Parton Distributions Functions and Correlated Uncertianties

Robert Thorne

November 18th, 2011

University College London

Thanks to Alan Martin, James Stirling and Graeme Watt

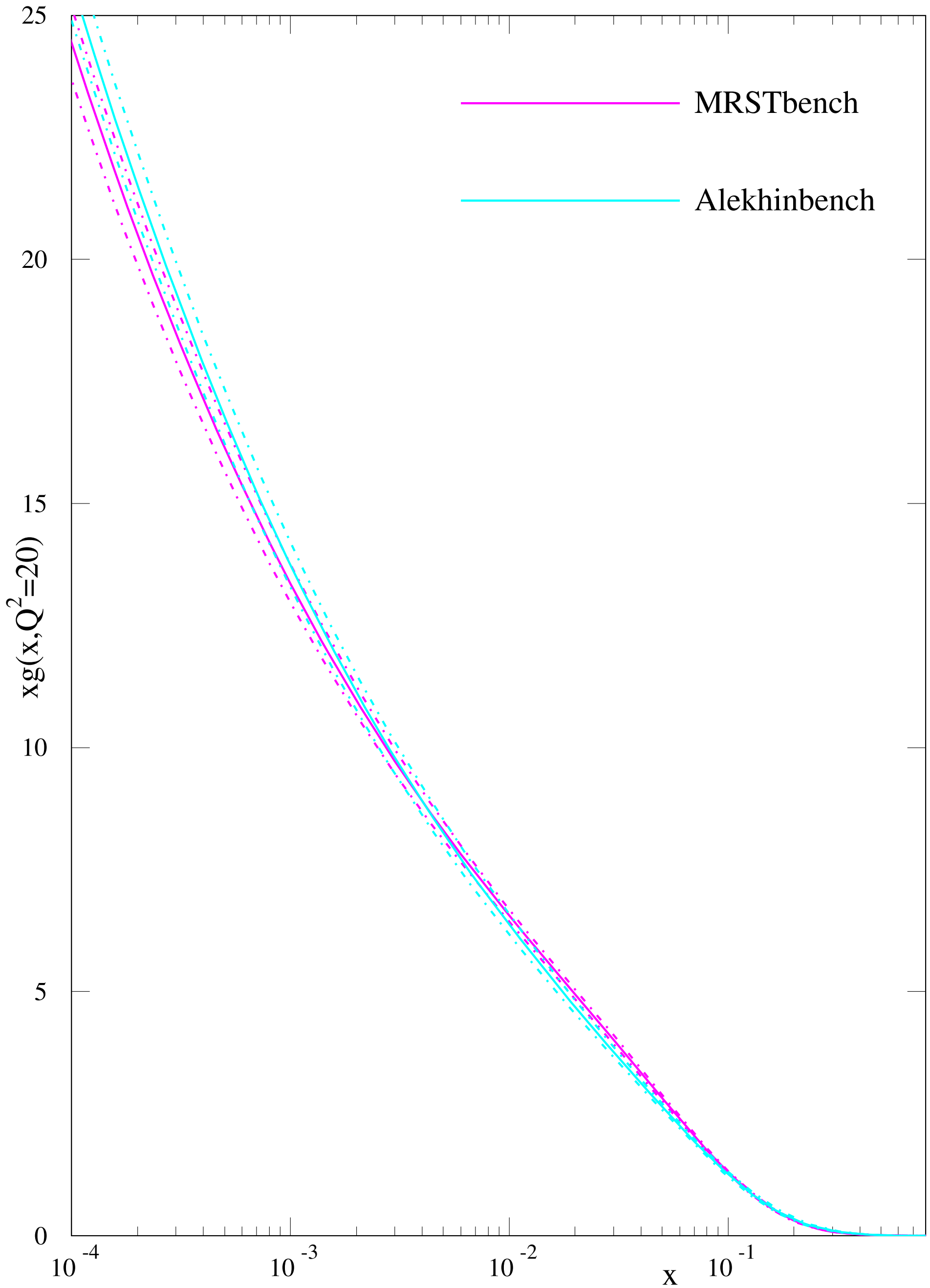
Do not yet include full correlations for DIS data. Add all errors in quadrature. Some observations.

First HERA-LHC benchmark study of PDFs, treatment without correlated errors gave $\Delta\alpha_S(M_Z^2)$ slightly lower and rather similar gluons (and other PDFs).

In MRST2001 fit use of correlated errors in HERA data led to $\Delta\alpha_S(M_Z^2) = -0.0003$ effect, and similarly small for partons. Most of HERA data in MSTW2008 fit.

Use of combined HERA data with much reduced effect of correlated errors pushes $\alpha_S(M_Z^2)$ up less than 0.001. HERAPDF and e.g. CT10 choose different default methods of dealing with correlations.

No reason to think correlated uncertainties have Gaussian distribution.



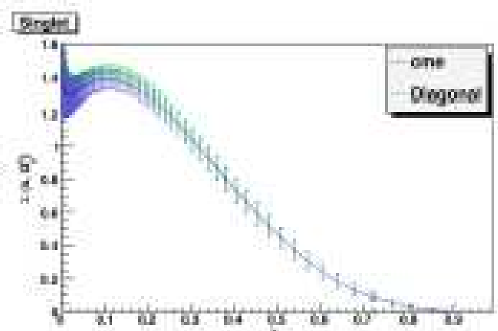
THE IMPACT OF CORRELATED UNCERTAINTIES

REPEAT THE FIT NEGLECTING ALL CORRELATIONS (A.Donati)

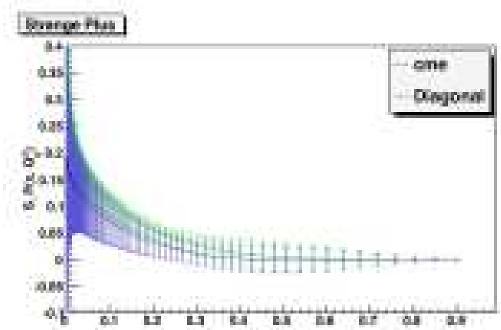
Experiment	Set	CME fit χ^2_{diag}	CME fit χ^2_{corr}	Diagonal fit χ^2_{diag}	Diagonal fit χ^2_{corr}
TOT (all exp)		0.988	1.323	0.844	1.321
NMC-pd		1.965	1.457	1.167	1.155
NMC		1.006	1.659	1.078	1.76
SLAC		0.836	1.183	1.008	1.406
	SLACp	1.018	1.307	1.132	1.525
	SLACd	0.651	0.912	0.882	1.275
BCDMS		0.777	1.646	0.552	1.604
	BCDMSp	0.873	1.808	0.617	1.703
	BCDMSd	0.648	1.296	0.465	1.23
ZEUS		0.770	1.053	0.742	1.048
	Z97lowQ2	0.474	1.294	0.434	1.367
	Z97NC	0.718	1.125	0.669	1.106
	Z97CC	0.912	0.800	1.021	0.894
	Z02NC	0.798	0.767	0.763	0.733
	Z02CC	0.619	0.592	0.593	0.569
	Z03NC	0.975	1.104	0.907	1.012
	Z03CC	1.131	1.001	1.259	1.115
H1		1.020	1.053	0.987	1.028
	H197mb	0.861	1.298	0.877	1.33
	H197lowQ2	0.666	0.948	0.774	0.97
	H197NC	1.071	0.903	0.986	0.852
	H197CC	0.758	0.764	0.831	0.824
	H199NC	1.229	1.109	1.171	1.068
	H199CC	0.621	0.646	0.644	0.668
	H199NChy	0.333	0.361	0.326	0.353
	H100NC	1.208	1.172	1.120	1.102
	H100CC	1.122	1.013	1.311	1.146
CHORUS		1.018	1.380	0.745	1.392
	CHORUSnu	1.082	1.449	0.628	1.403
	CHORUSnb	0.954	1.178	0.861	1.254
FLH108		0.984	1.729	0.946	1.7
NTVDMN		0.869	0.692	1.094	0.984
	NTVnuDMN	1.061	0.763	0.445	0.421
	NTVnbDMN	0.667	0.660	1.774	1.618
ZEUS-H2		1.392	1.509	1.373	1.512
	Z06NC	1.691	1.495	1.667	1.472
	Z06CC	0.664	1.230	0.659	1.252

- DIAGONAL χ^2 OF DIAGONAL FIT MUCH LOWER, CORREL. χ^2 OF TWO FITS UNCHANGED
- DIAGONAL FIT REWEIGHTS EXPERIMENTS $\Rightarrow$ EXPTS WITH LARGER SYST. (FIXED TARGET) GET SMALLER WEIGHT
- VALENCE & STRANGE PDFS AFFECTED AT THE $\frac{1}{4}\sigma$ LEVEL

SINGLET



STRANGE



Normalisation Uncertainties

The most significant correlated uncertainty for DIS data.

In old sets the normalization of each data set was determined by the best fit – and then fixed.

Now implement procedure of allowing normalisations of all sets to vary in best fit and scan over eigenvectors, with penalty term for each set

$$\chi^2_{\mathcal{N}} = \left(\frac{1-\mathcal{N}}{\sigma^{\mathcal{N}}}\right)^4$$

Quartic penalty avoids very large deviations. Still shift down at LO (fit failure) and slightly at NLO.

Rescale errors with normalization to avoid bias (D'Agostini).

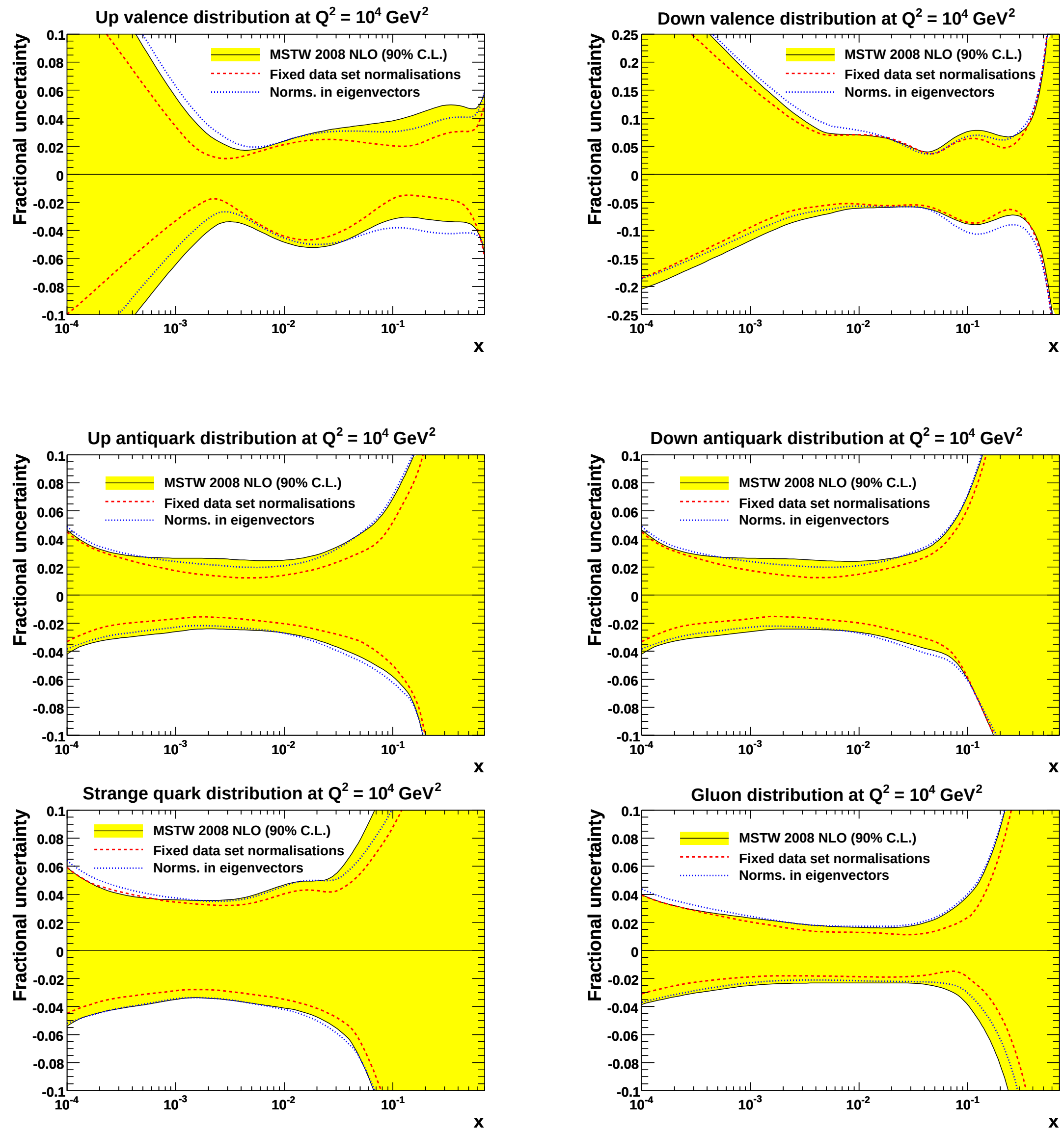
Data set	$\sigma_n^{\mathcal{N}}$	LO	NLO	NNLO
BCDMS μp F_2 [32]	3%	0.9667	0.9644	0.9678
BCDMS μd F_2 [102]	3%	0.9667	0.9644	0.9678
NMC μp F_2 [33]	2%	1.0083	0.9982	0.9999
NMC μd F_2 [33]	2%	1.0083	0.9982	0.9999
NMC $\mu n/\mu p$ [103]	—	1	1	1
E665 μp F_2 [104]	1.85%	1.0146	1.0052	1.0024
E665 μd F_2 [104]	1.85%	1.0146	1.0052	1.0024
SLAC ep F_2 [105, 106]	1.9%	1.0227	1.0125	1.0078
SLAC ed F_2 [105, 106]	1.9%	1.0227	1.0125	1.0078
NMC/BCDMS/SLAC F_L [32–34]	—	1	1	1
E866/NuSea pp DY [107]	6.5%	1.0629	1.0086	1.0868
E866/NuSea pd/pp DY [108]	—	1	1	1
NuTeV νN F_2 [37]	2.1%	0.9987	0.9997	0.9992
CHORUS νN F_2 [38]	2.1%	0.9987	0.9997	0.9992
NuTeV νN xF_3 [37]	2.1%	0.9987	0.9997	0.9992
CHORUS νN xF_3 [38]	2.1%	0.9987	0.9997	0.9992
CCFR $\nu N \to \mu\mu X$ [39]	2.1%	0.9987	0.9997	0.9992
NuTeV $\nu N \to \mu\mu X$ [39]	2.1%	0.9987	0.9997	0.9992
H1 MB 99 e^+p NC [31]	1.3%	0.9861	1.0098	1.0090
H1 MB 97 e^+p NC [109]	1.5%	0.9863	0.9921	0.9953
H1 low Q^2 96–97 e^+p NC [109]	1.7%	1.0029	1.0095	1.0172
H1 high Q^2 98–99 e^-p NC [110]	1.8%	0.9782	0.9851	0.9860
H1 high Q^2 99–00 e^+p NC [35]	1.5%	0.9762	0.9834	0.9842
ZEUS SVX 95 e^+p NC [111]	1.5%	0.9944	0.9948	1.0004
ZEUS 96–97 e^+p NC [112]	2%	0.9735	0.9811	0.9871
ZEUS 98–99 e^-p NC [113]	1.8%	0.9771	0.9855	0.9862
ZEUS 99–00 e^+p NC [114]	2.5%	0.9656	0.9761	0.9762
H1 99–00 e^+p CC [35]	1.5%	0.9762	0.9834	0.9842
ZEUS 99–00 e^+p CC [36]	2.5%	0.9656	0.9761	0.9762
H1/ZEUS ep $F_2^{\rm charm}$ [41–47]	—	1	1	1
H1 99–00 e^+p incl. jets [59]	1.5%	0.9762	0.9834	—
ZEUS 96–97 e^+p incl. jets [57]	2%	0.9735	0.9811	—
ZEUS 98–00 $e^\pm p$ incl. jets [58]	2.5%	0.9656	0.9761	—
DØ II $p\bar{p}$ incl. jets [56]	6.1%	0.9353	1.0596	1.0759
CDF II $p\bar{p}$ incl. jets [54]	5.8%	0.8779	0.9646	0.9900
CDF II $W \to \ell\nu$ asym. [48]	—	1	1	1
DØ II $W \to \ell\nu$ asym. [49]	—	1	1	1
DØ II Z rap. [53]	—	1	1	1
CDF II Z rap. [52]	5.8%	0.8779	0.9646	0.9900

Comparison of full uncertainty and that from no normalization uncertainties (except in best fit).

Contribution $\sim 1\!-\!1.5\%$, (for 90% C.L.) for all partons.

Varies for different eigenvectors – some very sensitive (size of quarks) others insensitive ($\bar{u} - \bar{d}$ determined from ratios).

Normalisation uncertainties increase uncertainties on partons significantly, but not dramatically.



For sets with larger, or dominant correlated uncertainties (all Tevatron data) we use

$$\chi^2_n(\{a\},\mathcal{N}_n) \;=\; \sum_{i=1}^{N_{\rm pts.}} \left(\frac{\hat{D}_{n,i} - T_{n,i}(\{a\})/\mathcal{N}_n}{\sigma^{\rm uncorr.}_{n,i}}\right)^2 \;+\; \sum_{k=1}^{N_{\rm corr.}} r^2_{n,k} \;+\; \chi^2_{\mathcal{N}_n},$$

where $\hat{D}_{n,i} \equiv D_{n,i} - \sum_{k=1}^{N_{\rm corr.}} r_{n,k}\,\sigma^{\rm corr.}_{n,k,i}$ are the data points allowed to shift by the systematic errors in order to give the best fit. Minimising χ^2_n with respect to $r_{n,k}$ gives the analytic result that

$$r_{n,k}(\{a\},\mathcal{N}_n) = \sum_{k'=1}^{N_{\rm corr.}} (A^{-1})_{kk'} B_{k'}(\{a\},\mathcal{N}_n),$$

where

$$A_{kk'} = \delta_{kk'} + \sum_{i=1}^{N_{\rm pts.}} \frac{\sigma^{\rm corr.}_{n,k,i}\,\sigma^{\rm corr.}_{n,k',i}}{\sigma^{\rm uncorr.}_{n,i}}, \quad B_k(\{a\},\mathcal{N}_n) = \sum_{i=1}^{N_{\rm pts.}} \frac{\sigma^{\rm corr.}_{n,k,i}\,(D_{n,i} - T_{n,i}(\{a\})/\mathcal{N}_n)}{(\sigma^{\rm uncorr.}_{n,i})^2}.$$

Therefore, the optimal shifts of the data points by the systematic errors are solved for analytically.

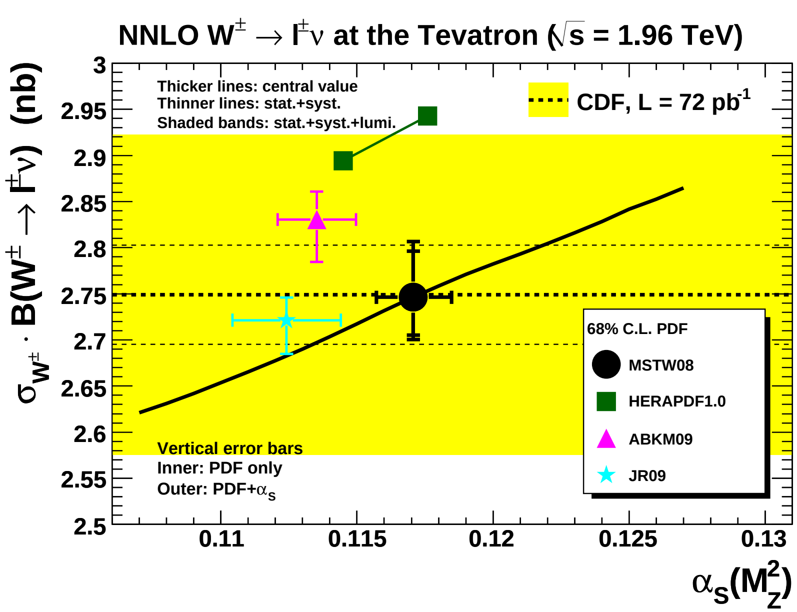
Important Implication

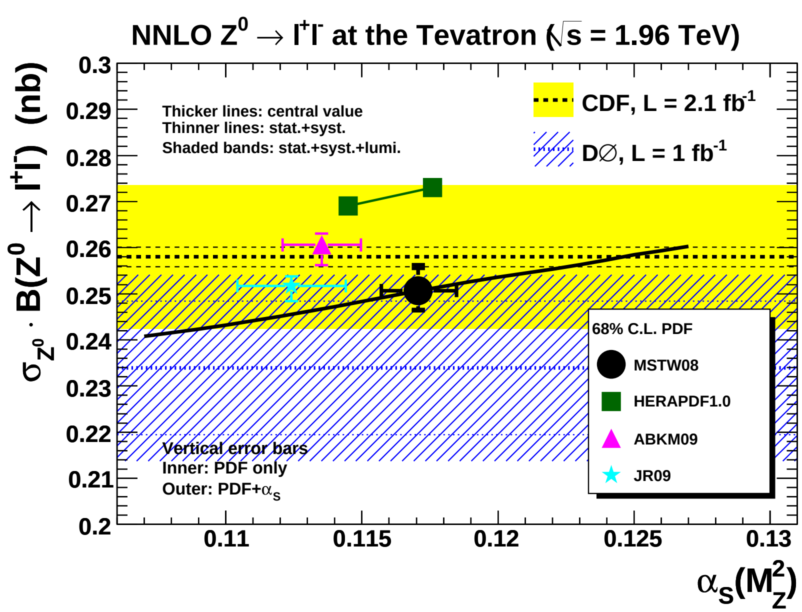
High-x gluon, at least to some extent, constrained by Tevatron jet data.

However, CDF Z-rapidity data, or total cross sections, sets Tevatron normalisation in a fit.

Only allows a few percent variation in normalisation.

Different PDF predictions for W and Z cross sections at the Tevatron compared to data.



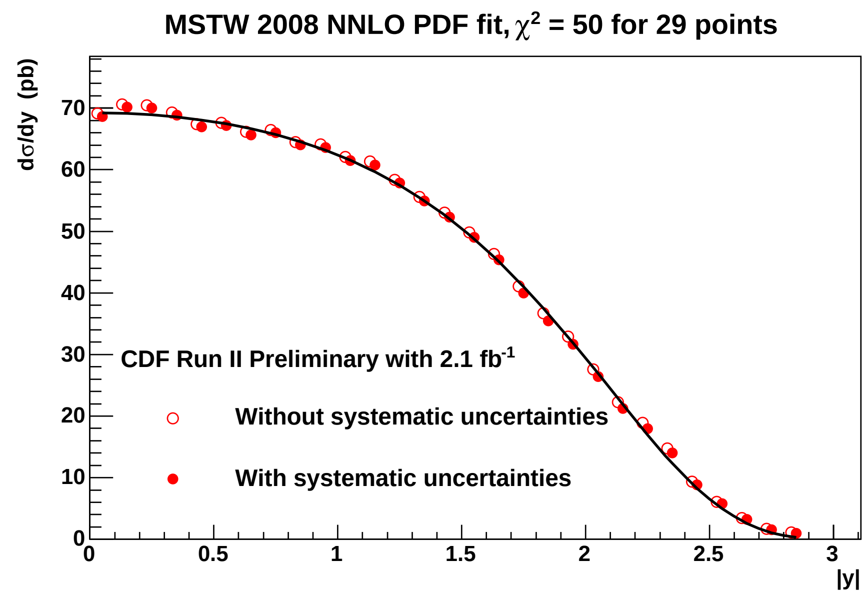


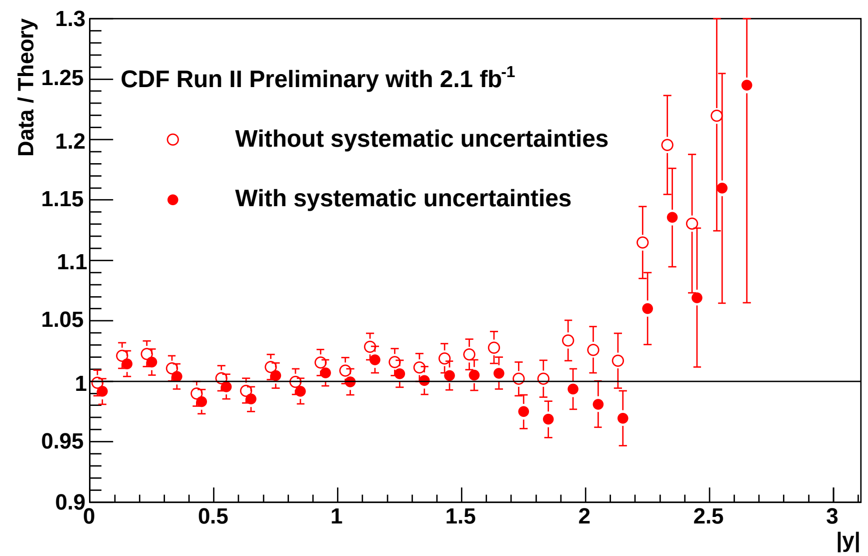
Everyone ok or a bit high. Normalisation no room to move down.

Fit to CDF Z-rapidity data within MSTW framework actually constrains normalisations to a couple of percent.

Movement due to other sources of correlated uncertainty in best fit. None much more than one standard deviation.

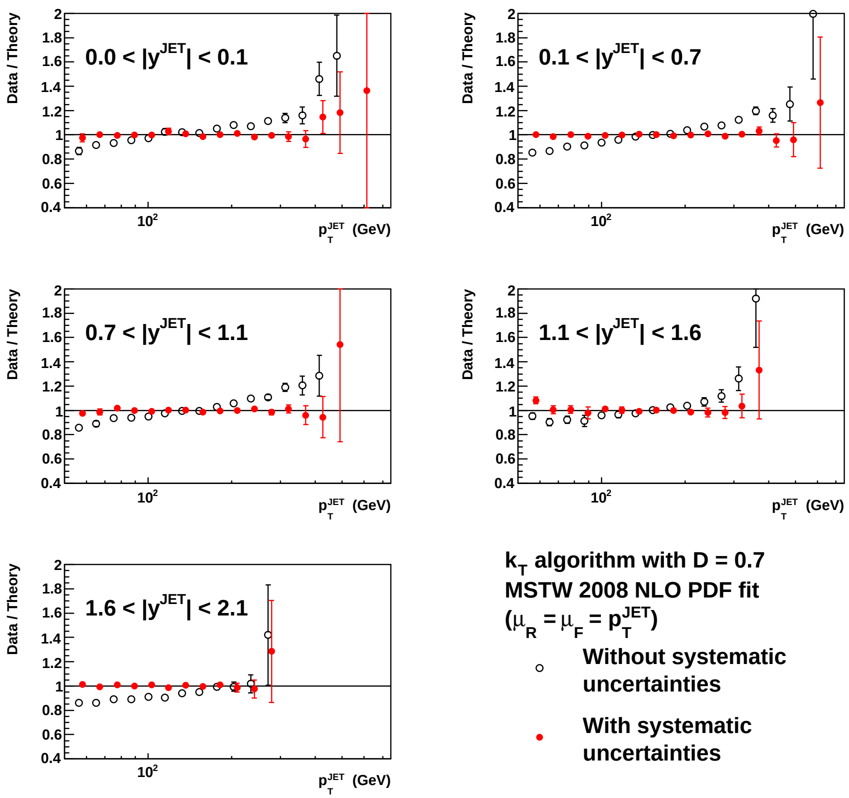
Z/γ^* rapidity distribution from CDF



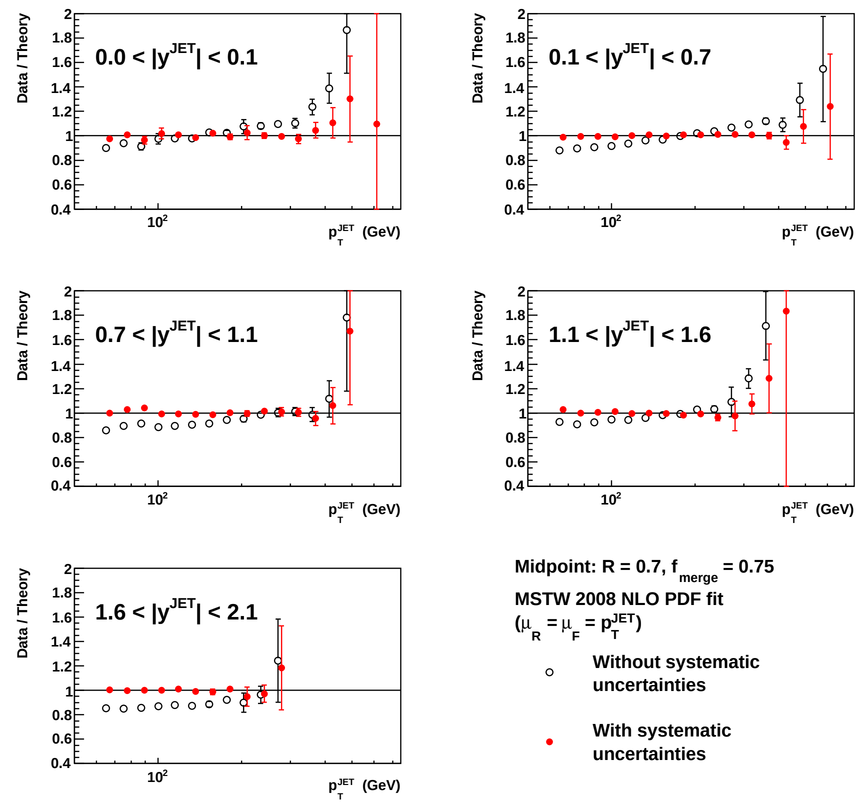


The comparison to CDF inclusive jet data using the k_T and cone algorithms.

CDF Run II inclusive jet data, χ^2 = 56 for 76 pts.



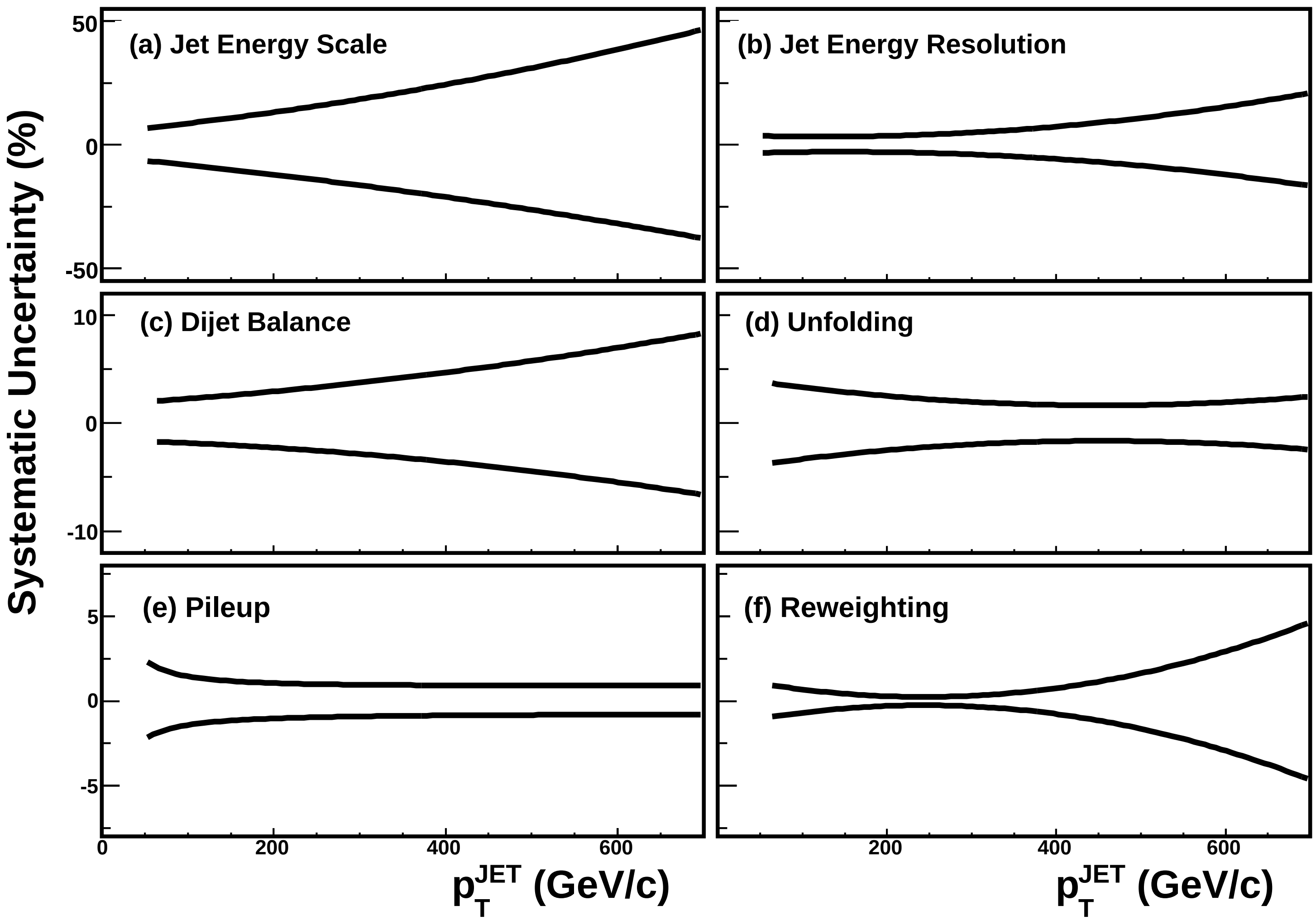
CDF Run II inclusive jet data, χ^2 = 108 for 72 pts.



Data/theory the same shape for both. Good compatibility. Verified by fits.

Shifts large compared to statistical uncertainties. Contribution to χ^2 from this source from all jet fits less than or (at most) about the number of correlated sources, and Gaussian distribution of shifts to good approx.

Systematic uncertainty plots for D0. Similar for CDF. Also vary with rapidity



No other systematic looks like normalisation, though some are bigger.

NLO PDF (with NLO $\hat{\sigma}$)	$\mu = p_T/2$	$\mu = p_T$	$\mu = 2p_T$
MSTW08	0.75 (+0.32)	0.68 (−0.88)	0.63 (−*2.69*)
CTEQ6.6	1.03 (−*2.47*)	1.04 (−**3.49**)	0.99 (−**4.75**)
CT10	0.99 (−*1.64*)	0.92 (−*2.69*)	0.86 (−**4.10**)
NNPDF2.1	0.74 (−0.33)	0.79 (−*1.60*)	0.80 (−**3.12**)
HERAPDF1.0	1.52 (−**4.07**)	1.57 (−**5.21**)	1.43 (−**6.22**)
HERAPDF1.5	1.48 (−**3.85**)	1.52 (−**5.00**)	1.39 (−**6.03**)
ABKM09	1.03 (−**3.49**)	1.01 (−**4.53**)	1.05 (−**5.80**)
GJR08	1.14 (+*2.47*)	0.93 (+*1.25*)	0.79 (−0.50)
NNLO PDF (with NLO+2-loop $\hat{\sigma}$)	$\mu = p_T/2$	$\mu = p_T$	$\mu = 2p_T$
MSTW08	1.39 (+0.35)	0.69 (−0.45)	0.97 (−*1.30*)
HERAPDF1.0, $\alpha_S(M_Z^2) = 0.1145$	2.37 (−*2.65*)	1.48 (−**3.64**)	1.29 (−**4.12**)
HERAPDF1.0, $\alpha_S(M_Z^2) = 0.1176$	2.24 (−0.48)	1.13 (−*1.60*)	1.09 (−*2.23*)
ABKM09	1.53 (−**4.27**)	1.23 (−**5.05**)	1.44 (−**5.65**)
JR09	0.75 (+0.13)	1.26 (−0.61)	2.20 (−*1.22*)

Table 1: Values of $\chi^2/N_{\rm pts.}$ (using FastNLO numbers) for the CDF Run II inclusive jet data using the k_T jet algorithm. No restriction is imposed on the shift in normalisation and the optimal value of "$-r_{\rm lumi.}$" is shown in brackets.

Imperative to tie W, Z normalisation to that of jets in fit. The same at the LHC.