PDF's for Event Generators

Hung-Liang Lai, Joey Huston

in collaboration with

S. Mrenna, D. Stump, W.K. Tung , C.-P. Yuan

Parton distribution functions and global fits

- Calculation of production cross sections at the LHC relies upon knowledge of pdf's in the relevant kinematic region
- Pdf's are determined by global analyses of data from DIS, DY and jet production
- Two* major groups that provide semi-regular updates to parton distributions when new data/theory becomes available
  - MRS->MRST98->MRST99 ->MRST2001->MRST2002 ->MRST2003->MRST2004 ->MSTW2008
  - CTEQ->CTEQ5->CTEQ6 ->CTEQ6.1->CTEQ6.5 ->CTEQ6.6->CT09 (see Pavel's plenary talk)

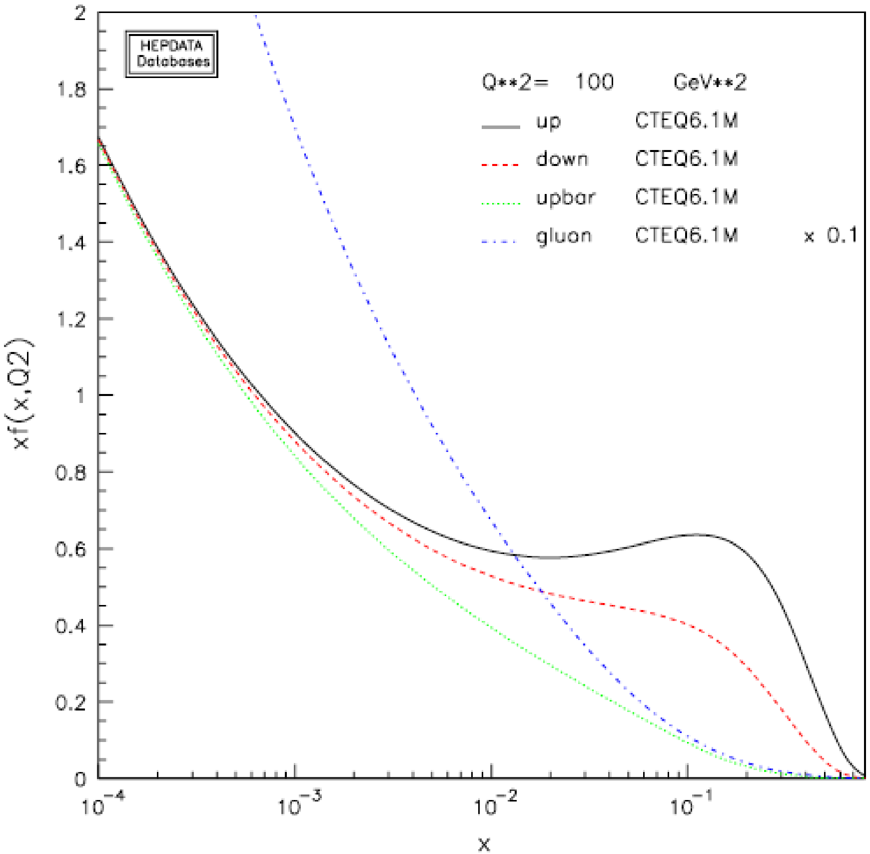


Figure 27. The CTEQ6.1 parton distribution functions evaluated at a Q of 10 GeV.

*now we also have the NNPDF collaboration

PDF's for Event Generators (Pdf4Eg)

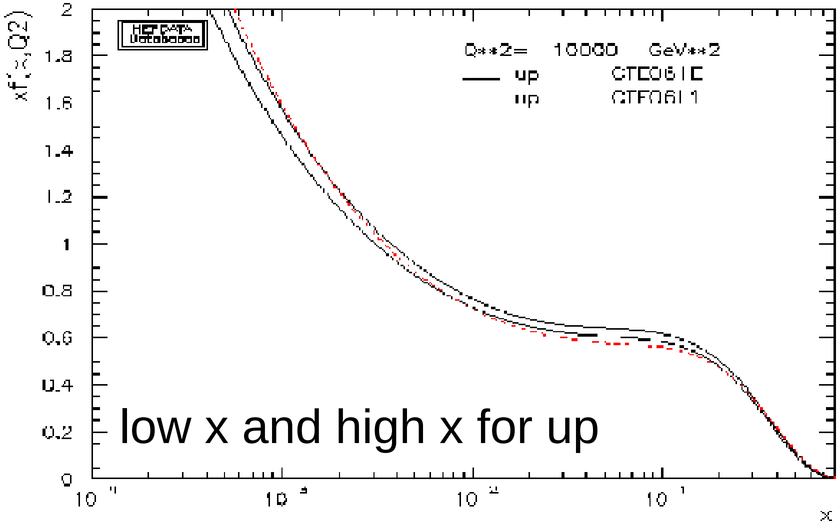
- What about pdf's for parton shower Monte Carlos?
  - standard has been to use LO pdf's, most commonly CTEQ5L/CTEQ6L, in Pythia, Herwig, Sherpa, ALPGEN/Madgraph+…
- …but
  - LO pdf's can create LHC cross sections/acceptances that differ in both shape and normalization from NLO
    - **due to influence of HERA data**
    - **and lack of ln(1/x) and ln(1-x) terms in leading order pdf's and evolution**
  - …and are often outside NLO error bands
  - experimenters use the NLO error pdf's in combination with the central LO pdf even with this mis-match
    - **causes an error in pdf re-weighting due to non-matching of Sudakov form factors**
  - predictions for inclusive observables from LO matrix elements for many of the collider processes that we want to calculate are not so different from those from NLO matrix elements (aside from a reasonably constant K-factor)

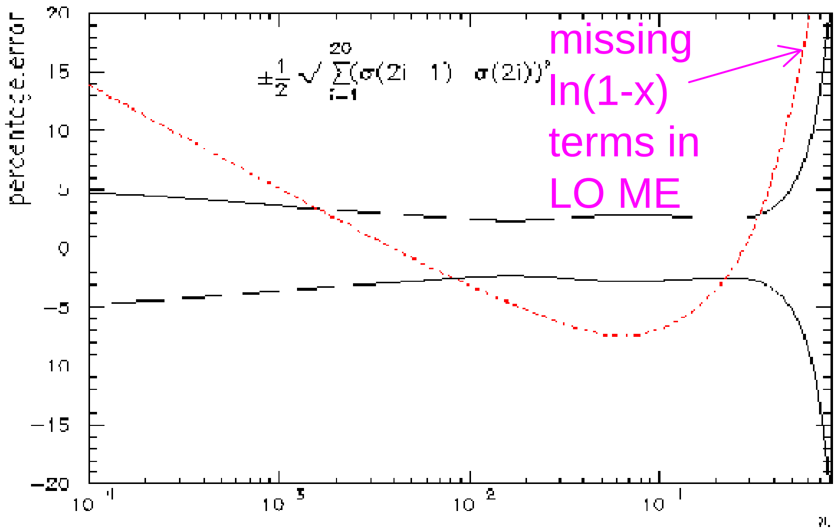
PDF's for Event Generators (Pdf4Eg)

- …but
  - we *like* the low x behavior of LO pdf's and rely upon them for our models of the underlying event at the Tevatron and its extrapolation to the LHC
  - as well as calculating low x cross sections at the LHC
- thus, the need for a special set of PDF's – Pdf4Eg

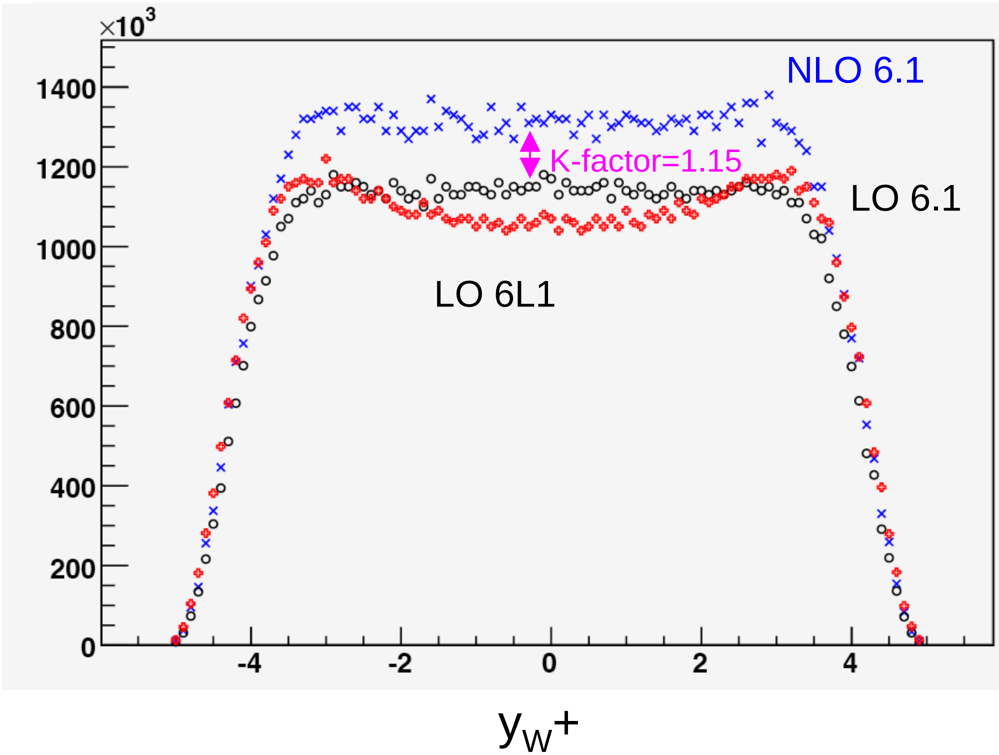
- Pdf4Eg should behave as LO as x->0; as close to NLO as possible as x->1
- Pdf4Eg should describe underlying event at Tevatron with a tune similar to CTEQ6L (for convenience) and extrapolate to a *reasonable* UE at the LHC

Where are the differences between LO and NLO partons?





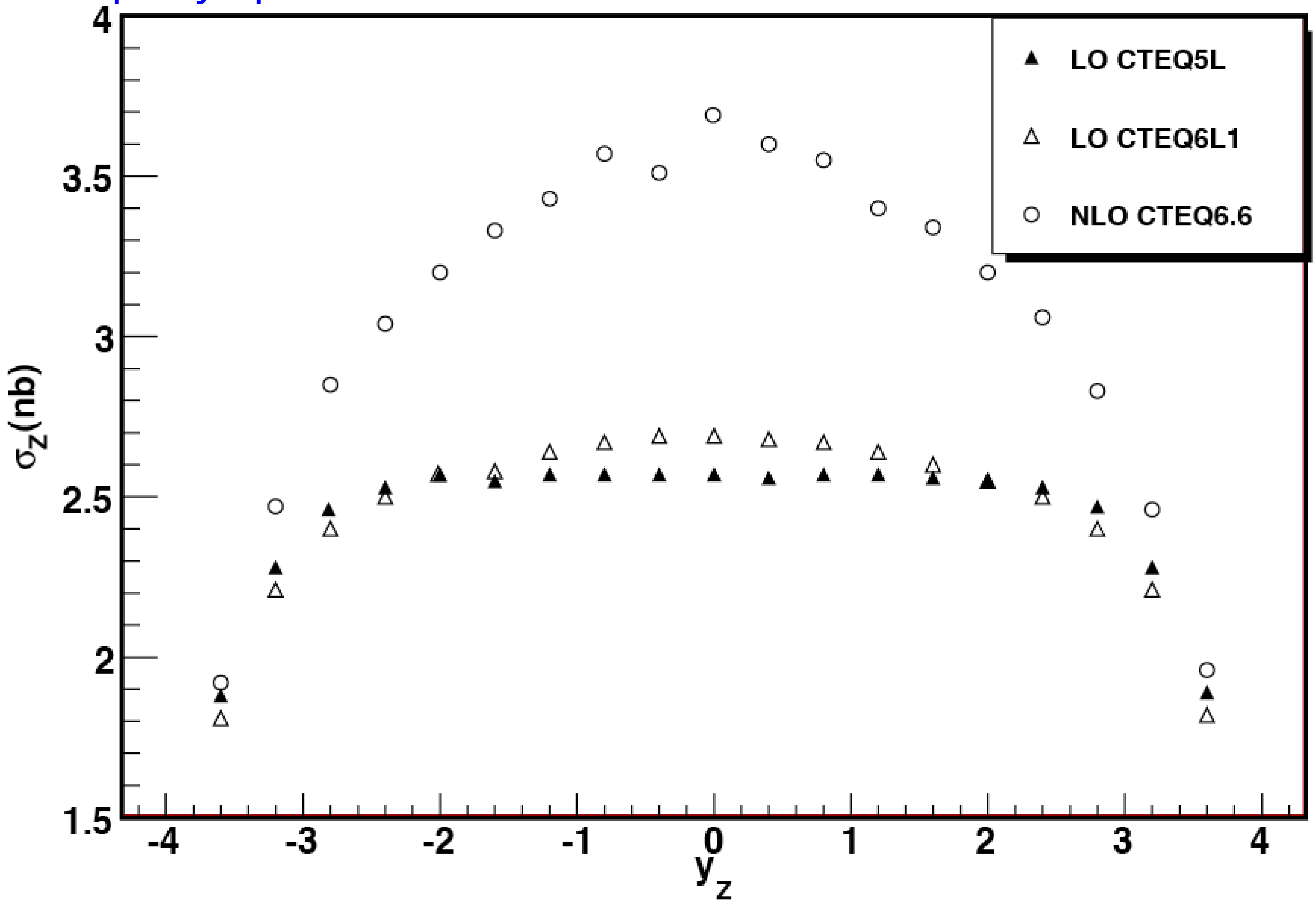
W+ rapidity distribution at LHC



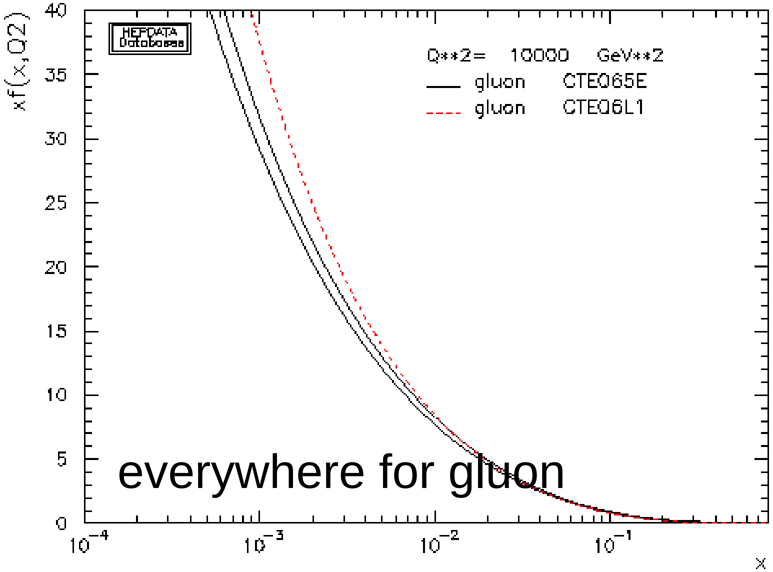
For example, the shape of the W^+ rapidity distribution is significantly different than the NLO result if the LO pdf is used, but very similar if the NLO pdf is used.

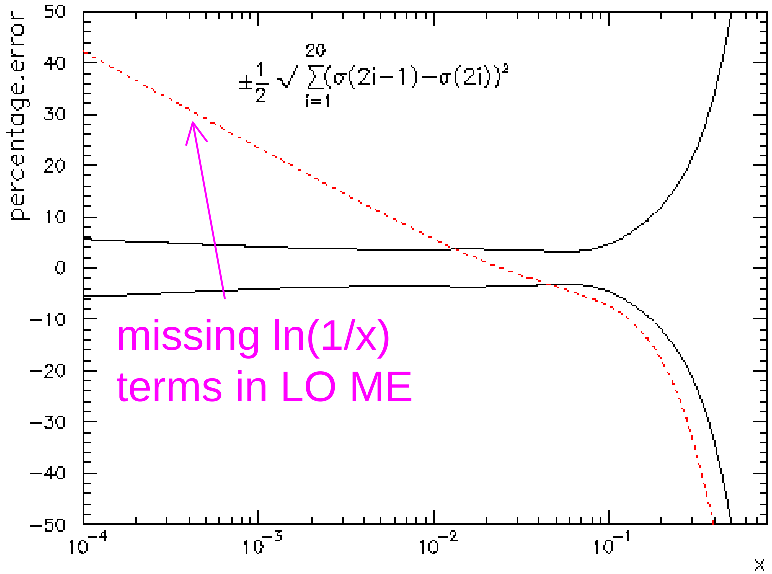
Z production at the LHC

- Consider Z production; CTEQ5L often used and predicts very flat rapidity spectrum

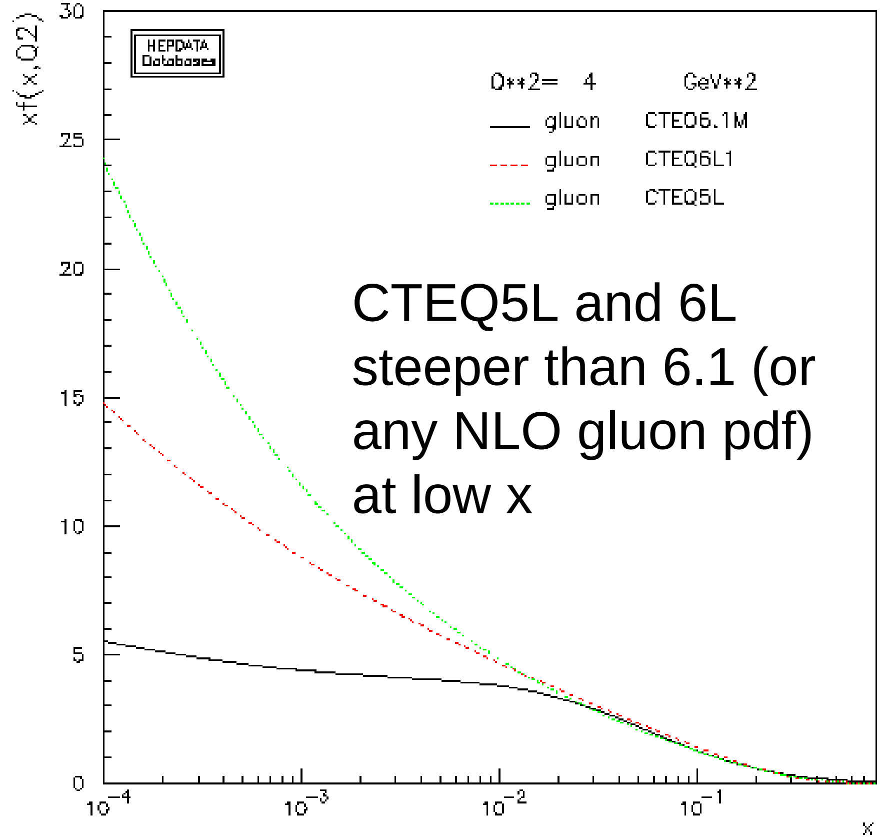


Where are the differences?





- at low Q



K-factors

- K-factors (NLO/LO) can depend on whether NLO or LO pdf's are used for the LO matrix element
- With modified pdf's, it's possible to approach closer to unity for some processes

	Typical scales		Tevatron K-factor			LHC K-factor		
Process	μ_0	μ_1	$\mathcal{K}(\mu_0)$	$\mathcal{K}(\mu_1)$	$\mathcal{K}'(\mu_0)$	$\mathcal{K}(\mu_0)$	$\mathcal{K}(\mu_1)$	$\mathcal{K}'(\mu_0)$
W	m_W	$2m_W$	1.33	1.31	1.21	1.15	1.05	1.15
W+1jet	m_W	p_T^{jet}	1.42	1.20	1.43	1.21	1.32	1.42
W+2jets	m_W	p_T^{jet}	1.16	0.91	1.29	0.89	0.88	1.10
WW+jet	m_W	$2m_W$	1.19	1.37	1.26	1.33	1.40	1.42
$t\bar{t}$	m_t	$2m_t$	1.08	1.31	1.24	1.40	1.59	1.48
$t\bar{t}$+1jet	m_t	$2m_t$	1.13	1.43	1.37	0.97	1.29	1.10
$b\bar{b}$	m_b	$2m_b$	1.20	1.21	2.10	0.98	0.84	2.51
Higgs	m_H	p_T^{jet}	2.33	–	2.33	1.72	–	2.32
Higgs via VBF	m_H	p_T^{jet}	1.07	0.97	1.07	1.23	1.34	1.09
Higgs+1jet	m_H	p_T^{jet}	2.02	–	2.13	1.47	–	1.90
Higgs+2jets	m_H	p_T^{jet}	–	–	–	1.15	–	–

Table 2: K-factors for various processes at the Tevatron and the LHC calculated using a selection of input parameters. In all cases, the CTEQ6M PDF set is used at NLO. $\mathcal{K}$ uses the CTEQ6L1 set at leading order, whilst $\mathcal{K}'$ uses the same set, CTEQ6M, as at NLO. For most of the processes listed, jets satisfy the requirements $p_T > 15$ GeV/c and $|\eta| < 2.5$ (5.0) at the Tevatron (LHC). For Higgs+1,2jets, a jet cut of 40 GeV/c and $|\eta| < 4.5$ has been applied. A cut of $p_T^{\text{jet}} > 20\ GeV/c$ has been applied for the $t\bar{t}$+jet process, and a cut of $p_T^{\text{jet}} > 50\ GeV/c$ for WW+jet. In the W(Higgs)+2jets process the jets are separated by $\Delta R > 0.52$, whilst the VBF calculations are performed for a Higgs boson of mass 120 GeV. In each case the value of the K-factor is compared at two often-used scale choices, where the scale indicated is used for both renormalization and factorization scales.

CTEQ techniques

- Include in Pdf4Eg fit (weighted) pseudo-data for characteristic LHC processes produced using CTEQ6.6 NLO pdf's with NLO matrix elements (using MCFM), along with full CTEQ6.6 dataset (2885 points)
  - low mass bB
    - fix low x gluon for UE
  - tT over full mass range
    - higher x gluon
  - W^+,W^-,Z^0 rapidity distributions
    - quark distributions
  - gg->H (120 GeV) rapidity distribution

Choices

- Use of 2-loop or 1-loop α_s
  - Herwig preference for 2-loop
  - Pythia preference for 1-loop
  - will provide both options, but will fix the value of α_s at world averages:0.118 (NLO),0.130 (LO)
- Fixed momentum sum rule, or not
  - re-arrange momentum within proton and/or add extra momentum
  - extra momentum appreciated by some of pseudo-data sets but not others and may lose some useful correlations
- Fix pseudo-data normalizations to K-factors expected from higher order corrections, or let float
- Scale variation within reasonable range for fine-tuning of agreement with pseudo-data
  - for example, let vector boson scale vary from 0.5 m_B to 2.0 m_B

3 pdf's: all using pseudo-data

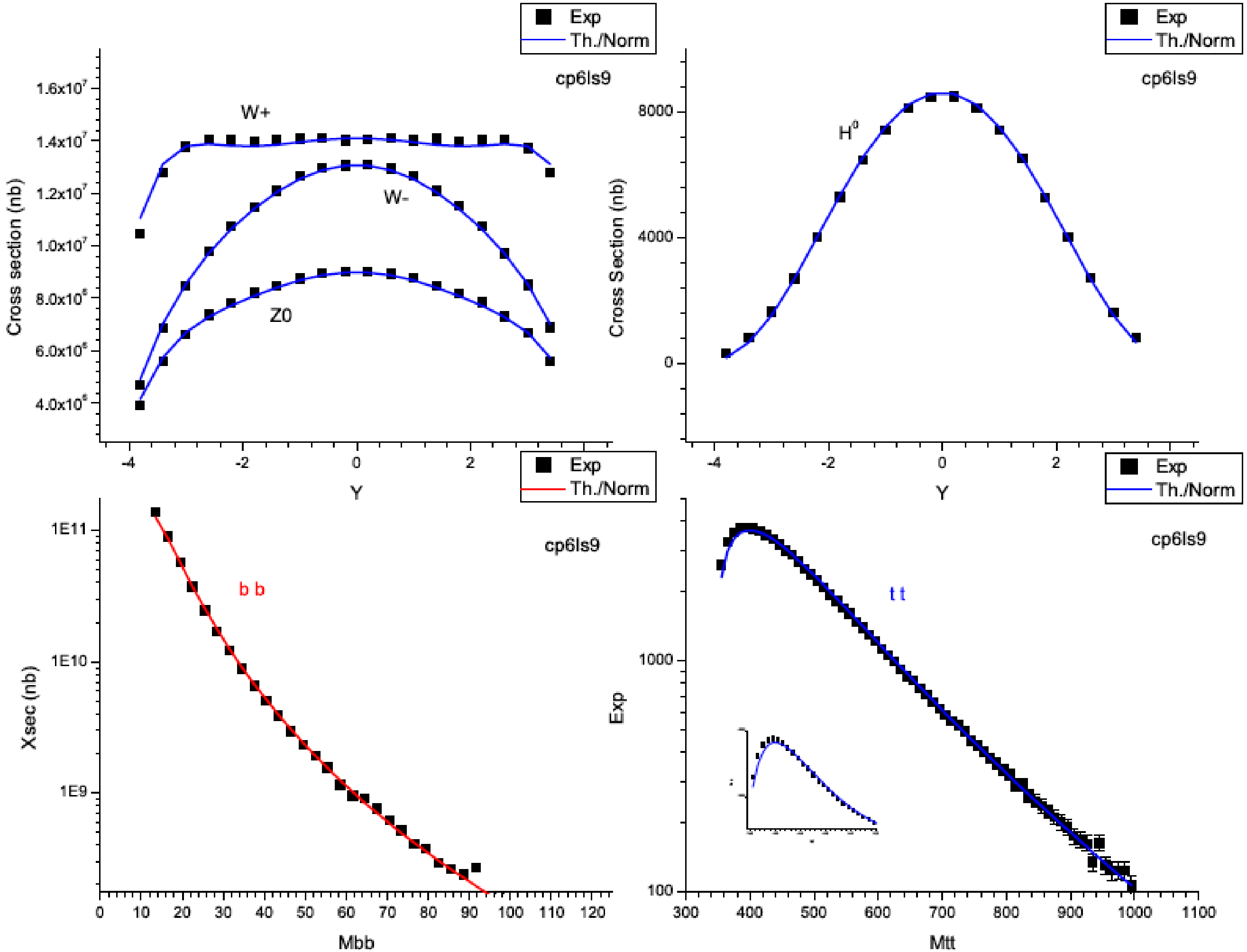
1. Keep momentum sum rule, vary scales used in matrix elements for pseudo-data (2-loop α_s)
2. Relax momentum sum rule, 2-loop α_s
3. Relax momentum sum rule, 1-loop α_s

Some observations

- Pseudo-data has conflicts with global data set
  - that's the motivation of the generating Pdf4Eg
- Requiring better fit to pseudo-data increases chisquare of LO fit to global data set by ~15% (although this is not the primary concern, the fit to the pseudo-data is; the amount of χ^2 increase depends on how much emphasis given to pseudo-data)
  - α_s
    - no strong preference for 1-loop or 2-loop α_s (with pseudo-data in the fit)
  - χ^2 improves with momentum sum rule free
    - prefers more momentum; mostly goes into the gluon distribution (other sum rules kept intact)
    - normalization of pseudo-data (needed K-factor) gets closer to 1 (since the χ^2 gets better if that happens)
    - still some conflicts with DIS data that don't prefer more momentum
    - …but we're not making these pdf's for DIS comparisons

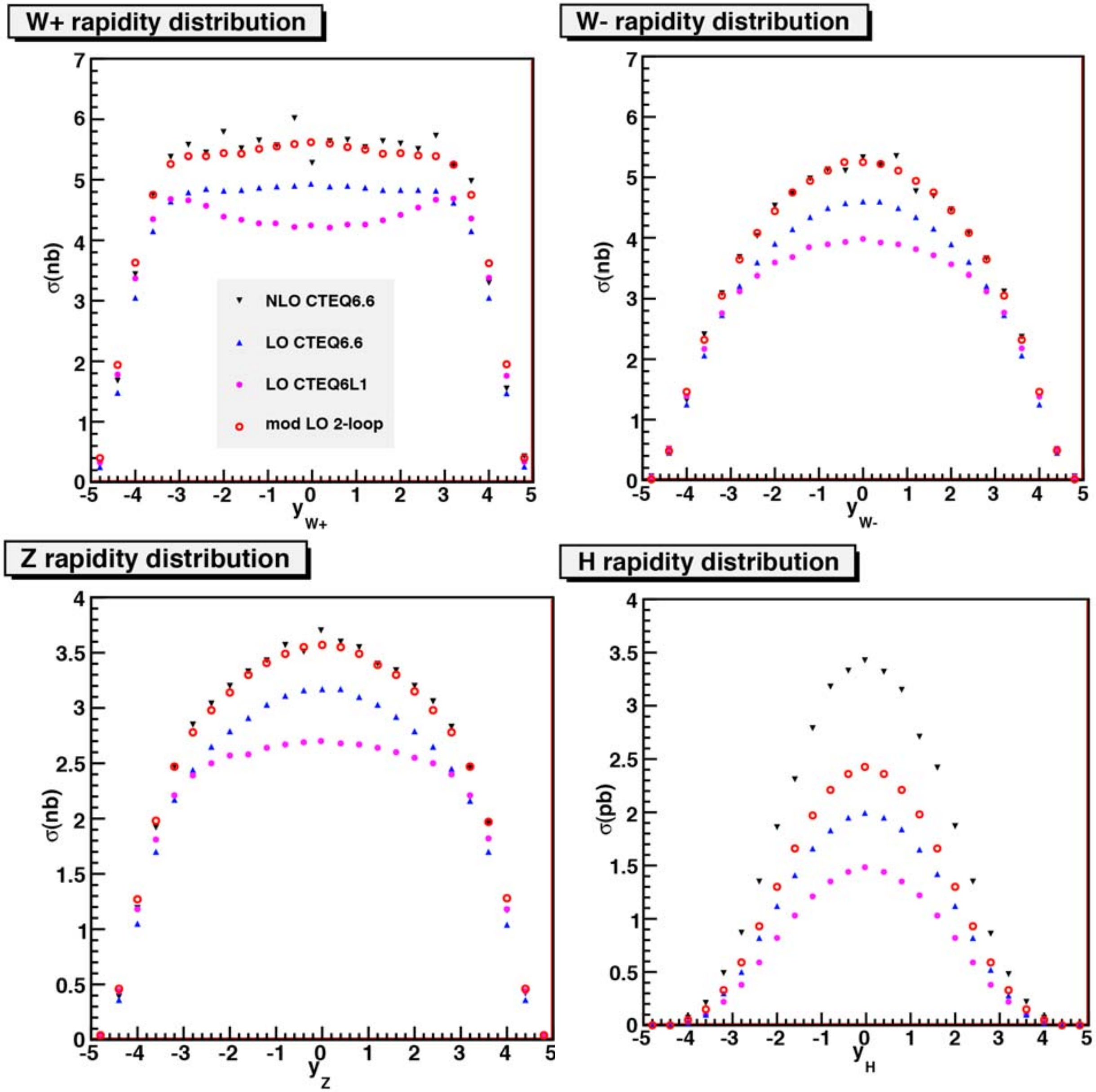
Fixed momentum sum rule, vary scale

2-loop α_s; fixed momentum sum; variable scales
good agreement with shapes; norms for H, tT lower



	N_i	scale
W^+	0.9	$1.96m_W$
W^-	0.92	$1.96m_W$
Z	0.92	$1.96m_Z$
H	0.53	$1.06m_H$
tT	0.48	$1.41m_t$

Allow momentum sum rule to float



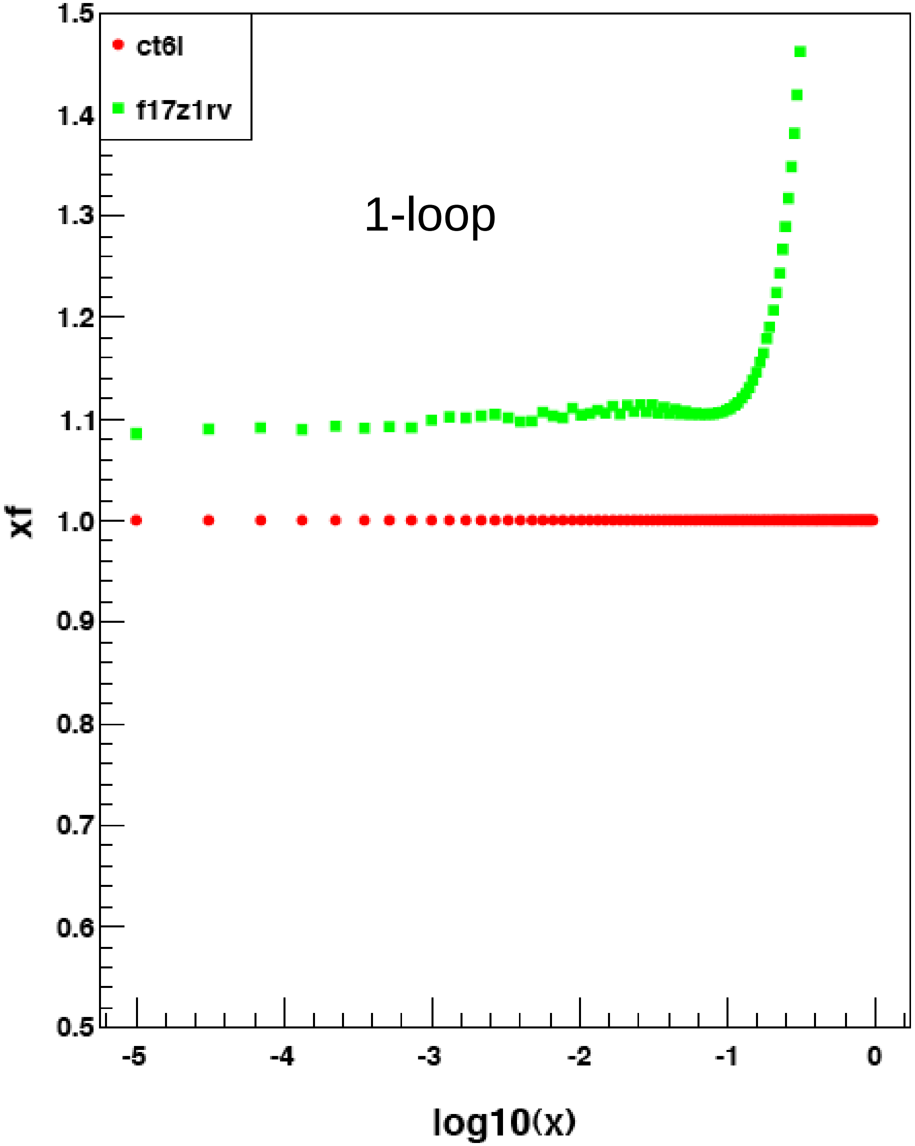
Pdf's with 1-loop or 2-loop α_s

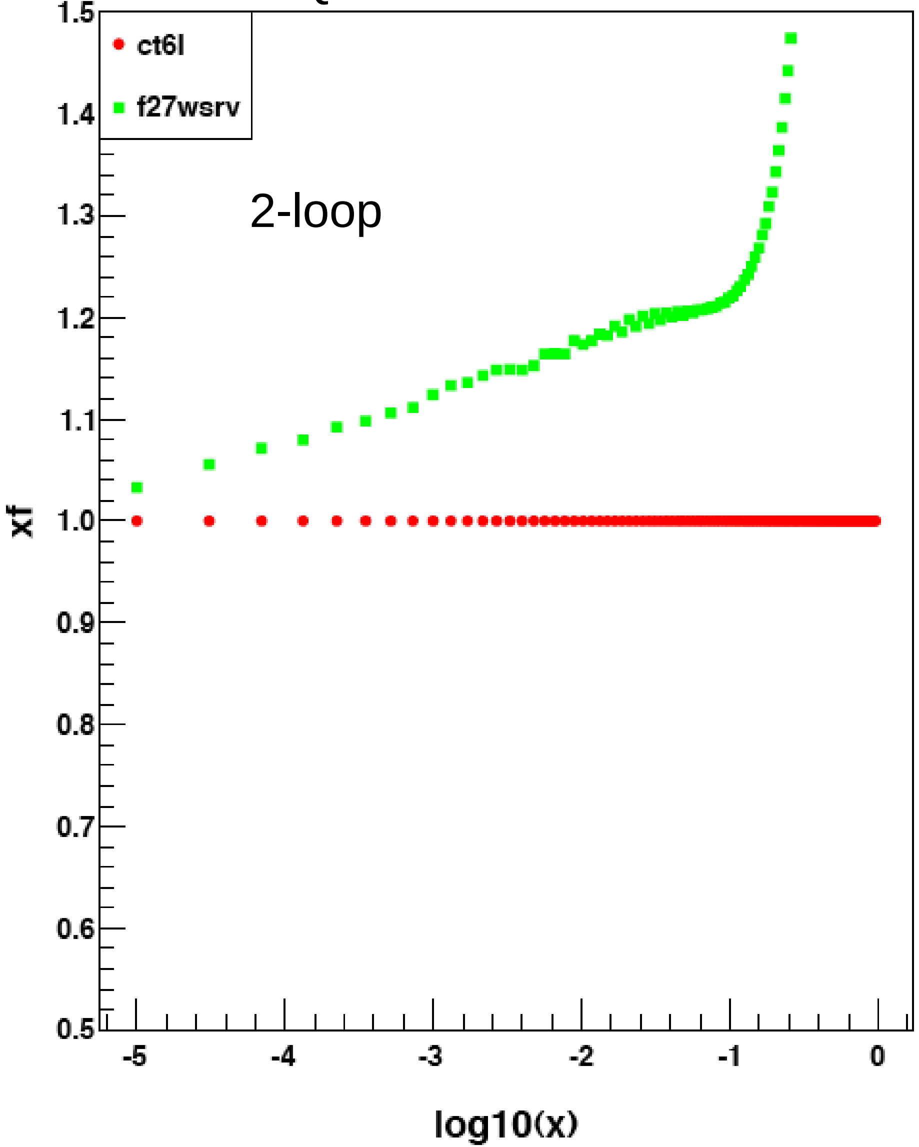
	1-loop	2-loop
mom sum	1.10	1.14
α_s	0.130	0.118
W^+ (norm)	1.00	0.98
W^-(norm)	1.01	1.00
Z (norm)	1.01	1.00
H(120)	0.82	0.76
tT (norm)	0.84	0.83

Plots look similar for 1-loop α_s

Behavior of gluon distributions

Take ratio with LO gluon from CTEQ6L





Summary

- Conventional ways of generating events with LO parton shower Monte Carlos have drawbacks from the point of view of parton distribution functions
- Here we have outlined two techniques for producing modified pdf's for use with LHC parton showering Monte Carlos
  - both provide better shapes for some key physics distributions
  - the second provides K-factors closer to unity
- Continuing LHC phenomenology with these pdf's
- This work was made possible by the insight and inspiration of our late colleague Wu-Ki Tung