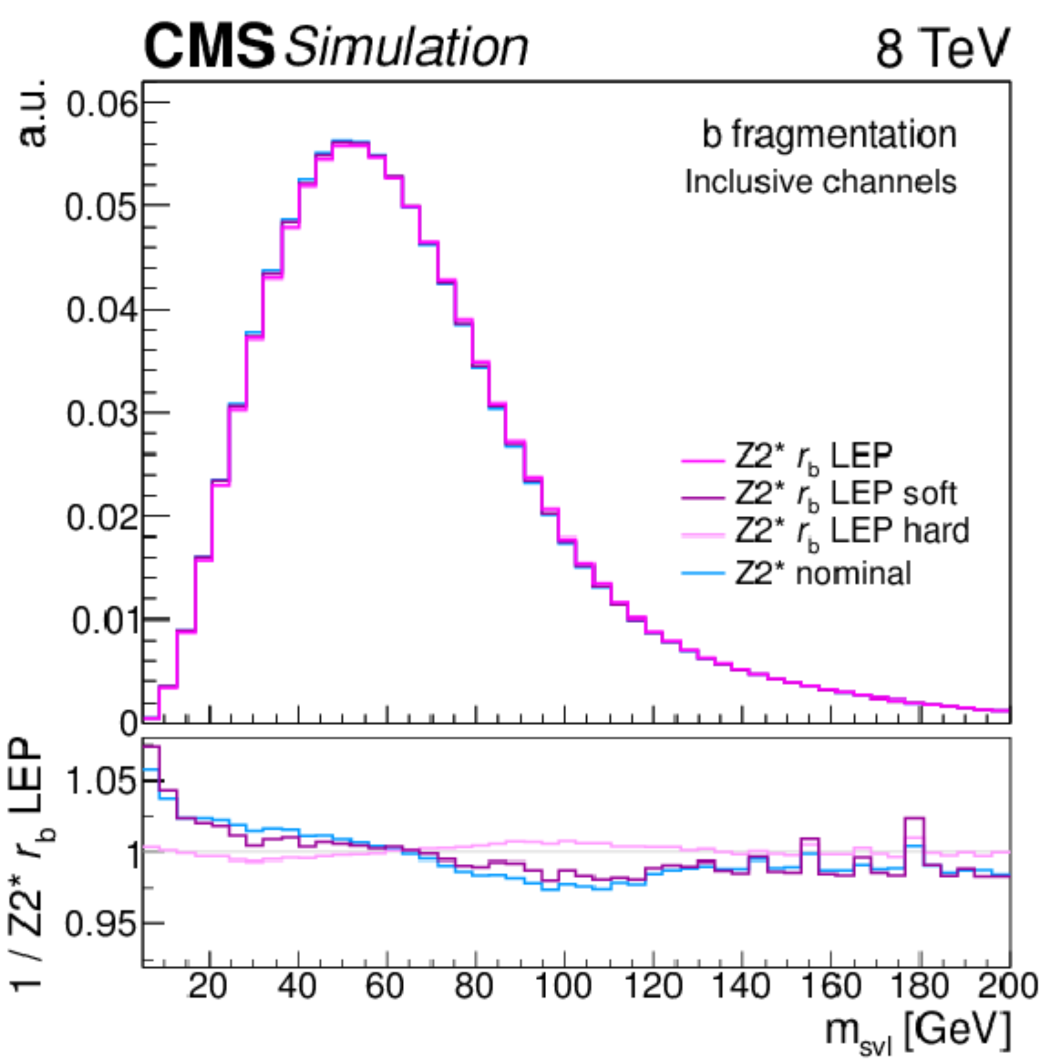
- In Pythia8 the fragmentation of heavy quarks is determined by the Bowler modification of the Lund string function
  - The ***a*** and ***b*** parameters are universal in the model and are constrained by measurements of charged multiplicity, event shapes
  - The ***r_b*** parameter acts as an effective mass for the heavy quark

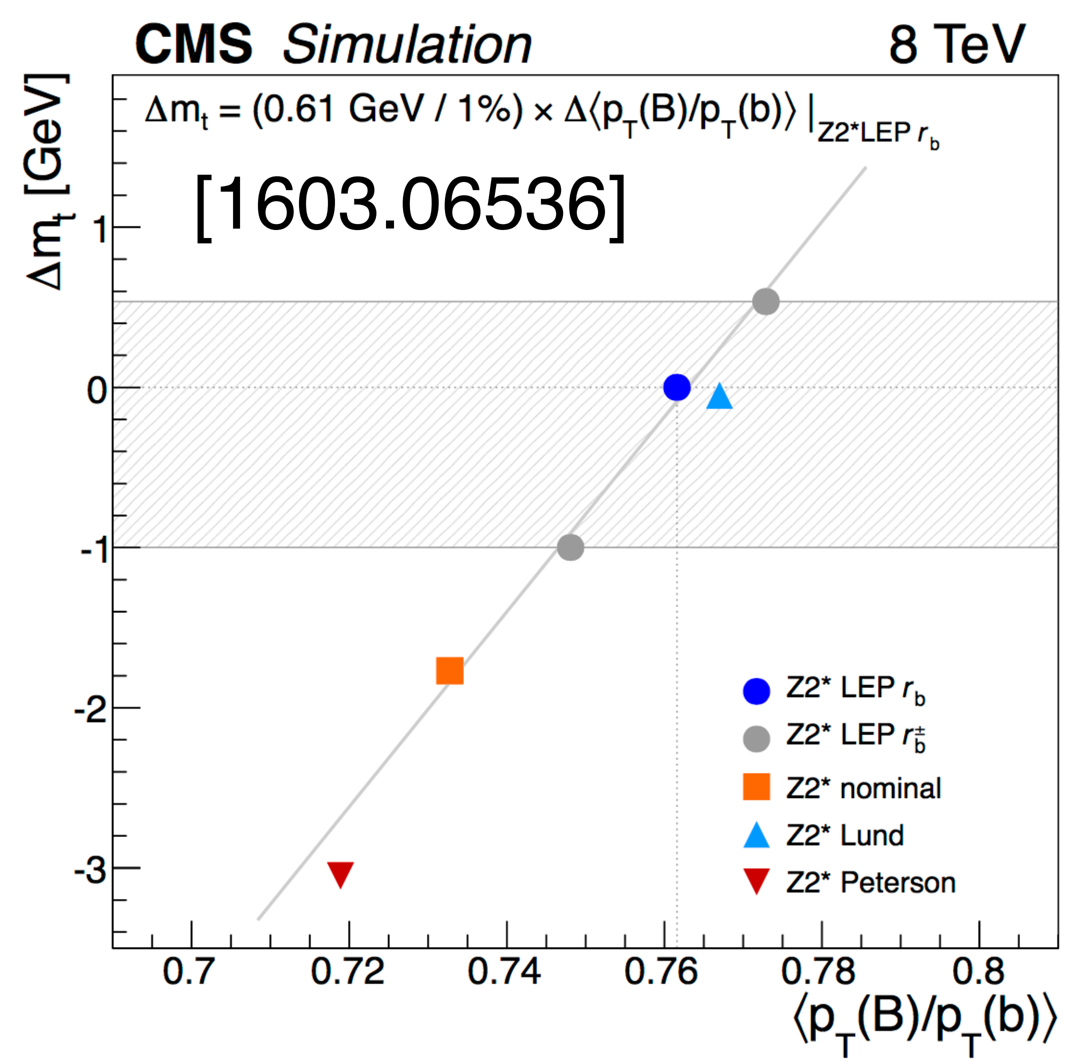
$$f_B(z) \propto \frac{1}{z^{1+br_b m_b^2}}(1-z)^a \exp(-bm_T^2/z)$$

- In Herwig no explicit fragmentation function
- But specific parameters to vary probabilities of heavy-quark cluster masses and splitting probabilities

Impact on Top mass

- Fragmentation uncertainties on the top mass are estimated by fitting r_b in Pythia6 to e^+e^- data $\quad r_b = 0.591^{+0.216}_{-0.274}$
  - Comparisons are also made using alternative functional form for the heavy quark fragmentation (Peterson) or with Herwig, smaller than fit uncertainties, so not used for the final uncertainty
- Top mass uncertainties from these methods 2 - 4 GeV, not yet competitive with traditional methods





Summary

- LHC measurements have reached outstanding precision
  - Sub-percent relative precision feasible for many processes, absolute precision only limited by luminosity
  - Could bring PDFs into an era of precision
  - Precision measurements probing effect of $o(\Lambda_{QCD})$

- But we are already pushing theory beyond its limits
  - And often the limitations come from our understanding of ambiguous non-perturbative effects

- With the HL-LHC program we expect to collect 3 ab^{-1} of data
- Crucial to not only make progress on the perturbative side but also on the non-perturbative part of computations
  - And should probably start to question some of our assumptions (e.g. hadronization universality)