The NNPDF Approach to Parton Determination: Ideas and Results

Stefano Forte
Università di Milano & INFN
for the NNPDF collaboration:
R. D. Ball, V. Bertone, F. Cerutti, L. del Debbio,
S.F., A. Guffanti, J. I. Latorre, J. Rojo, M. Ubiali

Università degli Studi di Milano
Dipartimento di Fisica

Loopfest 2010 Stony Brook, June 23, 2010

SUMMARY

- THE NNPDF APPROACH
- NNPDF2.0 PDFs: THE FIRST GLOBAL NLO FIT
- NNPDF2.0 PDFs: CONSISTENCY, STABILITY, PREDICTIVITY

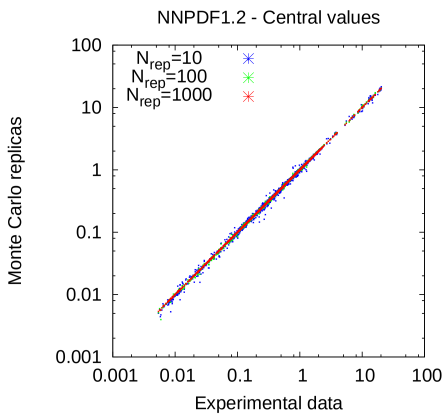
THE NNPDF APPROACH

THE NNPDF APPROACH:
THE NEURAL MONTE CARLO

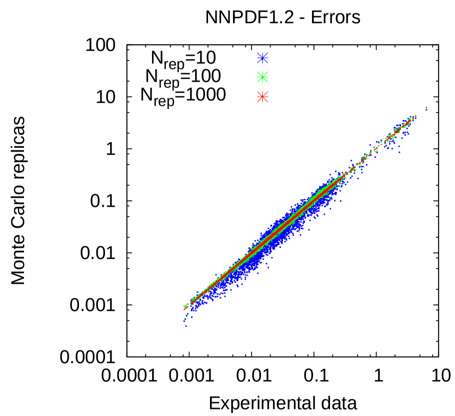
BASIC IDEA: MONTE CARLO SAMPLING OF THE PROBABILITY MEASURE IN THE (FUNCTION) SPACE OF PDFs

- START FROM MONTE CARLO SAMPLING OF DATA SPACE
- SPACE OF FUNCTIONS HUGE
  5 BINS FOR 10 PTS$\times$ 7 FCTNS $\rightarrow 5^{70} \sim 10^{49}$ BINS
- IMPORTANCE SAMPLING: DATA TELL US WHICH BINS ARE POPULATED

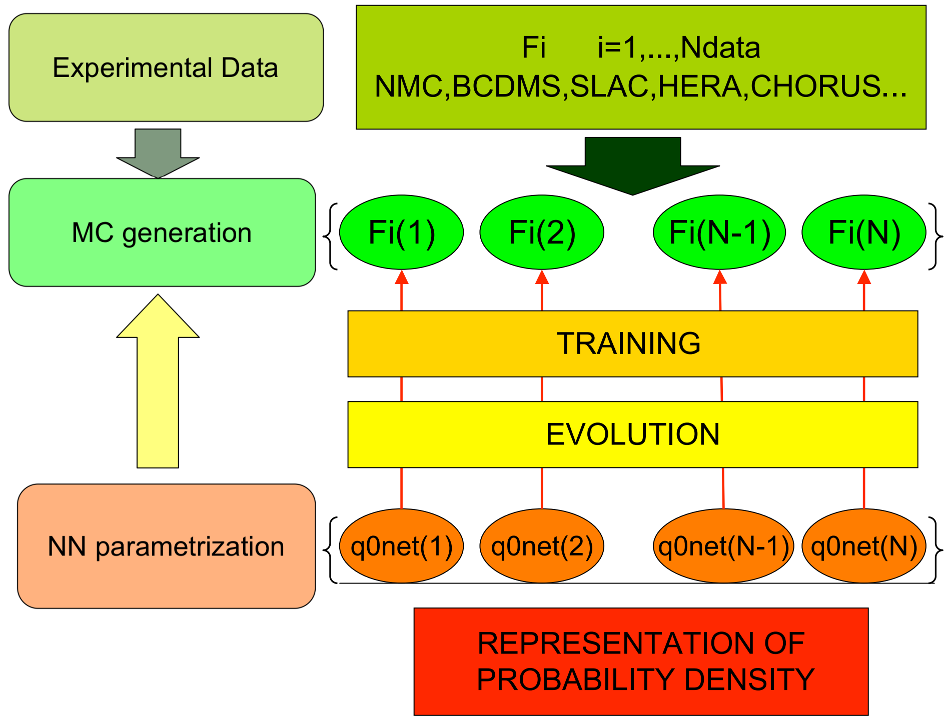
replica averages vs. central values



replica standard dev. vs. uncertainties

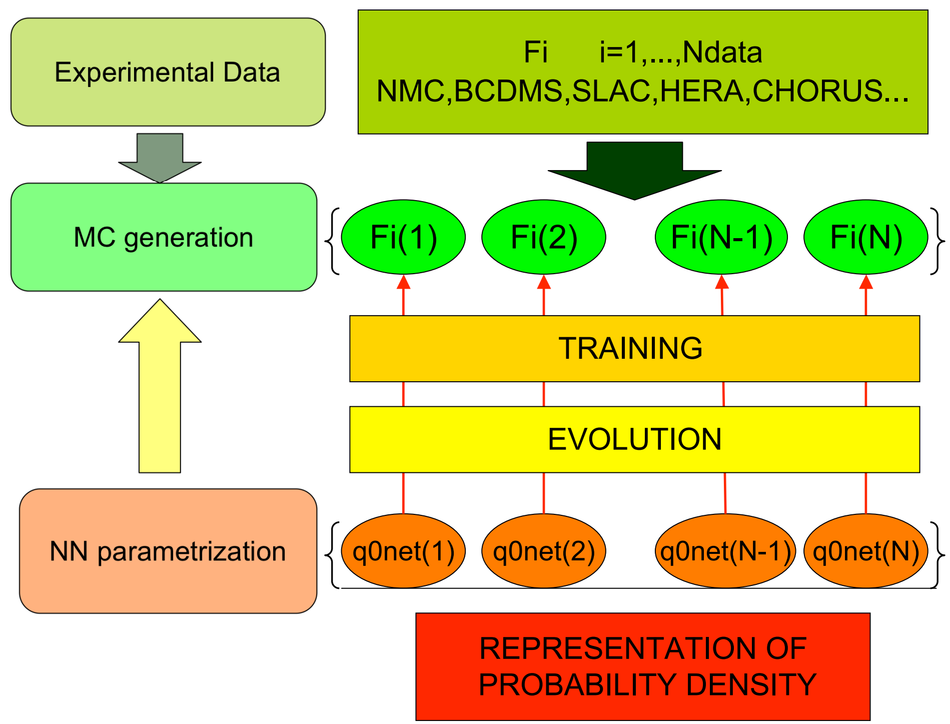


10 REPLICAS ENOUGH FOR CENTRAL VALS, 100 FOR UNCERTAINTIES, 1000 FOR CORRELNS



DATA MONTE CARLO $\Rightarrow$ PDF MONTE CARLO
NEURAL NETWORK PARM+ CROSS-VALIDATION METHOD

- EACH PDF$\leftrightarrow$ NEURAL NETWORK PARAMETRIZED BY 37 PARAMETERS
- NNPDF1.2: $37 \otimes 7 = 259$ PARMS (RECALL MSTW, CTEQ $\to$ 20 FREE PARAMETERS)
  "INFINITE" NUMBER OF PARAMETERS$\Rightarrow$ CAN REPRESENT ANY FUNCTION
- COMPLEX SHAPES (LARGE NO.OF PARAMETERS) REQUIRE LONGER FITTING
- FIT STOPS WHEN QUALITY OF FIT TO RANDOMLY SELECTED "VALIDATION" DATA (NOT FITTED) STOPS IMPROVING
- CAN OBTAIN A FIT WITH χ^2 LOWER THAN BEST FIT ("OVERLEARNING")

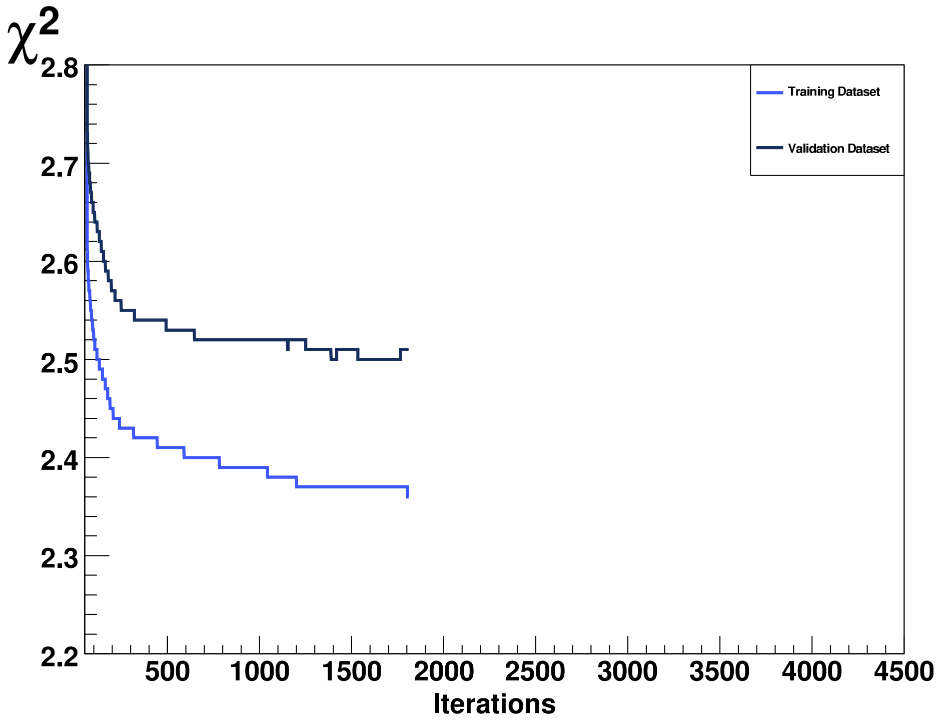


DATA MONTE CARLO ⇒ PDF MONTE CARLO
CROSS-VALIDATION

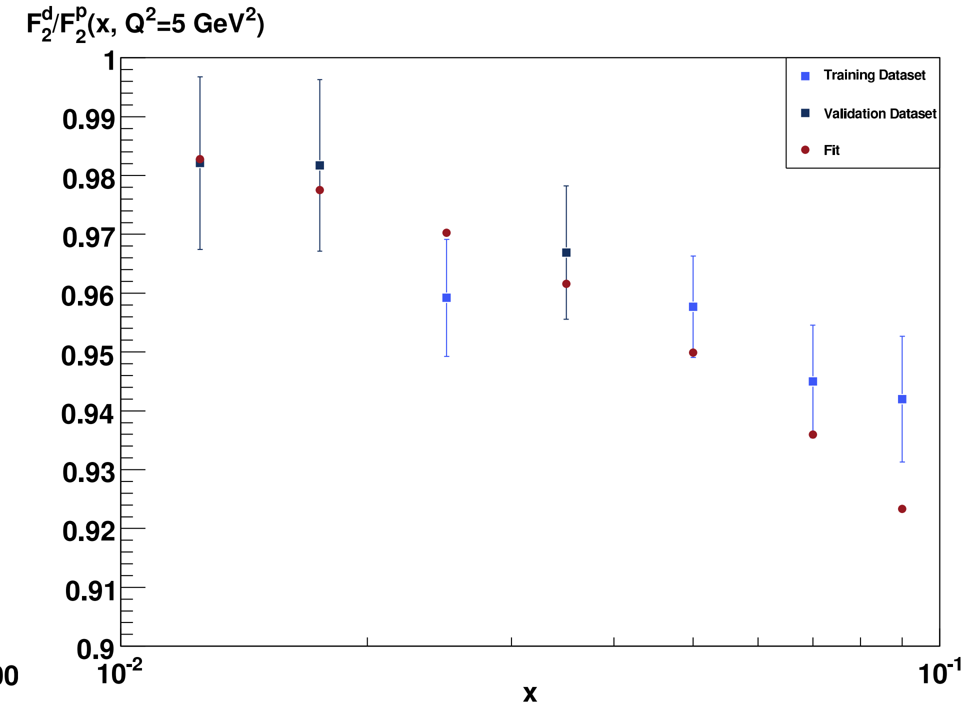
- REPLICAS ARE FITTED TO A DATA SUBSET
- A DIFFERENT SUBSET OF DATA USE FOR EACH REPLICA
- OPTIMAL FIT WHEN FIT TO VALIDATION (CONTROL) DATA STOPS IMPROVING
-

OPTIMAL FITTING

χ^2



FIT TO DATA

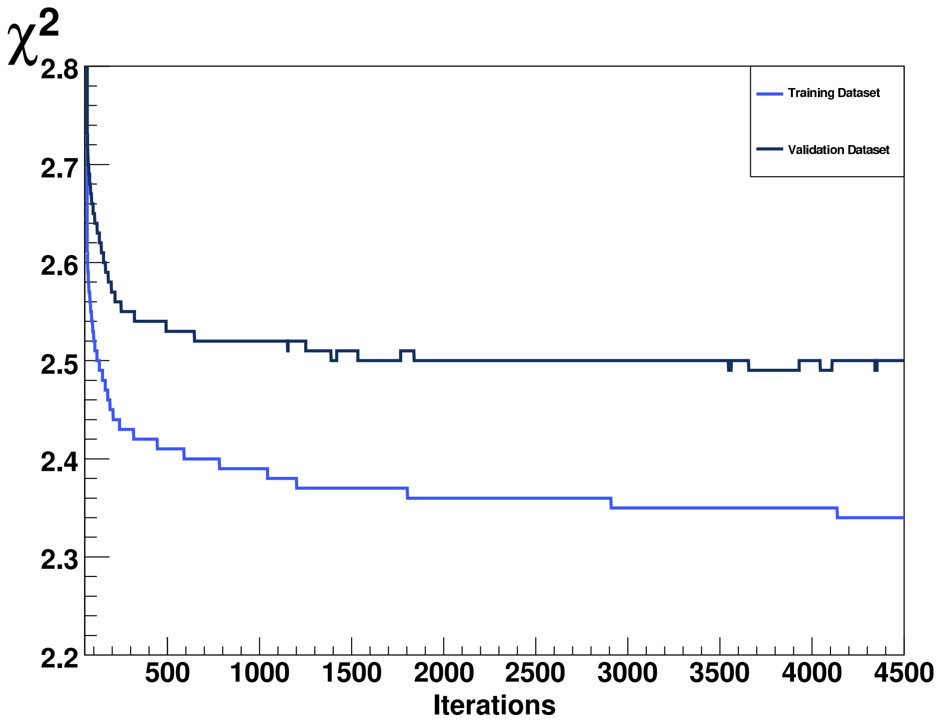


DATA MONTE CARLO ⇒ PDF MONTE CARLO
CROSS-VALIDATION

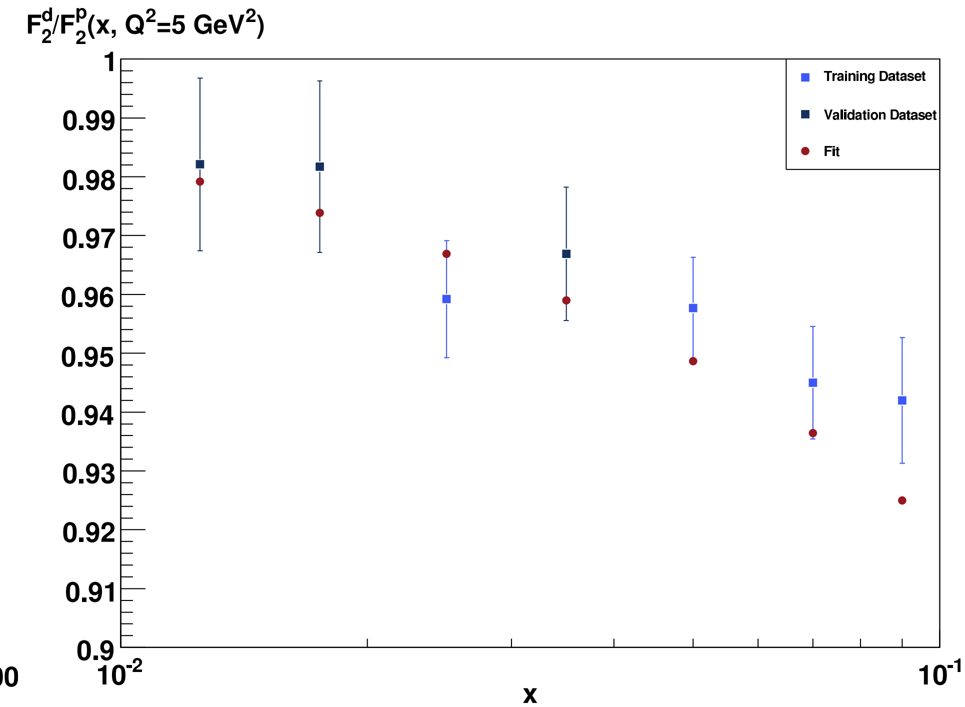
- REPLICAS ARE FITTED TO A DATA SUBSET
- A DIFFERENT SUBSET OF DATA USE FOR EACH REPLICA
- OPTIMAL FIT WHEN FIT TO VALIDATION (CONTROL) DATA STOPS IMPROVING
- THE BEST FIT IS NOT AT THE MINIMUM OF THE χ^2

OVERFITTING

χ^2



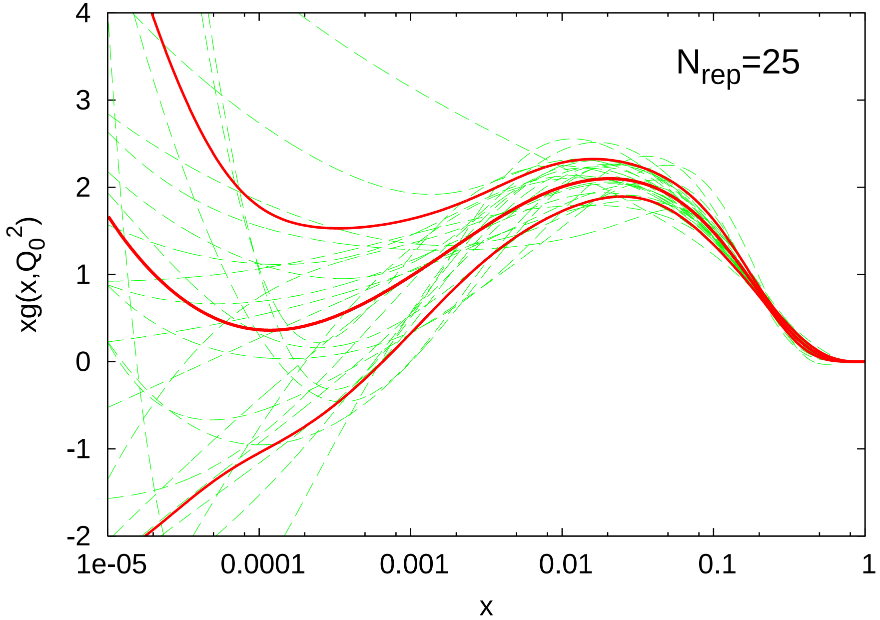
FIT TO DATA

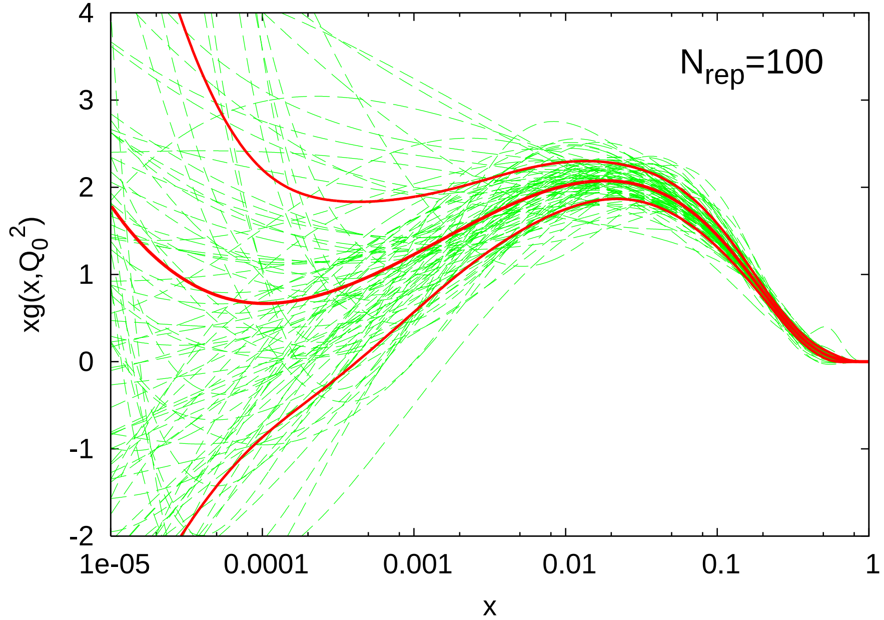


WHAT IS A ONE-σ UNCERTAINTY?

NNPDF: THE CENTRAL 68% OF THE MC DISTRIBUTION OF PDFs

Example: the gluon distribution in the NNPDF1.0 set





- ENSEMBLE OF REPLICAS $\leftrightarrow$ PROBABILITY DISTRIBUTION OF PDFs
- EXPECTED CENTRAL VALUE $\leftrightarrow$ MEAN; UNCERTAINTY $\leftrightarrow$ STANDARD DEVIATION
- ANY FEATURES OF DISTRIBUTION CAN BE DETERMINED (C.L. INTERVALS, CORRELATIONS...)

THE NNPDF TIMELINE

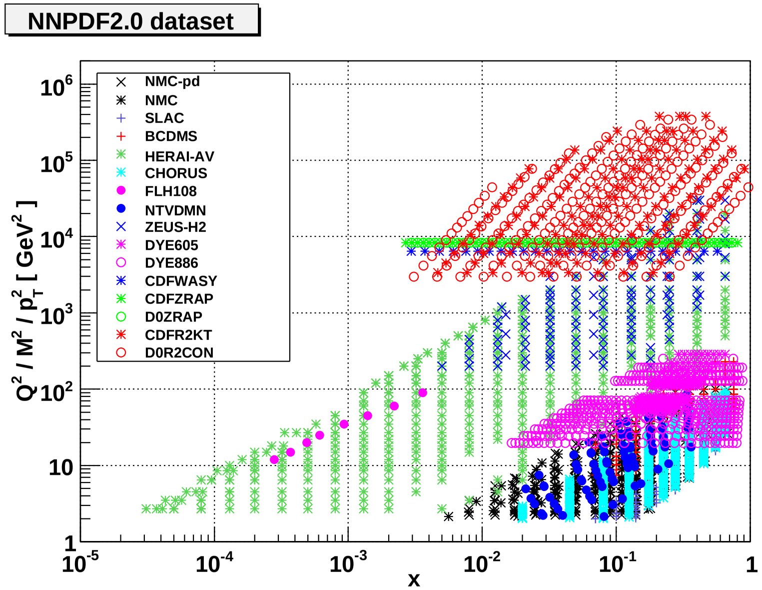
- 2002 FIT OF NUCLEON STRUCTURE FUNCTION TO TWO DATASETS
- 2005 FIT OF NUCLEON STRUCTURE FUNCTION TO MANY DATASETS
- 2007 FIT OF ONE SINGLE (ISOTRIPLET) PARTON DISTRIBUTION TO TWO DATASETS
- 2008 FIRST FULL PARTON SET: NNPDF1.0 (WORLD DIS DATA)
- 2009 PARTON SET WITH ACCURATE STRANGENESS: NNPDF1.2 (DIS+ DIMUON DATA)
- 2010 FIRST GLOBAL PARTON SET: NNPDF2.0

NNPDF2.0 PDFs:
THE FIRST GLOBAL NLO FIT

NNPDF VS. CURRENT PDF SETS

$$\sigma_X(s, M_X^2) = \sum_{a,b} \int_{x_{\min}}^1 dx_1\, dx_2\, f_{a/h_1}(x_1) f_{b/h_2}(x_2) \hat{\sigma}_{q_a q_b \to X}\left(x_1 x_2 s, M_X^2\right)$$

THE NNPDF2.0 DATASET



- fixet p, d target charged lepton DIS data (SLAC, NMC, BCDMS)
- HERA I combined ZEUS+H1 DIS data
- neutrino DIS data (CHORUS)
- neutrino dimuon production (NuTeV)
- fixed p, d target Drell-Yan (E605, E886)
- CDF W asymmetry
- CDF & D0 Z rapidity distn.
- CDF & D0 high E_T jets

- CTEQ6.6: GLOBAL, NLO, VFN WITH HQ MASS, SEVERAL α_s VALUES
- MSTW08: GLOBAL, NLO & NNLO, VFN WITH HQ MASS, SEVERAL α_s VALUES
- NNPDF2.0: GLOBAL, NLO, VFN WITHOUT HQ MASS, SEVERAL α_s VALUES
- ALEKHIN ABKM: DIS+SOME DY, NLO & NNLO, BMSN (FFN), SINGLE α_s VALUE
- HERAPDF1.0: ONLY HERA DATA, NLO, VFN WITH HQ MASS, SEVERAL α_s VALS.
- SETS BASED ON MODEL ASSUMPTIONS (GRV/GJR, STATISTICAL PDFS,...)

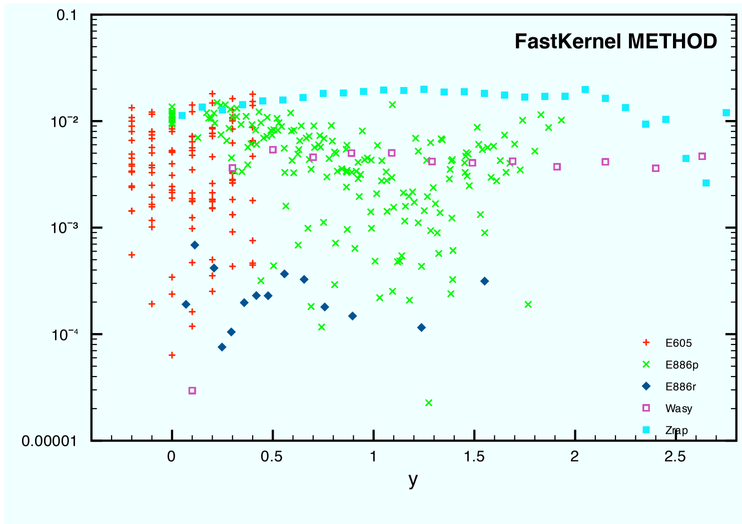
THE PROBLEM OF NLO FITTING

- OTHER EXISTING GLOBAL FITS ARE NOT FULLY NLO
  MSTW, CTEQ: DRELL-YAN TREATED AT LO+ K–FACTORS
- OTHER EXISTING NLO FITS ARE NOT GLOBAL
  HERAPDF ONLY DIS, ALEKHIN (ABKM) DIS+SOME FIXED-TARGET DRELL-YAN
- BOTTLENECK: FAST COMPUTATION OF DOUBLE CONVOLUTIONS FOR HADRONIC PROCESSES
  MELLIN SPACE APPROACH INCOVENIENT FOR JETS, & FOR GENERAL PARTON PARAMETRIZATIONS

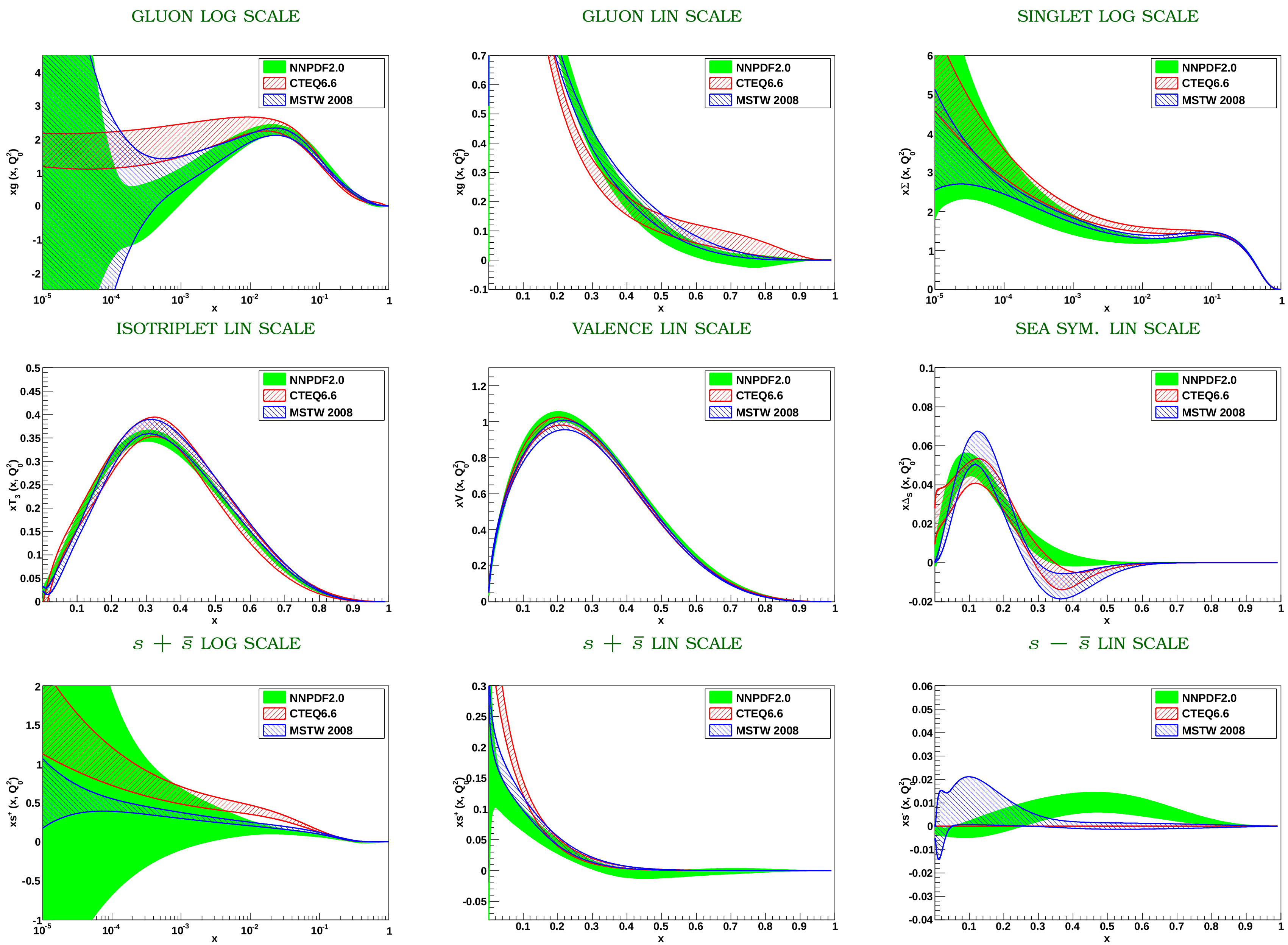
NNPDF2.0: THE FIRST GLOBAL NLO FIT

- GRID-BASED METHODS: EXPANSION OF PDFS ON BASES OF POLYNOMIALS, PRECOMPUTE CONVOLUTION WITH BASIS FUNCTIONS (Pascaud, Zomer, 2001)
- FASTNLO: FAST INTERFACE FOR JET CROSS SECTIONS (Kluge, Rabbertz, Wobisch 2006)
- NNPDF2.0 USES FASTKERNEL: GRID METHOD INTERFACED TO N-SPACE COMPUTATION OF GLAP GREEN FUNCTIONS, INTERFACED TO FASTNLO FOR JETS AND TO SUITABLE FAST-DY (NNPDF, 2010)
- MORE RECENTLY APPLGRID: OPTIMIZED GRID, POTENTIALLY UNIVERSAL INTERFACE, IMPLEMENTED FOR JETS, W AND Z PRODUCTION (Carli et al., 2010)

FASTKERNEL PERC. ACCURACY



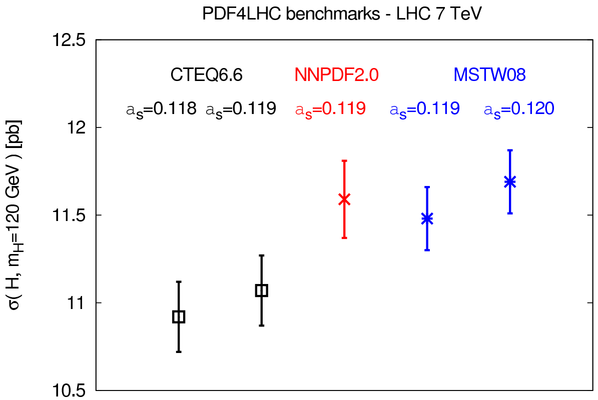
NNPDF2.0 PDFS VS. OTHER GLOBAL SETS



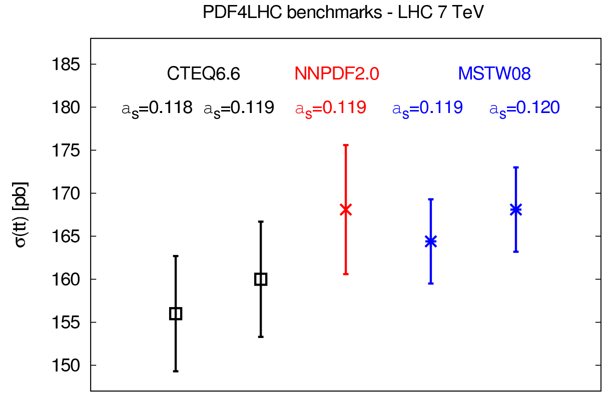
NNPDF2.0 STANDARD CANDLES
VS. OTHER GLOBAL SETS

- RESULTS FOR LHC 7 TeV
- PDF4LHC BENCHMARKS: COMMON VERSION OF MCFM AND PARAMETERS
- GOOD AGREEMENT AT THE ONE-SIGMA LEVEL
- VALUE OF α_s IMPORTANT

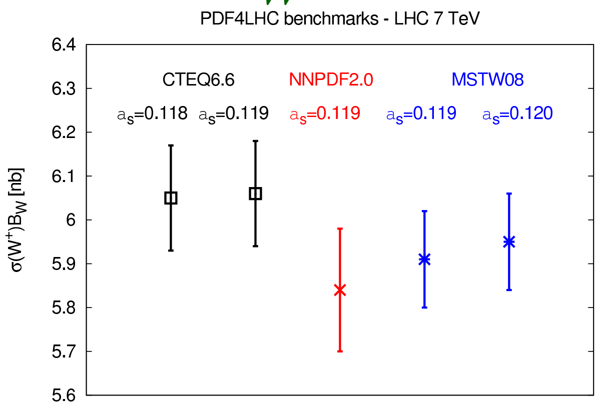
HIGGS 120 GeV



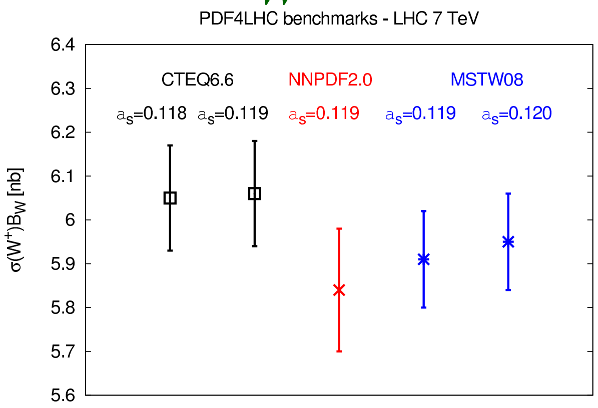
TOP



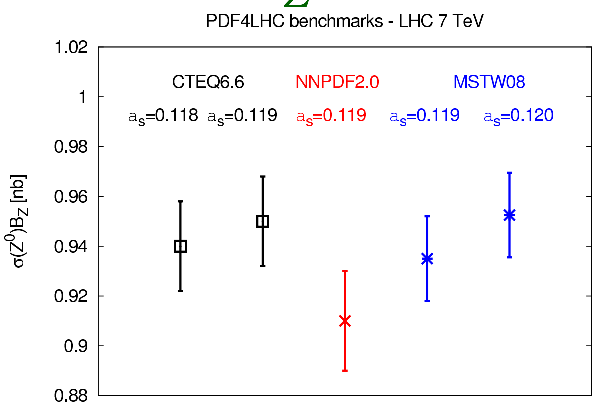
W^+



W^-



Z



THE NNPDF2.0 PDFs: CONSISTENCY, STABILITY, PREDICTIVITY

DATA VS. THEORY:
DETAILED TESTS

- GLOBAL FITS INCLUDE DATASETS IN DISPARATE KINEMATIC REGIONS
- MUTUAL CONSISTENCY $\Rightarrow$ TEST OF DGLAP EVOLUTION
- INCONSISTENT DATA $\Rightarrow$ NO IMPROVEMENT IN PDF ACCURACY
- CAN TEST BY COMPARING FITS TO DIFFERENT DATA SUBSETS

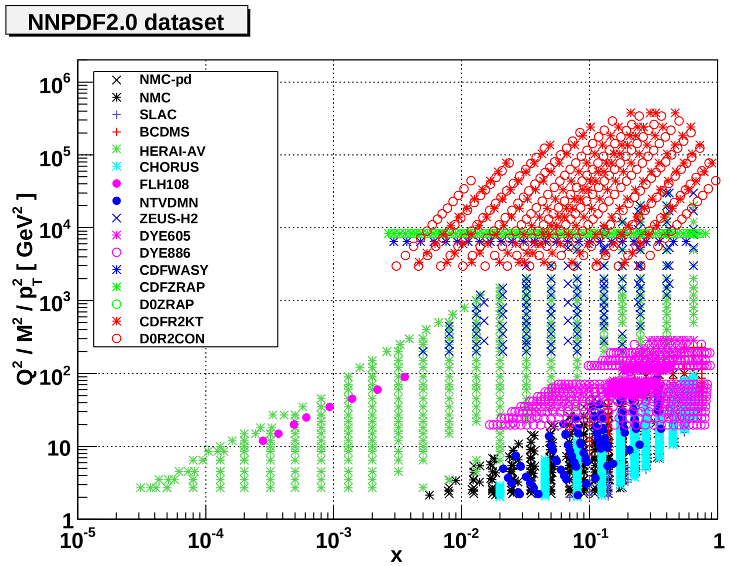
A USEFUL TOOL:
DISTANCE BETWEEN PDF SETS

$$d^2\left(\langle q^{(1)}\rangle, \langle q^{(2)}\rangle\right) = \frac{\left(\langle q^{(1)}\rangle_{(1)} - \langle q^{(2)}\rangle_{(2)}\right)^2}{\sigma^2_{(1)}[\langle q^{(1)}\rangle] + \sigma^2_{(2)}[\langle q^{(2)}\rangle]}; \quad \sigma^2_{(i)}[\langle q^{(i)}\rangle] = \frac{1}{N^{(i)}_{\rm rep}}\sigma^2_{(i)}[q^{(i)}]$$

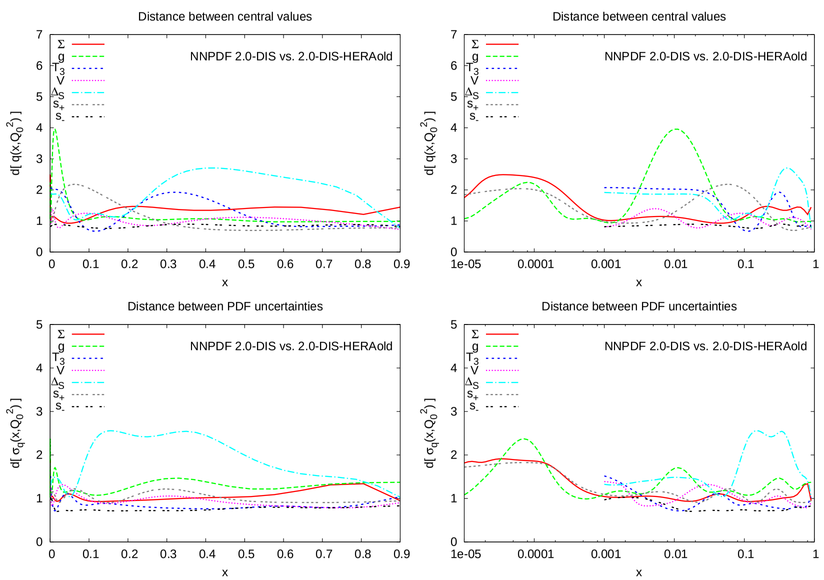
- MEAN PDF EXTRACTED FROM MONTECARLO SET OF SIZE $N_{\rm rep}$ FLUCTUATES WITH STANDARD DEVIATION $\frac{\sigma}{N_{\rm rep}}$
- IF TWO MC SETS EXTRACTED FROM SAME UNDERLYING DISTRIBUTION, THEN $d \sim 1$

FLEXIBILITY $\Rightarrow$ STABILITY: OLD HERA DATA VS. COMBINED HERA DATA

	DIS-HERAold	DIS	NNPDF2.0
χ^2_{tot}	1.12	1.20	1.21
NMC-pd	0.87	0.85	0.99
NMC	1.65	1.69	1.69
SLAC	1.33	1.37	1.34
BCDMS	1.25	1.26	1.27
HERAI	**0.96**	**1.13**	1.14
CHORUS	1.12	1.13	1.18
FLH108	1.53	1.51	1.49
NTVDMN	0.71	0.71	0.67
ZEUS-H2	1.49	1.50	1.51
CDFR2KT	0.78	0.91	0.80
D0R2CON	0.94	1.00	0.93
DYE605	8.21	7.32	0.88
DYE866	2.46	2.24	1.28
CDFWASY	20.32	13.06	1.85
CDFZRAP	3.13	3.12	2.02
D0ZRAP	0.65	0.65	0.47

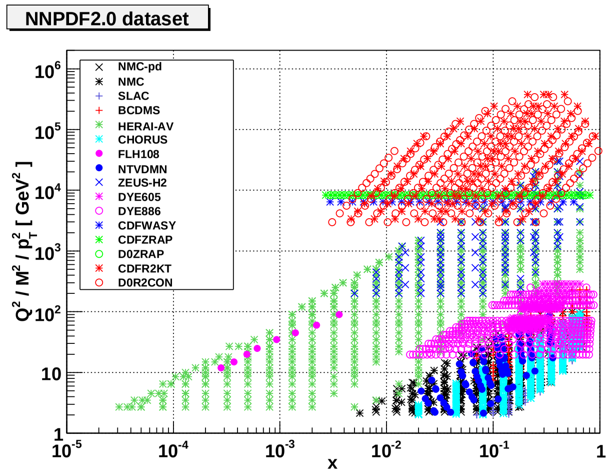


- PDFs UNCHANGED, MARGINAL IMPROVEMENT IN GLUON
- QUALITY OF FIT TO FIXED-TARGET DIS UNCHANGED
- SLIGHT DETERIORATION IN QUALITY OF FIT TO HERA DATA
- $\Rightarrow$ COULD BE EVIDENCE FOR (MARGINALLY) INADEQUATE THEORY

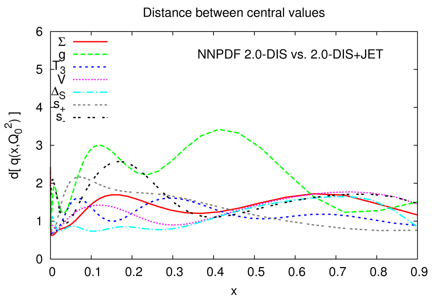


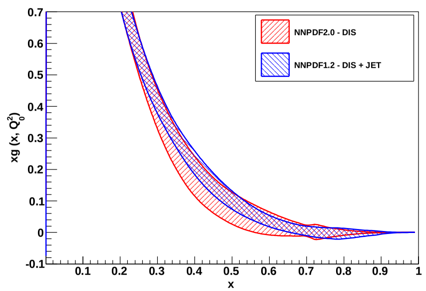
FLEXIBILITY $\Rightarrow$ CONSISTENCY: DIS DATA VS. JET DATA

	DIS	DIS+JET	NNPDF2.0
$\chi^2_{\rm tot}$	1.20	1.18	1.21
NMC-pd	0.85	0.86	0.99
NMC	1.69	1.66	1.69
SLAC	1.37	1.31	1.34
BCDMS	1.26	1.27	1.27
HERAI	1.13	1.13	1.14
CHORUS	1.13	1.11	1.18
FLH108	1.51	1.49	1.49
NTVDMN	0.71	0.75	0.67
ZEUS-H2	1.50	1.49	1.51
CDFR2KT	**0.91**	**0.79**	0.80
D0R2CON	**1.00**	**0.93**	0.93
DYE605	7.32	10.35	0.88
DYE866	2.24	2.59	1.28
CDFWASY	13.06	14.13	1.85
CDFZRAP	3.12	3.31	2.02
D0ZRAP	0.65	0.68	0.47



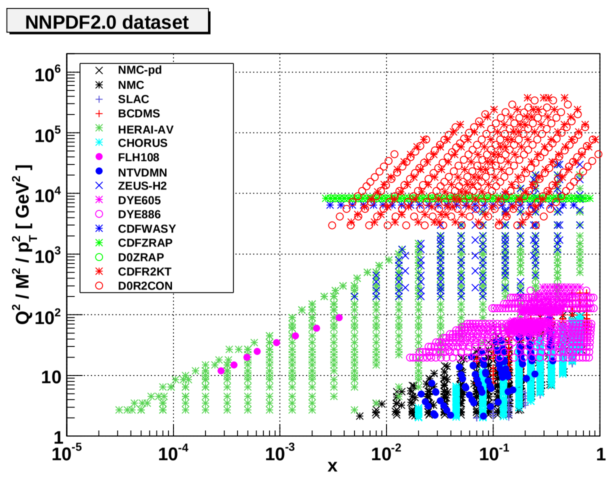
- HIGH E_T JET DATA WELL REPRODUCED EVEN WHEN NOT FITTED $\Rightarrow$ LARGE x GLUON WELL DETERMINED BY SCALING VIOLATIONS!
- SIGNIFICANT IMPROVEMENT IN LARGE x GLUON ACCURACY
- OTHER PDFS UNCHANGED



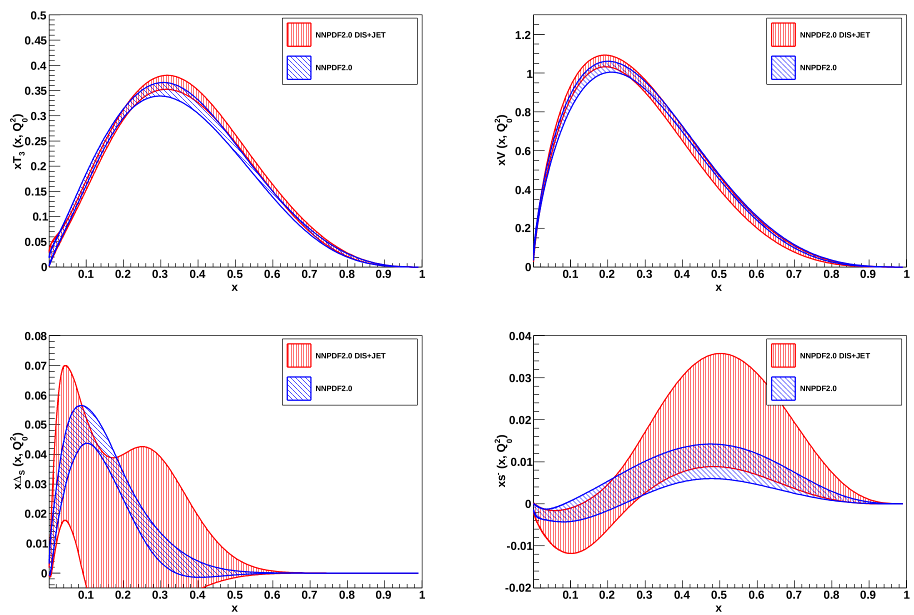


CONSISTENCY ⇒ ACCURACY: DIS+JETS VS. DRELL-YAN (AND W, Z) DATA

	DIS	DIS+JET	NNPDF2.0
$\chi^2_{\rm tot}$	1.20	1.18	1.21
NMC-pd	0.85	0.86	0.99
NMC	1.69	1.66	1.69
SLAC	1.37	1.31	1.34
BCDMS	1.26	1.27	1.27
HERAI	1.13	1.13	1.14
CHORUS	1.13	1.11	1.18
FLH108	1.51	1.49	1.49
NTVDMN	0.71	0.75	0.67
ZEUS-H2	1.50	1.49	1.51
CDFR2KT	0.91	0.79	0.80
D0R2CON	1.00	0.93	0.93
DYE605	7.32	**10.35**	**0.88**
DYE866	2.24	**2.59**	**1.28**
CDFWASY	13.06	**14.13**	**1.85**
CDFZRAP	3.12	**3.31**	**2.02**
D0ZRAP	0.65	**0.68**	**0.47**



- VERY SUBSTANTIAL IMPROVEMENT IN FIT QUALITY WHEN DATA INCLUDED ⇒ SOME PDF COMBINATIONS POORLY DETERMINED WITHOUT THESE DATA
- HUGE IMPROVEMENT IN SEA ASYM $\bar{u}-\bar{d}$ & STRANGENESS $s-\bar{s}$
- SIGNIFICANT IMPROVEMENT IN TOTAL VALENCE $(\sum_i (q_i - \bar{q}_i))$ & ISOTRIPLET $(u+\bar{u}-(d+\bar{d}))$

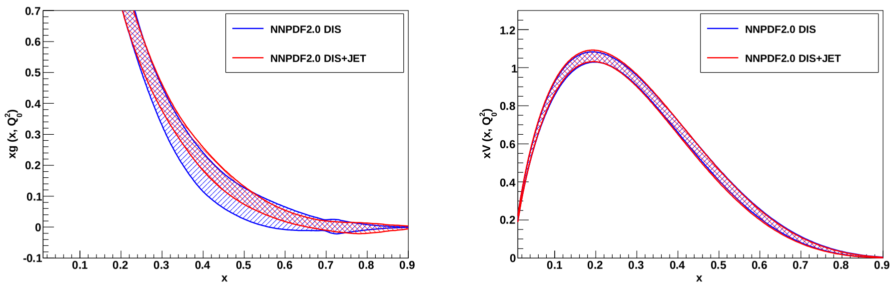


CONSISTENCY ⇒ NO TENSION!
DIS VS. HADRONIC DATA

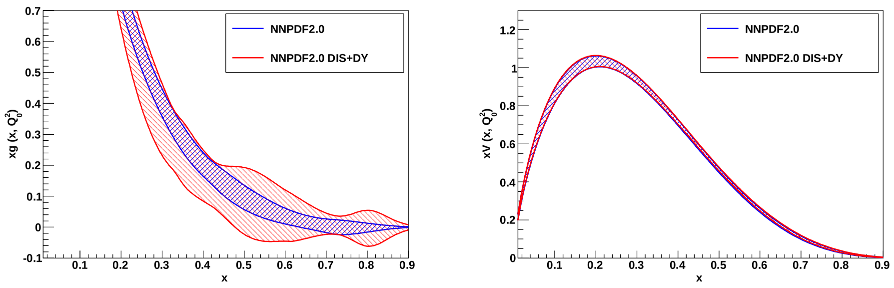
A SENSITIVE TEST: IS THE IMPACT OF A DATASET INDEP. OF THE DATA IT IS ADDED TO?

ADDING JET DATA...

. . . TO DIS DATA

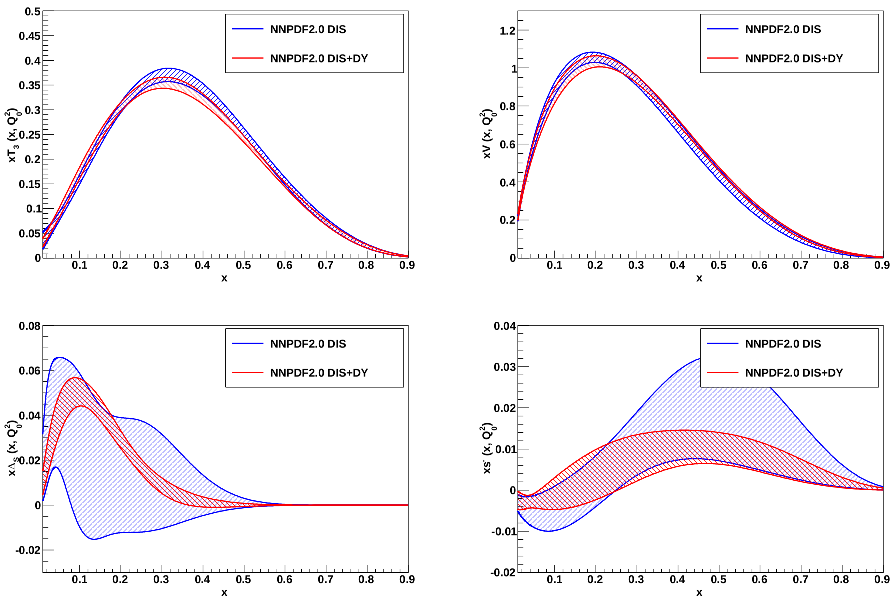


. . . TO DIS+DY DATA

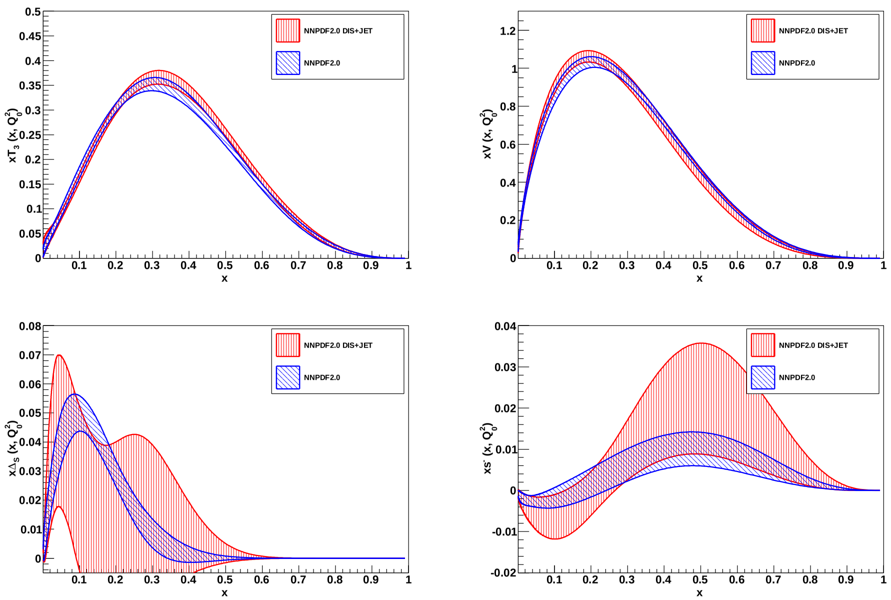


ADDING DRELL-YAN DATA...

. . . TO DIS DATA



. . . TO DIS+JET DATA



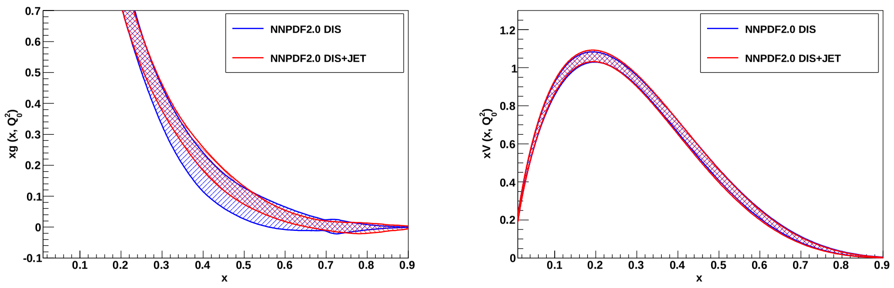
CONSISTENCY ⇒ NO TENSION!

DIS VS. HADRONIC DATA

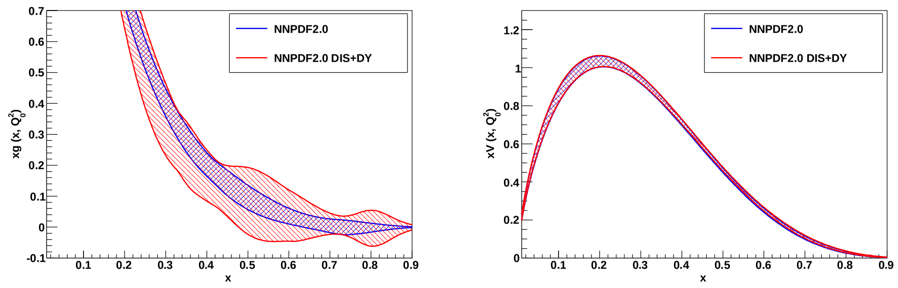
A SENSITIVE TEST: IS THE IMPACT OF A DATASET INDEP. OF THE DATA IT IS ADDED TO?

ADDING JET DATA. . .

. . . TO DIS DATA

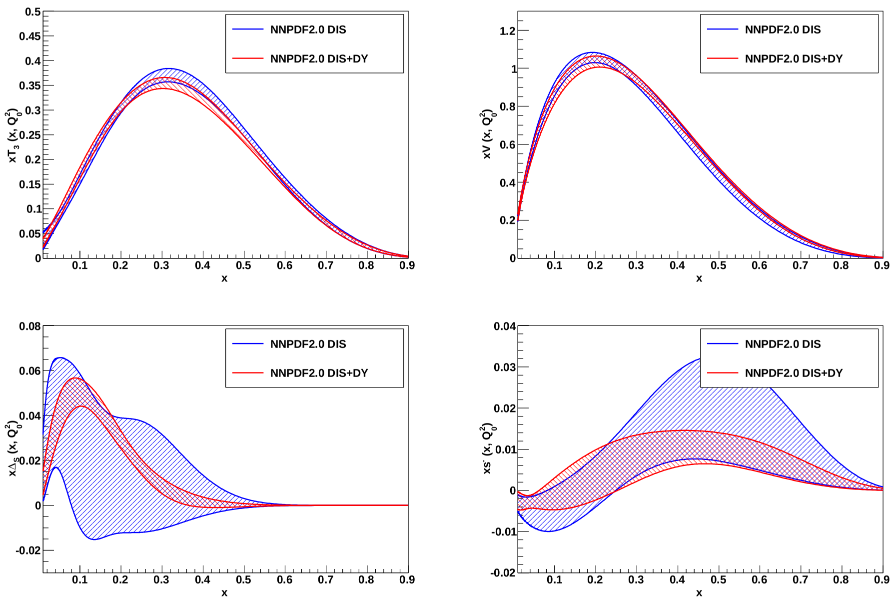


. . . TO DIS+DY DATA

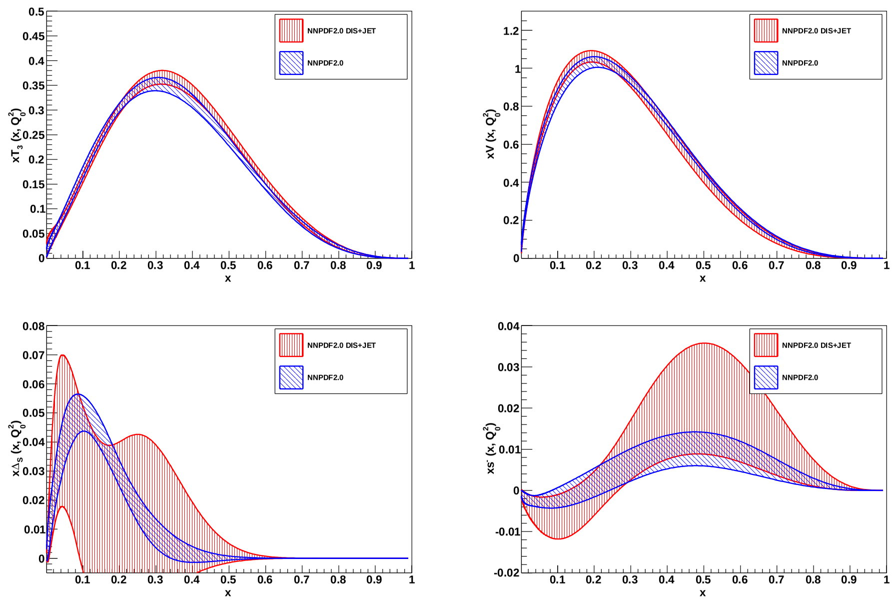


ADDING DRELL-YAN DATA. . .

. . . TO DIS DATA



. . . TO DIS+JET DATA



FITS COMMUTE ⇒ PERFECT COMPATIBILITY!

PRECISION PHENOMENOLOGY: AN EXAMPLE
THE NUTEV ANOMALY...

$$
\begin{aligned}
R_{\mathrm{PW}} &\equiv \frac{\sigma(\nu\mathcal{N}\to\nu X)-\sigma(\bar\nu\mathcal{N}\to\bar\nu X)}{\sigma(\nu\mathcal{N}\to\ell X)-\sigma(\bar\nu\mathcal{N}\to\bar\ell X)} \\
&= \frac{1}{2}-\sin^2\theta_{\mathrm{W}}+\left(\frac{(U^- - D^-)+(C^- - S^-)}{\mathcal{Q}^-}\frac{1}{6}\left(3-7\sin^2\theta_{\mathrm{W}}\right)\right),
\end{aligned}
$$

- PASCHOS-WOLFENSTEIN RATIO CAN BE MEASURED IN NEUTRINO DIS
- RESULT DEPENDS ON EW MIXING ANGLE, VALENCE ISOSPIN BREAKING (WITH ISOSINGLET TARGET), STRANGENESS VALENCE MOMENTUM ASYMMETRY
- STRANGENESS VALENCE MOMENTUM
  ASSUMED BY NUTEV TO VANISH $\Rightarrow$ THREE σ DISCREPANCY WITH GLOBAL FIT

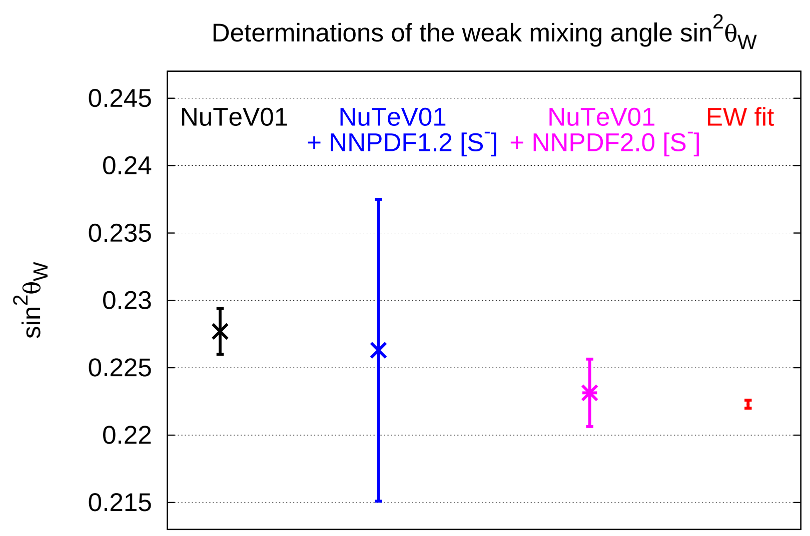
PRECISION PHENOMENOLOGY: AN EXAMPLE
THE NUTEV ANOMALY...

$$R_{\rm PW} \equiv \frac{\sigma(\nu\mathcal{N}\to\nu X)-\sigma(\bar\nu\mathcal{N}\to\bar\nu X)}{\sigma(\nu\mathcal{N}\to\ell X)-\sigma(\bar\nu\mathcal{N}\to\bar\ell X)} = \frac{1}{2}-\sin^2\theta_{\rm W}+\left(\frac{(U^- - D^-)+(C^- - S^-)}{\mathcal{Q}^-}\frac{1}{6}\left(3-7\sin^2\theta_{\rm W}\right)\right),$$

- PASCHOS-WOLFENSTEIN RATIO CAN BE MEASURED IN NEUTRINO DIS
- RESULT DEPENDS ON EW MIXING ANGLE, VALENCE ISOSPIN BREAKING (WITH ISOSINGLET TARGET), STRANGENESS VALENCE MOMENTUM ASYMMETRY
- STRANGENESS VALENCE MOMENTUM ASSUMED BY NUTEV TO VANISH ⇒ THREE σ DISCREPANCY WITH GLOBAL FIT

...IS GONE

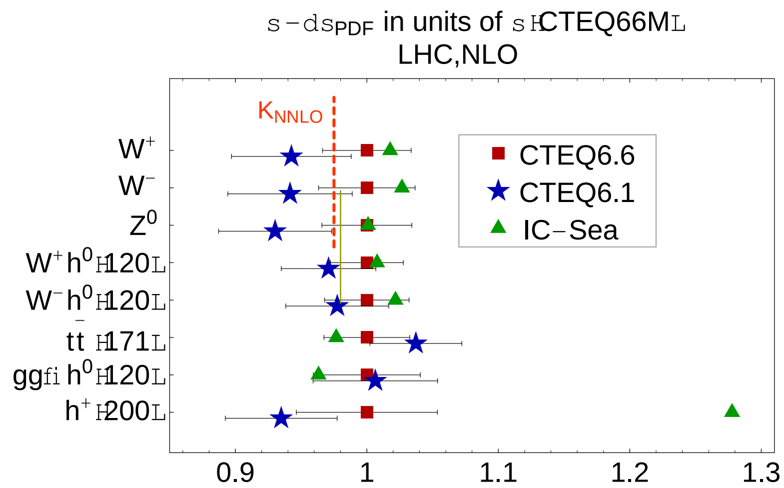
- NNPDF: s, $\bar s$ LIKE ANY OTHER PDF → 37 PARMS. EACH (CTEQ6.6: $s=\bar s$, TWO PARMS; MSTW08, S4S4, $\bar s$ TWO PARMS EACH)
- IF STRANGENESS UNCERTAINTY KEPT INTO ACCOUNT (DIS ONLY FIT: NNPDF1.2) ⇒ EFFECT LOSES STAT. SIGNIFICANCE
- IF HADRONIC DATA INCLUDED (NNPDF2.0 GLOBAL FIT) ⇒ STRANGENESS ASYMMETRY DETERMINED QUITE ACCURATELY → CORRECTED RESULT IN IMPRESSIVE AGREEMENT WITH SM GLOBAL FIT



(NNPDF2.0, 2010)

THE PROBLEM OF HEAVY QUARKS

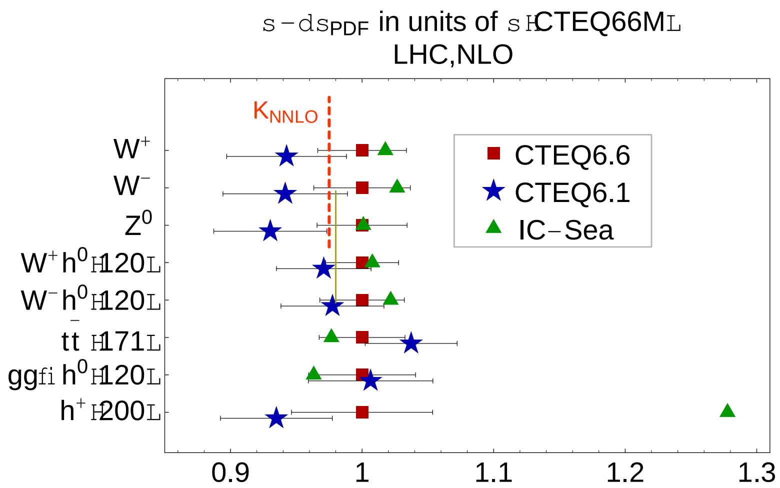
- MANY FITS (CTEQ<6, NNPDF, ALEKHIN<09) TREAT CHARM AS MASSLESS ABOVE THRESHOLD $\Rightarrow$ "ZMVFN" SCHEME
- TR/TR' PROCEDURE IMPLEMENTED SINCE '98 IN MRST
- WHEN CTEQ IMPLEMENTED ACOT IN 2008, SURPRISING CHANGE CTEQ61→CTEQ6.5 IN σ_W, & AGREEMENT WITH MRST SPOILED (LATER RESTORED)



(Nadolsky et al., 2008)

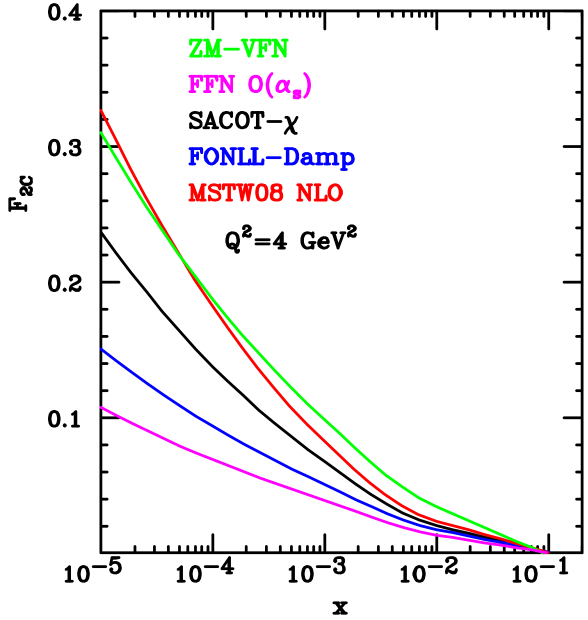
THE PROBLEM OF HEAVY QUARKS

- MANY FITS (CTEQ<6, NNPDF, ALEKHIN<09) TREAT CHARM AS MASSLESS ABOVE THRESHOLD $\Rightarrow$ "ZMVFN" SCHEME
- TR/TR' PROCEDURE IMPLEMENTED SINCE '98 IN MRST
- WHEN CTEQ IMPLEMENTED ACOT IN 2008, SURPRISING CHANGE CTEQ61→CTEQ6.5 IN σ_W, & AGREEMENT WITH MRST SPOILED (LATER RESTORED)



(Nadolsky et al., 2008)

RECENT PROGRESS: THE LES HOUCHES 2009 BENCHMARKS



(Rojo et al., 2010)

- TR, FONLL AND ACOT FOR DIS BENCHMARKED AT NLO AND NNLO
- $O(\alpha_s)$ FONLL, ACOT COINCIDE EXACTLY, TR' DIFFERS BY SUBLEADING $O(\alpha_s^2(m_c))$ Q^2–INDEP. TERM
- DIFFERENCES BETWEEN DAMPING PRESCRIPTIONS SIZABLE
- DIFFERENCE BETWEEN DIFFERENT PRESCRIPTIONS (ACOT-χ–SCALING, FONLL-DAMPING, MSTW-MATCHING) AS LARGE AS DIFFERENCE BETWEEN FFN (NO DGLAP RESUMMATION FOR CHARM) AND ZMVFN (NO CHARM MASS)

HEAVY QUARKS IN NNPDF

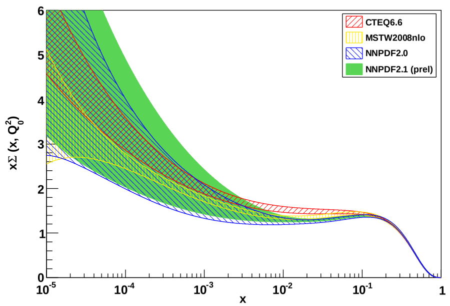
FLEXIBILITY ⇒ STABILITY (PDFs)

Set	NNPDF2.0	NNPDF2.1
NMC-pd	1.04	0.96
NMC	1.69	1.61
SLACp	1.57	1.41
SLACd	1.14	1.13
BCDMSp	1.30	1.22
BCDMSd	1.16	1.11
HERA1-NCep	1.32	1.29
HERA1-NCem	0.85	0.82
HERA1-CCep	0.97	0.96
HERA1-CCem	0.57	0.57
CHORUSnu	1.26	1.26
CHORUSnb	1.04	1.05
FLH108	1.50	1.32
NTVnuDMN	0.72	0.54
NTVnbDMN	0.70	0.81
Z06NC	1.17	1.14
Z06CC	1.27	1.25
DYE605	0.85	0.82
DYE886p	1.27	1.32
DYE886r	0.87	0.70
CDFWASY	1.84	2.10
CDFZRAP	1.98	1.83
D0ZRAP	0.56	0.56
CDFR2KT	0.79	0.79
D0R2CON	0.92	0.91

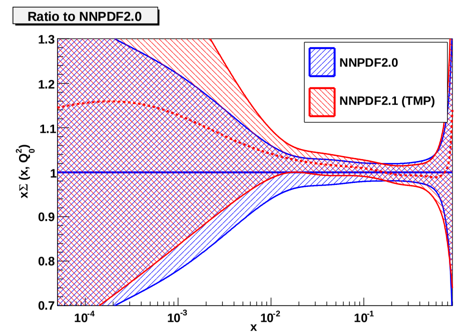
- QUALITY OF FIT TO FIXED-TARGET DATA IMPROVES
- QUALITY OF FIT TO HERA DATA UNCHANGED $\to$ HQ EFFECTS ABSORBED IN THE PDFS: IF HQ INCLUDED BUT NOT FITTED, $\chi^2 = 1.29 \Rightarrow \chi^2 = 1.56$ (NO OTHER DATA IN HERA REGION)
- QUALITY OF FIT TO HADRONIC DATA UNCHANGED
- SMALL x SINGLET GOES UP BY $1/3\sigma$, OTHER PDFS UNCHANGED

SINGLET

2.1 VS 2.1



2.1/2.0



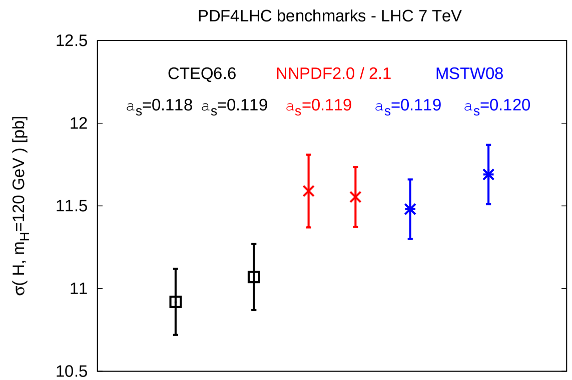
HEAVY QUARKS IN NNPDF

FLEXIBILITY ⇒ STABILITY (STANDARD CANDLES)

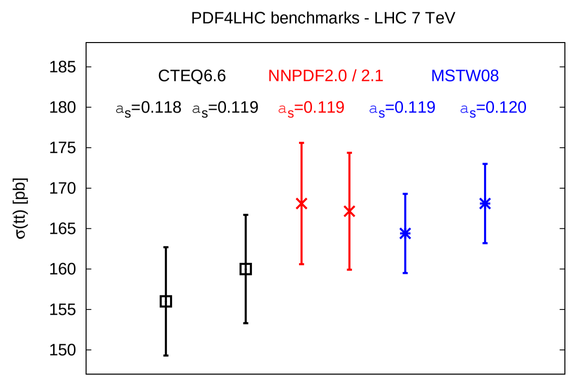
LHC 7 TeV:

- INCREASE IN W AND Z CROSS SECTION (AS EXPECTED)
- BUT IMPACT QUITE MODERATE: ABOUT $1/3\ \sigma$ (AS FOR PDFS)

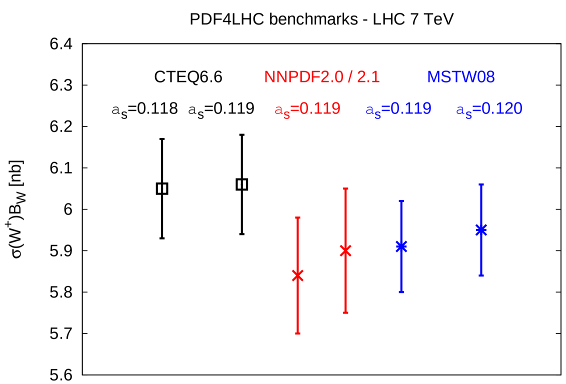
HIGGS 120 GeV



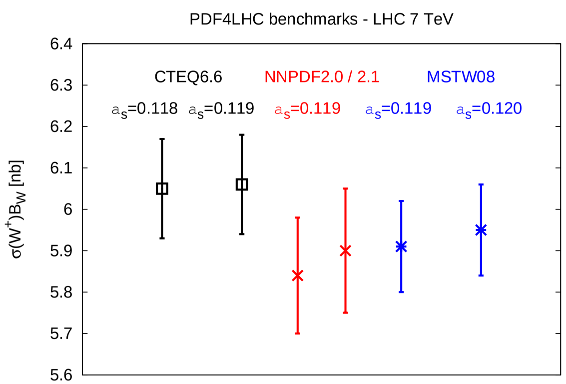
TOP



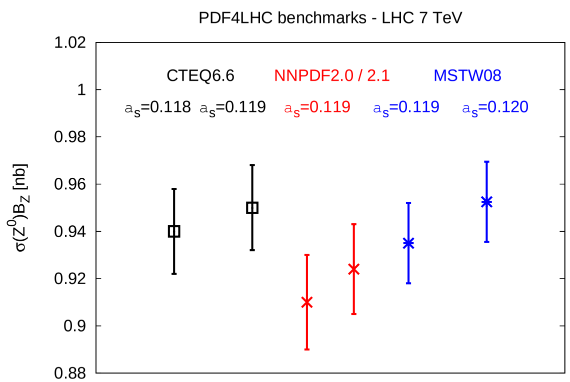
W^+



W^-



Z



Leading order PDFs

- LO PDFs needed for Monte Carlos
- MSTW+CTEQ LO → bad fit (χ^2 deteriorates by 25-30%)
- LO* PDFs suggested as compromise: LO PDFS with NLO α_s & no momentum sum rule

LEADING ORDER PDFs

- LO PDFs NEEDED FOR MONTE CARLOS
- MSTW+CTEQ LO → BAD FIT (χ^2 DETERIORATES BY 25-30%)
- LO* PDFs SUGGESTED AS COMPROMISE: LO PDFS WITH NLO α_s & NO MOMENTUM SUM RULE

NNPDF2.0 AT LEADING ORDER

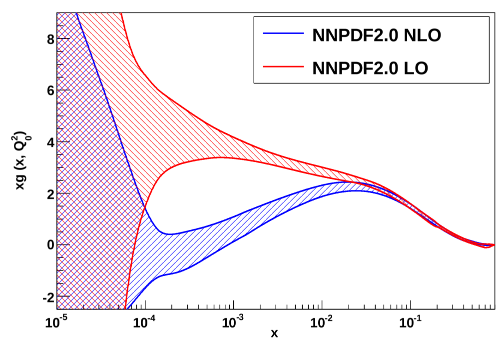
FLEXIBILITY ⇒ STABILITY (PDFs)

Fit	total χ^2	momentum
LO	1.29	1.0 (fix.)
LOmod	1.28	1.0 (fix.)
LO*	1.33	1.08 ± 0.05
NLO	1.21	1.0 (fix.)
NLO*	1.21	1.007 ± 0.023

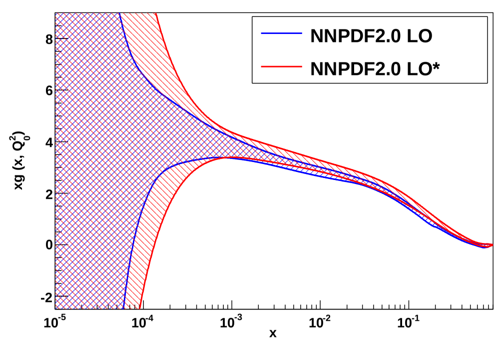
- LO FIT MARGINALLY WORSE THAN NLO
- NLO α_s IN LO FIT (LOMOD) → NO EFFECT
- MOMENTUM SUM RULE OK AT 1.5σ IN LO FIT
- LO* VS. LO/LOMOD ⇒ SIZABLY LARGER UNCERTAINTY

THE GLUON

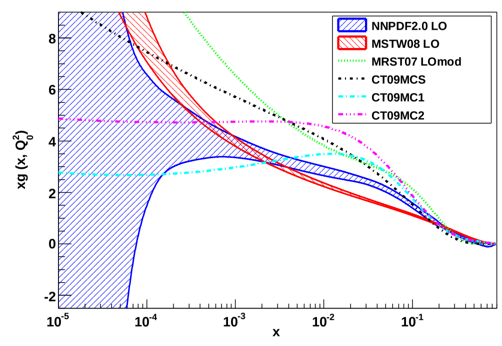
LO VS. NLO



LO VS. LO*



NNPDF VS. MSTW/TEA



LEADING ORDER PDFs

- LO PDFs NEEDED FOR MONTE CARLOS
- MSTW+CTEQ LO → BAD FIT (χ^2 DETERIORATES BY 25-30%)
- LO* PDFs SUGGESTED AS COMPROMISE: LO PDFS WITH NLO α_s & NO MOMENTUM SUM RULE

NNPDF2.0 AT LEADING ORDER

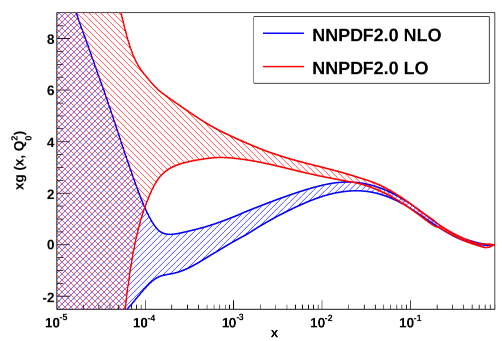
FLEXIBILITY ⇒ STABILITY (PDFs)

Fit	total χ^2	momentum
LO	1.29	1.0 (fix.)
LOmod	1.28	1.0 (fix.)
LO*	1.33	1.08 ± 0.05
NLO	1.21	1.0 (fix.)
NLO*	1.21	1.007 ± 0.023

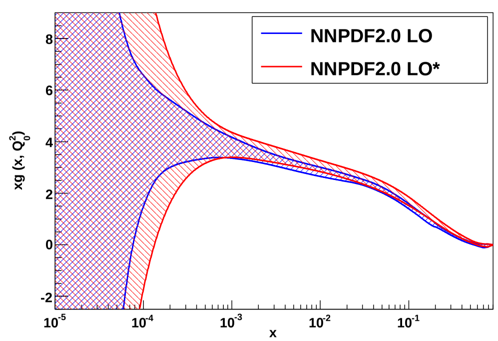
- LO FIT MARGINALLY WORSE THAN NLO
- NLO α_s IN LO FIT (LOMOD) → NO EFFECT
- MOMENTUM SUM RULE OK AT 1.5σ IN LO FIT
- LO* VS. LO/LOMOD ⇒ SIZABLY LARGER UNCERTAINTY

THE GLUON

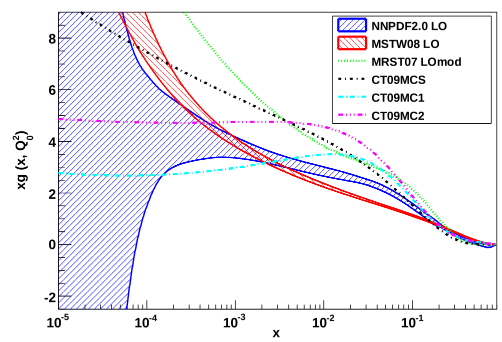
LO VS. NLO



LO VS. LO*



NNPDF VS. MSTW/TEA



NO NEED FOR LO*

DATA TEST THE MOMENTUM SR AT LO AND NLO

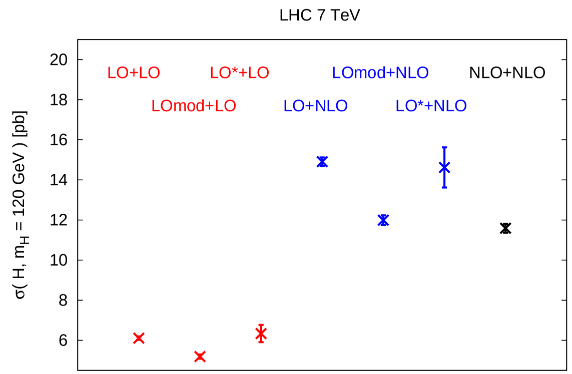
NNPDF2.0 AT LEADING ORDER

FLEXIBILITY $\Rightarrow$ STABILITY (STANDARD CANDLES)

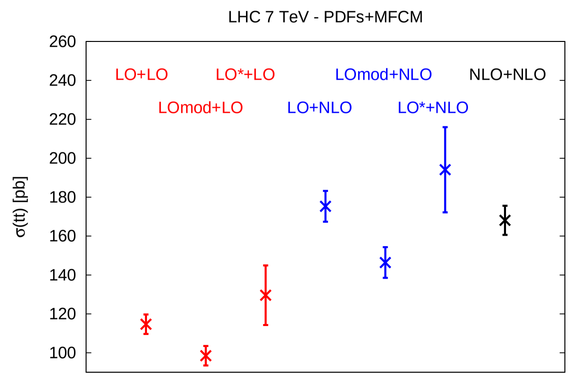
LHC 7 TeV:

- NLO MATRIX ELEMENTS NEEDED FOR PROCESSES WITH LARGE K FACTORS
- LO PDFS HAVE MODERATE IMPACT
- LO* $\to$ POOR MAN'S WAY OF ESTIMATING LO UNCERTAINTY (NOT RECOMMENDED)

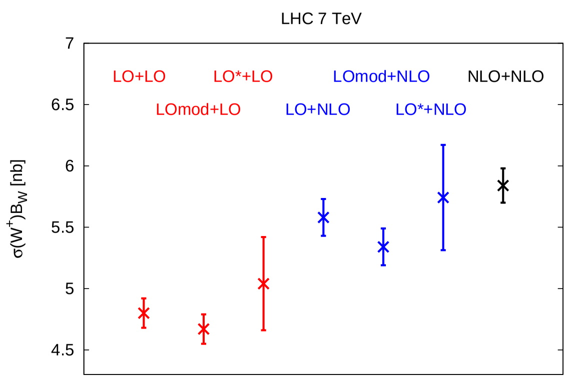
HIGGS 120 GeV



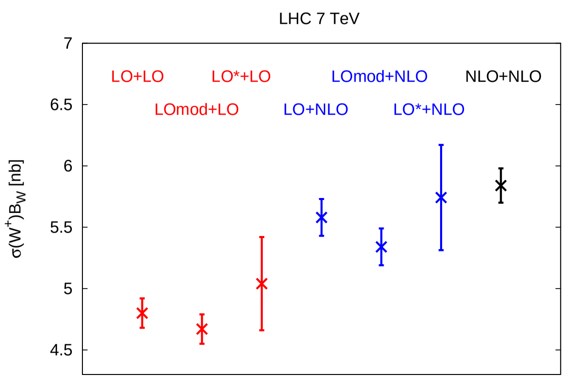
TOP



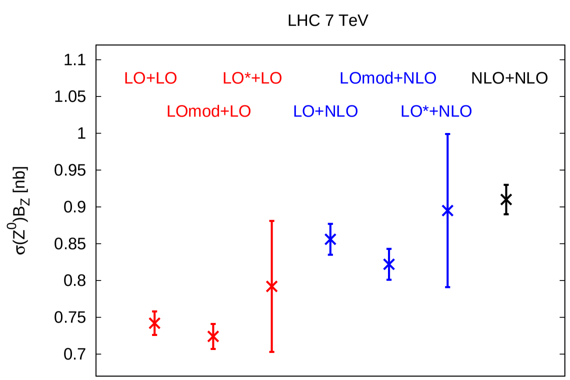
W^+



W^-



Z

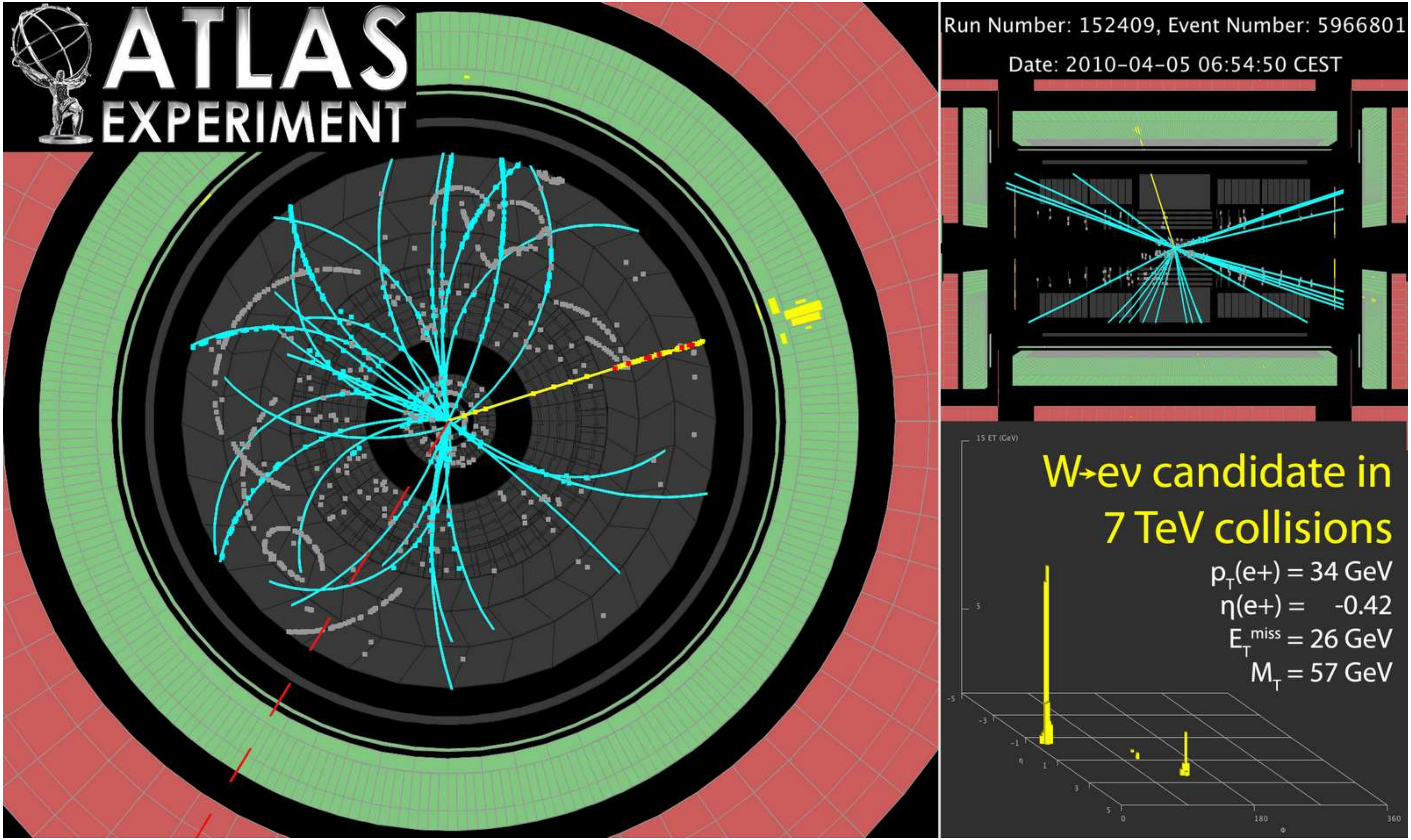


THE NNPDF TIMELINE

- 2002 FIT OF NUCLEON STRUCTURE FUNCTION TO TWO DATASETS
- 2005 FIT OF NUCLEON STRUCTURE FUNCTION TO MANY DATASETS
- 2007 FIT OF ONE SINGLE (ISOTRIPLET) PARTON DISTRIBUTION TO TWO DATASETS
- 2008 FIRST FULL PARTON SET: NNPDF1.0 (WORLD DIS DATA)
- 2009 PARTON SET WITH ACCURATE STRANGENESS: NNPDF1.2 (DIS+ DIMUON DATA)
- 2010 FIRST GLOBAL PARTON SET: NNPDF2.0
- JULY 2010: NNPDF2.1, A GLOBAL SET WITH FONLL/ACOT GENERAL MASS SCHEME
- END 2010: NNLO GLOBAL SET
- 2011: RESUMMED (SMALL AND LARGE x) SETS

CONCLUSION

THIS IS JUST THE BEGINNING!

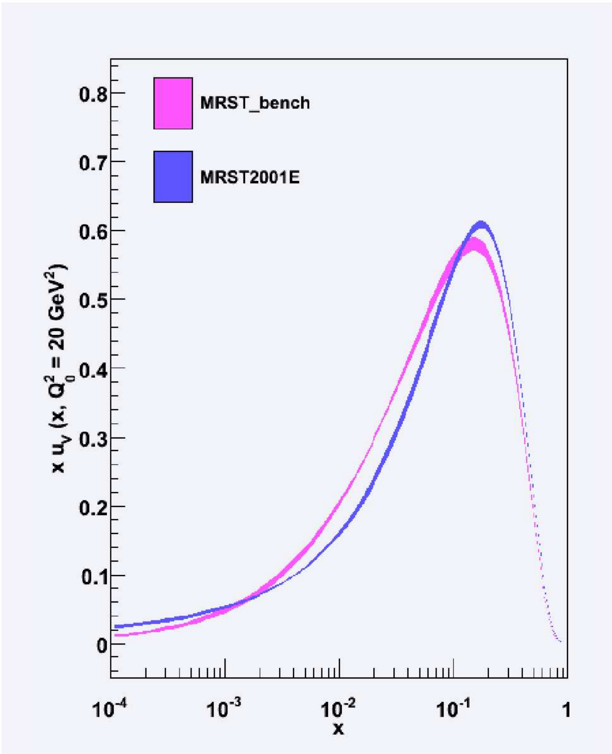


EXTRAS

WHAT DETERMINES PDF UNCERTAINTIES?

THE PROBLEM OF BENCHMARK FITS (HERALHC 2005-2008)

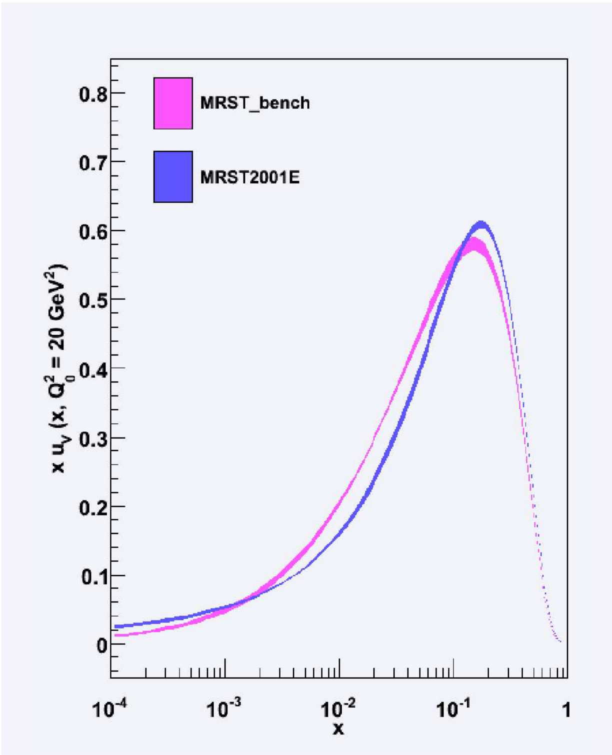
- PERFORM A MRST (MRSTBENCH) FIT TO A CONSISTENT SUBSET OF DATA, USE $\Delta\chi^2 = 1$
  $\Rightarrow$ RESULTS NOT CONSISTENT, UNCERTAINTY DOES NOT GROW AS DATASET DECREASES

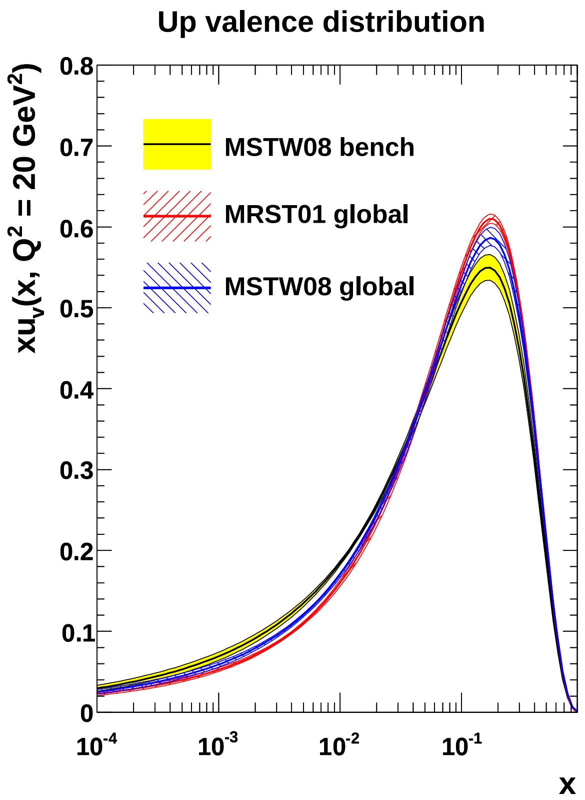


WHAT DETERMINES PDF UNCERTAINTIES?

THE PROBLEM OF BENCHMARK FITS (HERALHC 2005-2008)

- PERFORM A MRST (MRSTBENCH) FIT TO A CONSISTENT SUBSET OF DATA, USE $\Delta\chi^2 = 1$
  $\Rightarrow$ RESULTS NOT CONSISTENT, UNCERTAINTY DOES NOT GROW AS DATASET DECREASES
- ...BUT MRST WAS DONE WITH TOLERANCE 50: REPEAT WITH DYNAMICAL TOLERANCE (MSTW08BENCH)
- IMPROVEMENT, BUT PROBLEM NOT SOLVED
  $\Rightarrow$ MUST TUNE PARAMETRIZATION AND STATISTICAL TREATMENT TO DATASET

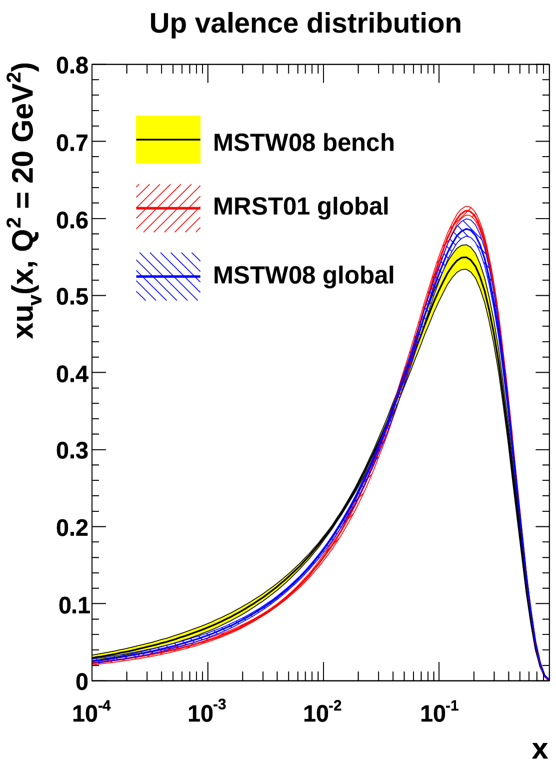




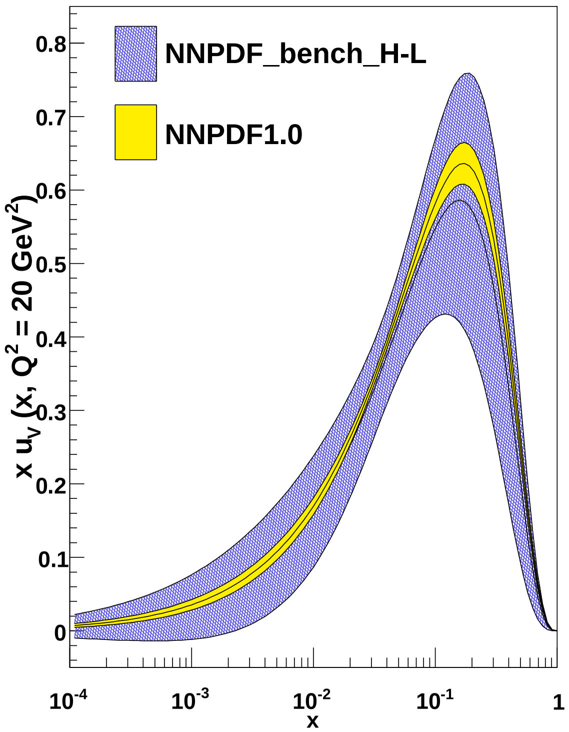
WHAT DETERMINES PDF UNCERTAINTIES?

THE NNPDF SOLUTION (HERALHC 2008)

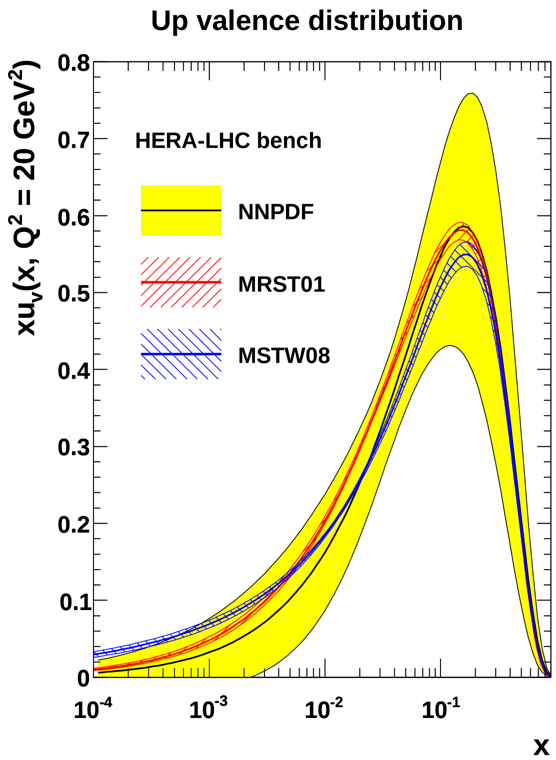
MRST/MSTW: BENCH VS REF



NNPDF: BENCH VS REF



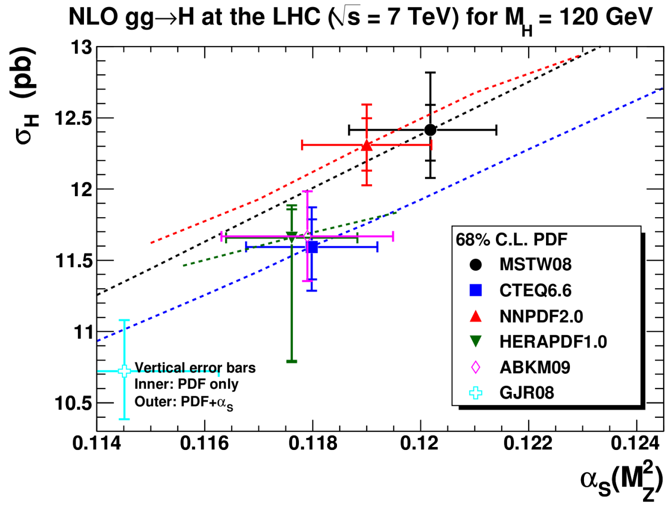
NNPDF BENCH VS MRST/MSTW BENCH



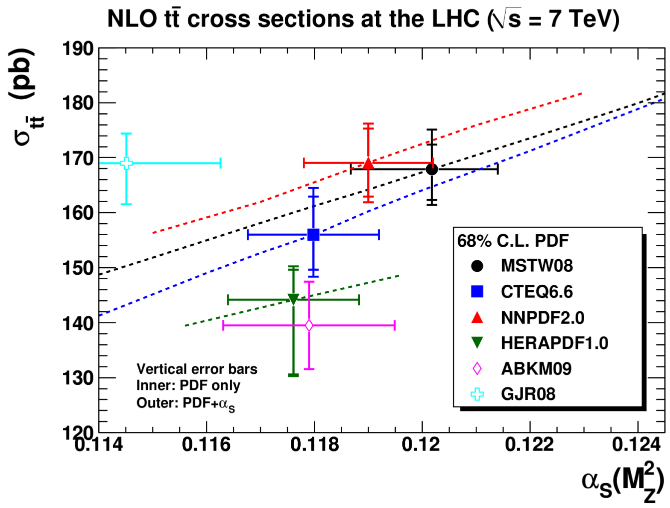
- SINGLE PARAMETRIZATION AND STAT. TREATMENT CAN ACCOMMODATE DIFFERENT DATASETS
- IMPACT OF DATA CAN BE STUDIED INDEPENDENT OF THEORETICAL FRAMEWORK

LHC STANDARD CANDLES: WHERE DO WE STAND?

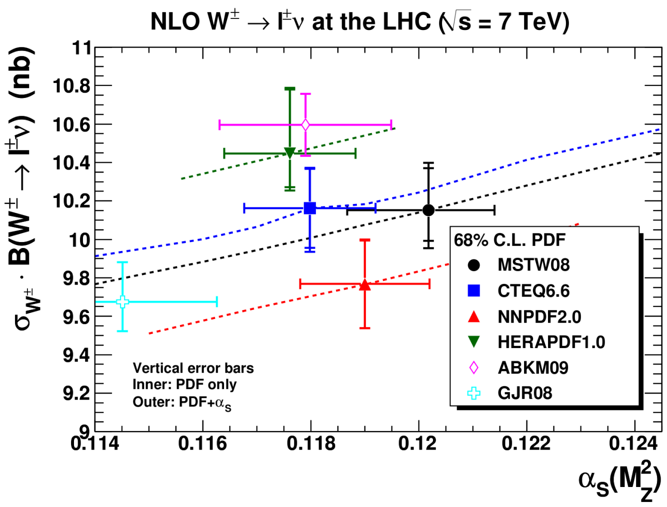
HIGGS



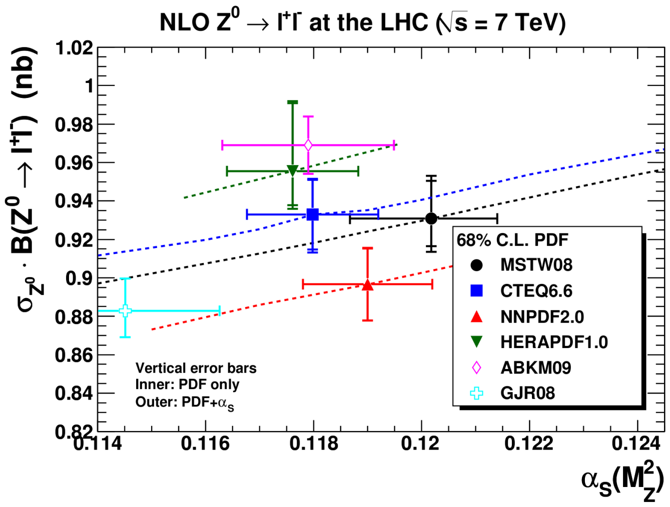
TOP



$W^{\pm}$



Z

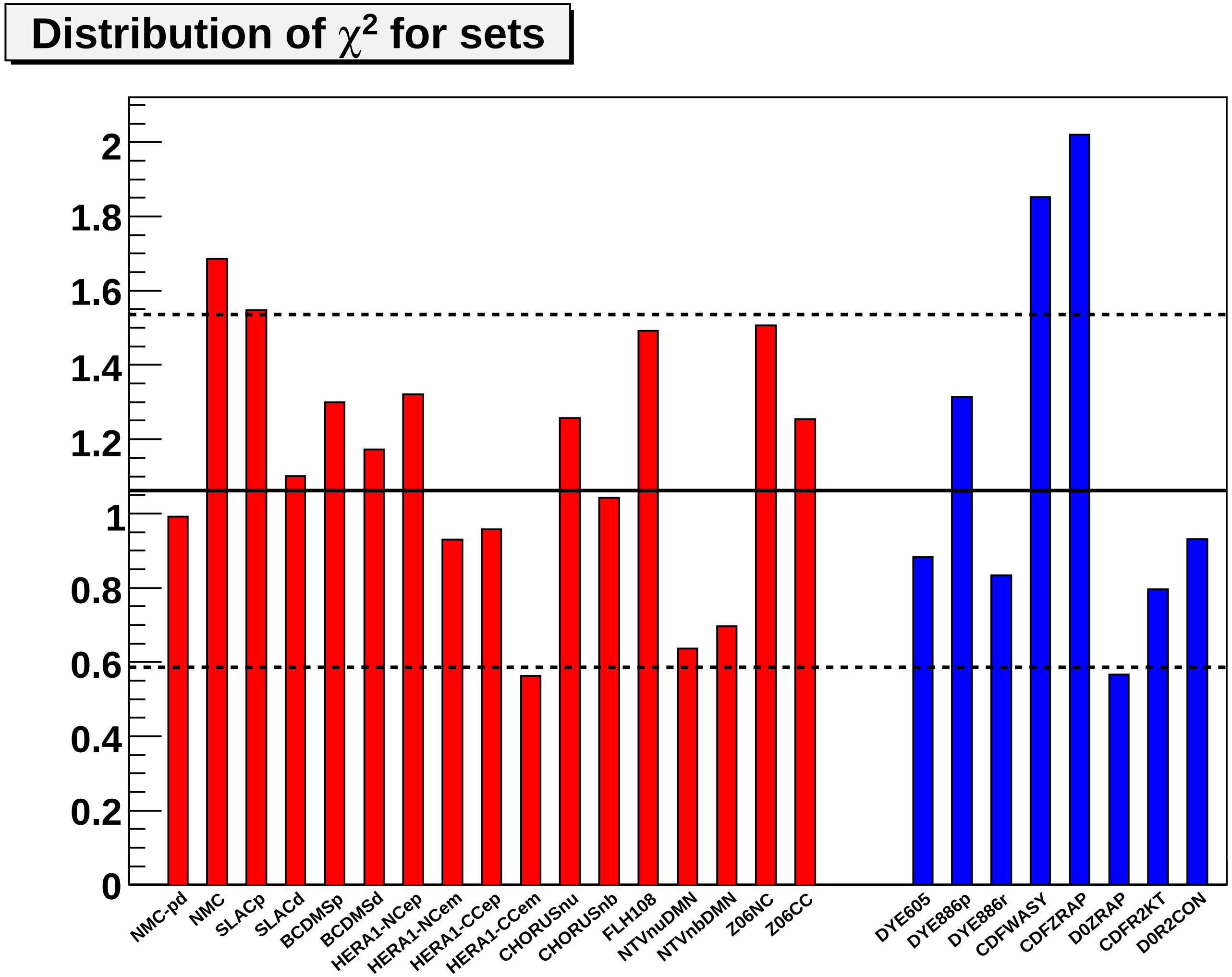


(G. Watt, 2010)

- GLOBAL FITS (MSTW, NNPDF, CTEQ) IN GOOD MUTUAL AGREEMENT (CENTRAL VALUE & UNCERTAINTY
- CHOICE OF α_s VALUE IMPORTANT

HOW WELL DOES IT WORK?

DIS DATA VS. HADRONIC DATA
NNPDF2.0

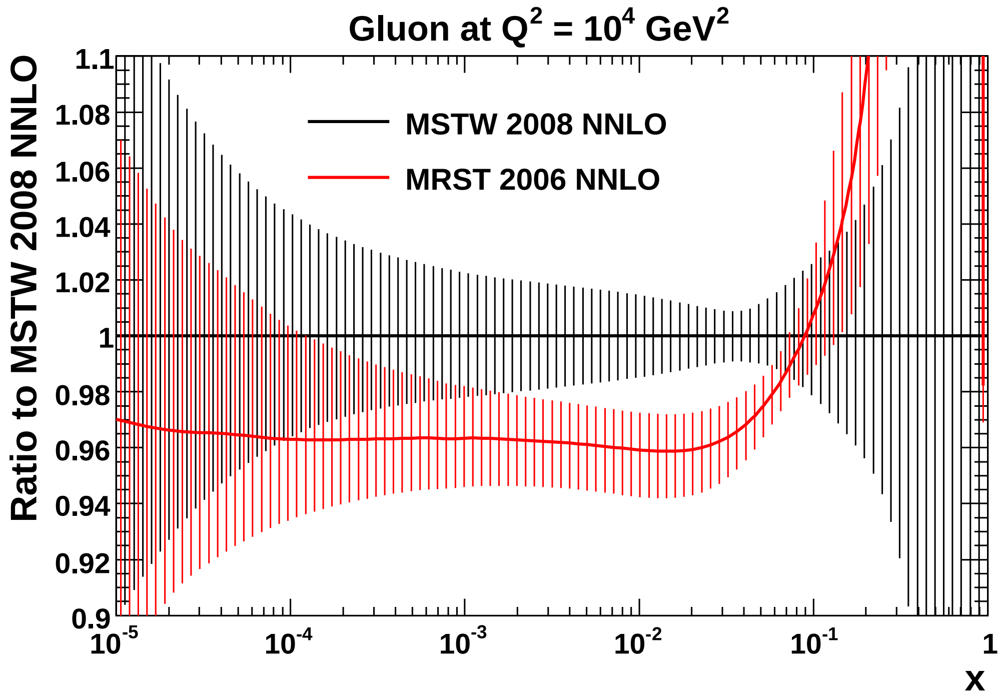


- NO OBVIOUS MUTUAL TENSION BETWEEN DIS AND HADRONIC DATA
- SOME INTERNAL DATA INCONSISTENCIES
  (NMC DIS DATA, CDF Z AND W RAPIDITY DISTRIBUTIONS)
- REASONABLE DISTRIBUTION OF χ^2 VALUES

WHAT DETERMINES PDF UNCERTAINTIES?

UNCERTAINTIES IN MSTW/CTEQ FITS OFTEN GO UP WHEN DATA ARE ADDED, BECAUSE OF THE NEED TO ADD PARAMETERS

Smaller high-x gluon (and slightly smaller α_S) results in larger small-x gluon – now shown at NNLO.



R. THORNE, HERALHC2008

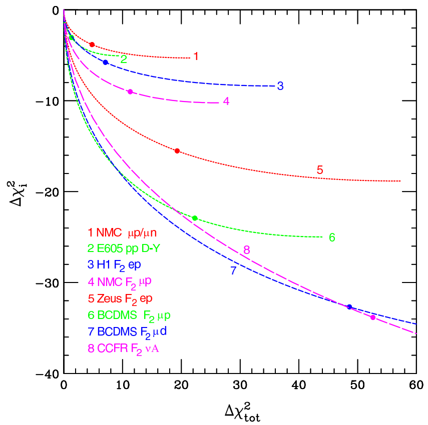
Larger small-x uncertainty due to extrat free parameter.

WHAT IS A ONE-σ UNCERTAINTY?

MSTW/CTEQ: THE SPREAD OF PDFS WITHIN AN ACCEPTABLE TOLERANCE

- STANDARD $\Delta\chi^2 = 1$ BANDS TOO NARROW $\Rightarrow$ LARGE DISCREPANCIES FOR INDIVIDUAL EXPERIMENTS

MINIMUM χ_i^2
VS GLOBAL χ^2



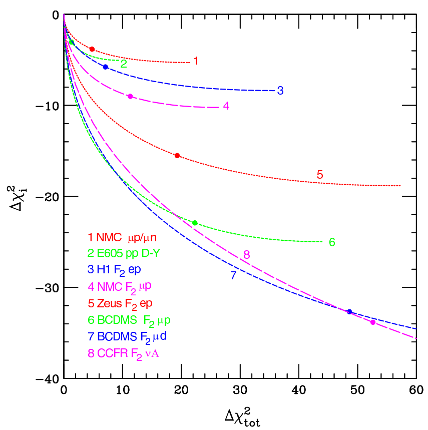
Collins, Pumplin
2001

WHAT IS A ONE-σ UNCERTAINTY?

MSTW/CTEQ: THE SPREAD OF PDFS WITHIN AN ACCEPTABLE TOLERANCE
CTEQ TOLERANCE CRITERION &

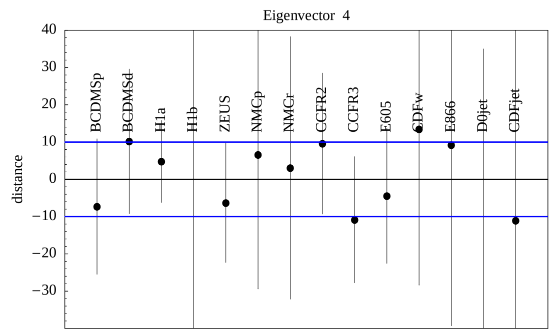
- STANDARD $\Delta\chi^2 = 1$ BANDS TOO NARROW $\Rightarrow$ LARGE DISCREPANCIES FOR INDIVIDUAL EXPERIMENTS
- TOLERANCE $\Rightarrow$ ENVELOPE OF UNCERTAINTIES OF EXPERIMENTS

MINIMUM χ_i^2 VS GLOBAL χ^2



Collins, Pumplin
2001

CTEQ TOLERANCE PLOT FOR 4TH EIGENVEC.



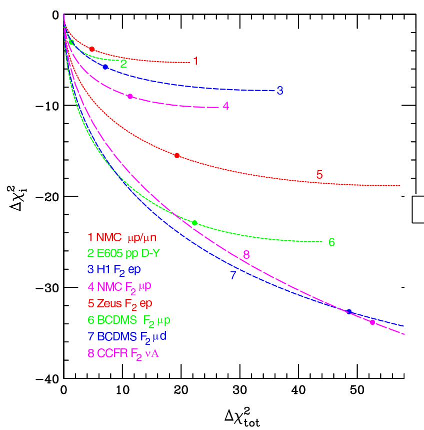
WHAT IS A ONE-σ UNCERTAINTY?

MSTW/CTEQ: THE SPREAD OF PDFS WITHIN AN ACCEPTABLE TOLERANCE

CTEQ TOLERANCE CRITERION & MSTW DYNAMICAL TOLERANCE

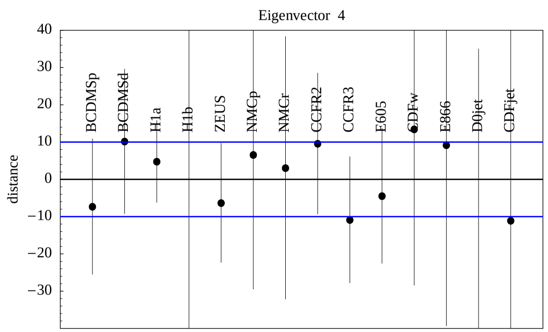
- STANDARD $\Delta\chi^2 = 1$ BANDS TOO NARROW $\Rightarrow$ LARGE DISCREPANCIES FOR INDIVIDUAL EXPERIMENTS
- TOLERANCE $\Rightarrow$ ENVELOPE OF UNCERTAINTIES OF EXPERIMENTS
- DYNAMICAL $\Rightarrow$ SEPARATELY DETERMINED FOR EACH HESSIAN EIGENVECTOR

MINIMUM χ_i^2 VS GLOBAL χ^2

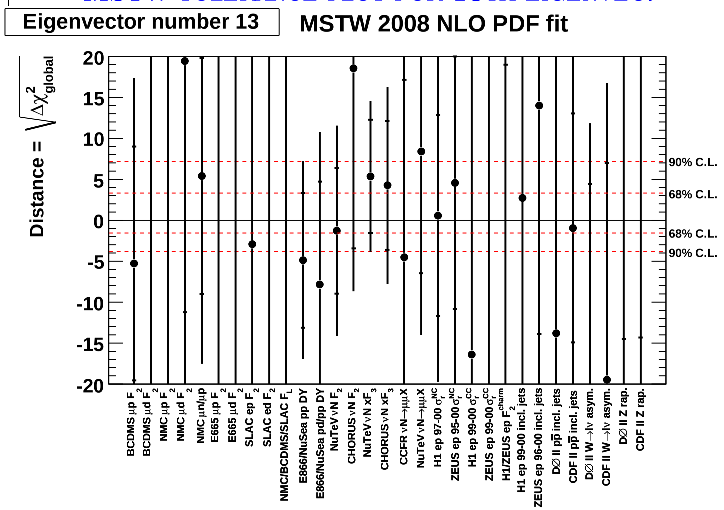


Collins, Pumplin 2001

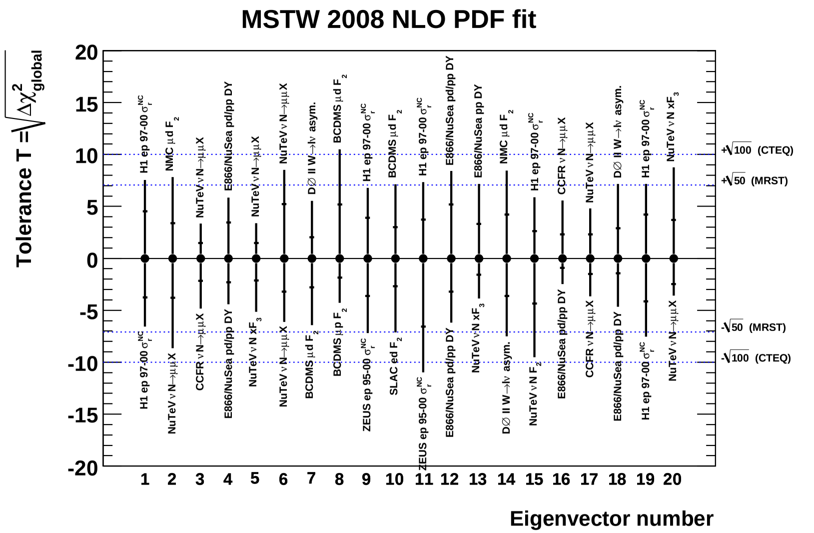
CTEQ TOLERANCE PLOT FOR 4TH EIGENVEC.



MSTW TOLERANCE PLOT FOR 13TH EIGENVEC.



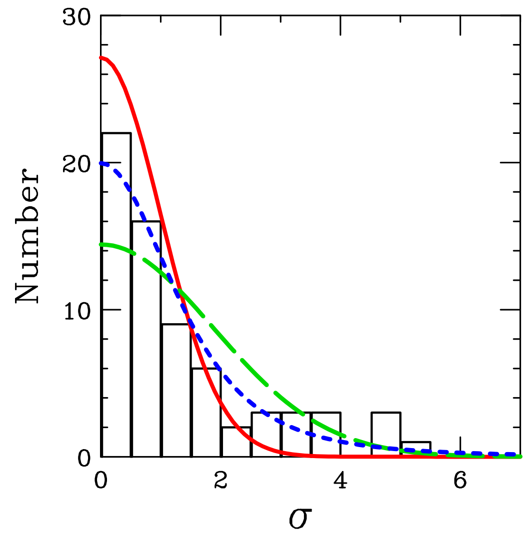
GLOBAL MSTW TOLERANCE



WHERE IS THE HESSIAN UNCERTAINTY COMING FROM?
WHY DOES ONE NEED LARGE TOLERANCES?

DATA INCOMPATIBILITY (Pumplin, 2009)

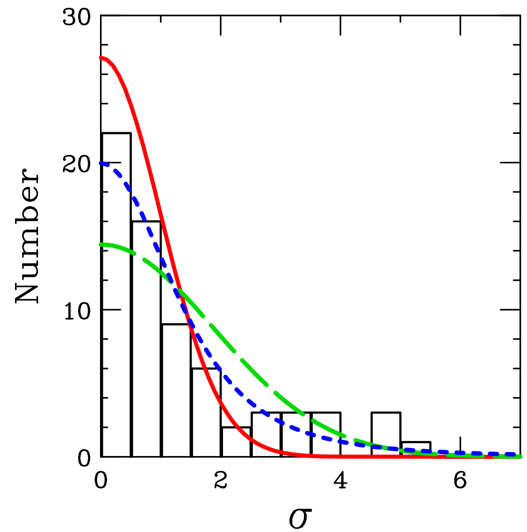
- CAN "REDIAGONALIZE": DIAGONALIZE SIMULTANEOUSLY χ^2 FOR TOTAL AND i–TH EXPT $\Rightarrow$ COMPATIBILITY OF EACH EXPT WITH GLOBAL FIT
- STUDY DISTRIBUTION OF DISCREPANCIES
- APPROX. GAUSSIAN WITH UNCERTAINTIES RESCALED BY 2 $\Rightarrow$ $\Delta\chi^2 \sim 10$ FOR 90%C.L.



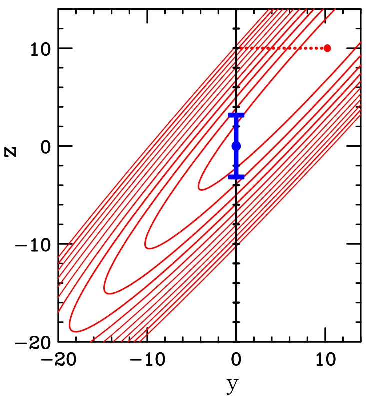
WHERE IS THE HESSIAN UNCERTAINTY COMING FROM?
WHY DOES ONE NEED LARGE TOLERANCES?

DATA INCOMPATIBILITY (Pumplin, 2009)

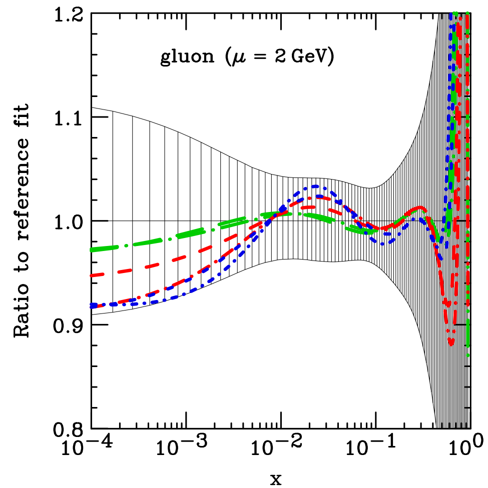
- CAN "REDIAGONALIZE": DIAGONALIZE SIMULTANEOUSLY χ^2 FOR TOTAL AND i–TH EXPT $\Rightarrow$ COMPATIBILITY OF EACH EXPT WITH GLOBAL FIT
- STUDY DISTRIBUTION OF DISCREPANCIES
- APPROX. GAUSSIAN WITH UNCERTAINTIES RESCALED BY 2 $\Rightarrow$ $\Delta\chi^2 \sim 10$ FOR 90%C.L.



FUNCTIONAL BIAS (Pumplin, 2009)



- IF PARM. NOT GENERAL ENOUGH, GLOBAL MIN. IS NOT TRUE MIN.
- ONE-σ VARIATION ABOUT FAKE MIN CORRESP. TO LARGE χ^2 VARIATION
- USE OF CHEBYSHEV POLYNOMIALS SUGGESTS "MOST GENERAL" PARM. WITHIN $\Delta\chi^2 = 10$ OF CTEQ6.6 PARM.



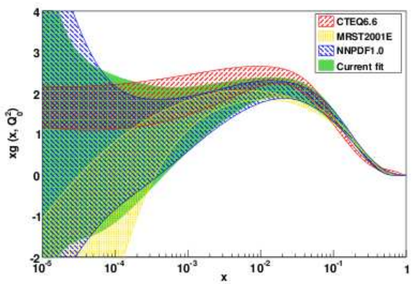
WHERE IS THE MONTE CARLO UNCERTAINTY COMING FROM?

FIT TO REPLICAS VS RANDOM SUBSET OF CENTRAL VAL.S

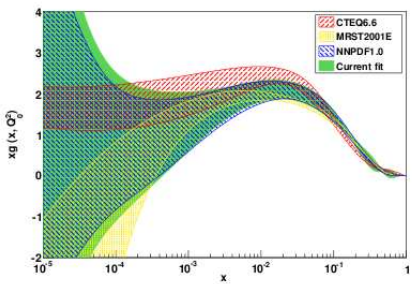
	REPLICAS	CENTRAL V.
χ^2	1.32	1.32
$\langle\chi^2\rangle_{\rm rep}$	2.79 ± 0.24	1.65 ± 0.20
$\langle\sigma^{\rm dat}\rangle$	0.039	0.035

GLUE

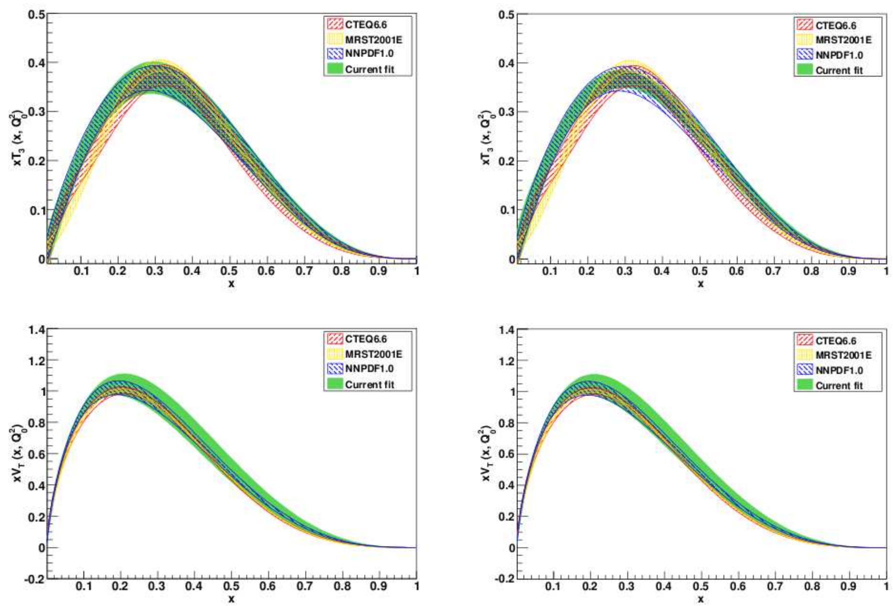
replicas



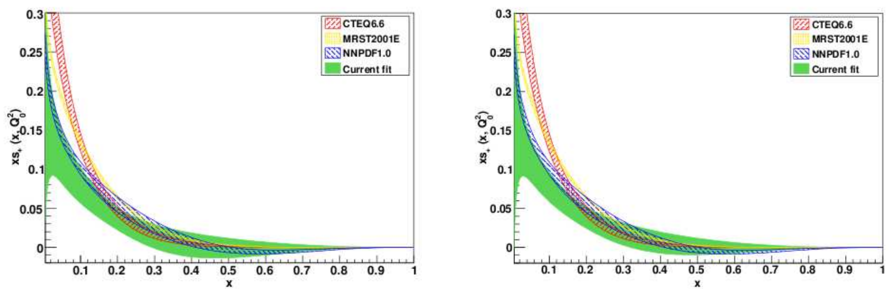
c. vals.



LIGHT QUARKS



STRANGE



- QUALITY OF FIT &PDFS UNCHANGED
- REDUCTION OF $\langle\chi^2\rangle_{\rm rep}$ BY FACTOR $\sim 2 \Rightarrow$ FLUCTUATIONS ABOUT TRUE VALUE HALVED
- UNCERTAINTY ON DATA ONLY REDUCED BY 1.1 $\Rightarrow$ EXPT. UNCERTAINTIES UNDERESTIMATED OR UNDERLYING INCOMPRESSIBLE UNCERTAINTY

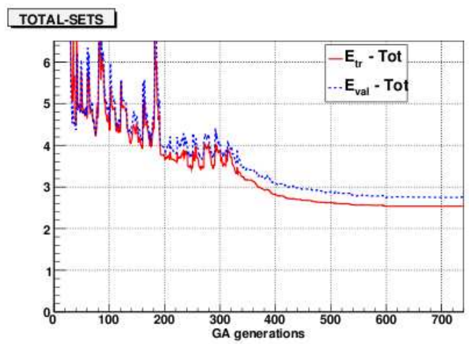
WHERE IS THE MONTE CARLO UNCERTAINTY COMING FROM?

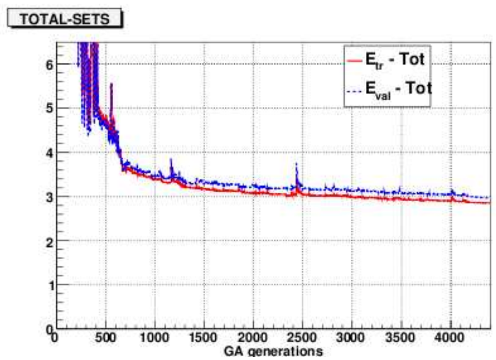
CENTRAL VALUES: VARYING PARTITION VS FIXED PARTITION

	REPLICAS	CENTRAL VALUE	FIXED PARTITION
χ^2	1.32	1.32	$\sim$1.3
$\langle\chi^2\rangle_{\rm rep}$	2.79 ± 0.24	1.65 ± 0.20	$\sim 1.6 \pm 0.2$
$\langle\sigma^{\rm dat}\rangle$	0.039	0.035	$\sim$0.03

fixed partition results obtained averaging over 5 different choices of partition (100 replicas each); more partitions needed for accurate results

- QUALITY OF FIT UNCHANGED
- $\langle\chi^2\rangle_{\rm rep}$ UNCHANGED $\Rightarrow$ CENTRAL FIT UNCHANGED
- UNCERTAINTY ON PREDICTION (I.E. ON PDFS) REDUCED

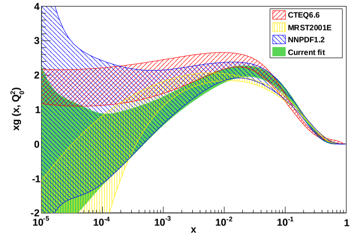




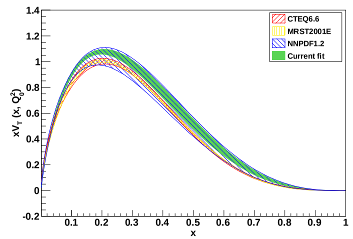
FUNCTIONAL UNCERTAINTY

- MORE THAN HALF OF UNCERTAINTY DUE TO "FUNCTIONAL FORM": $\langle\sigma^{\rm dat}\rangle =\sim 0.03$ SMALLER FOR HERA DATA
- REMAINING UNCERTAINTY ROUGHLY SCALES WITH DATA UNCERTAINTY: $\langle\sigma^{\rm dat}\rangle =\sim 0.005$ CENT.; $\langle\sigma^{\rm dat}\rangle =\sim 0.009$ REP.

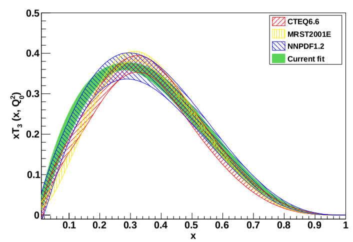
GLUE



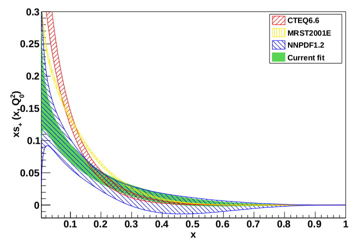
VALENCE



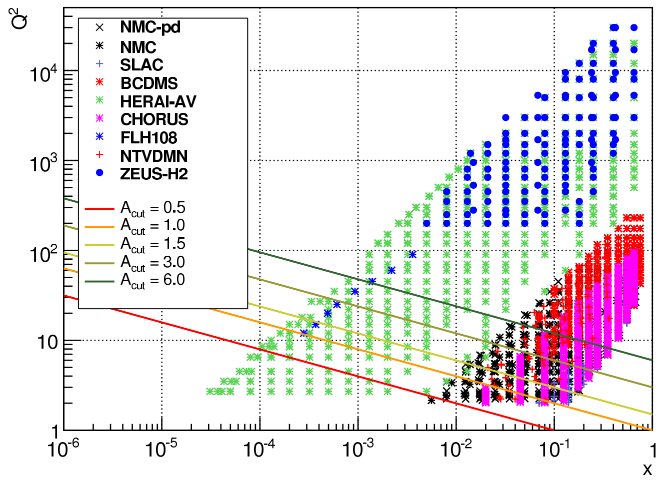
TRIPLET



STRANGE



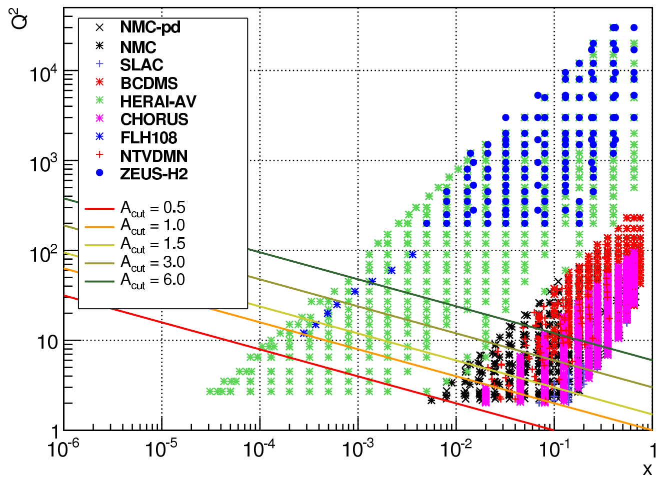
BEYOND DGLAP: TESTING FOR DEVIATIONS



IDEA: (Gélis, 2008, $\Rightarrow$ Caola, s.f. ,Rojo 2010)

- CUT OUT DATA IN THE "DANGEROUS" (SMALL τ) REGION
- DETERMINE PDFs IN THE "SAFE" (LARGE x AND Q^2) REGION
- EVOLVE BACKWARDS AND COMPARE TO DATA

BEYOND DGLAP: TESTING FOR DEVIATIONS

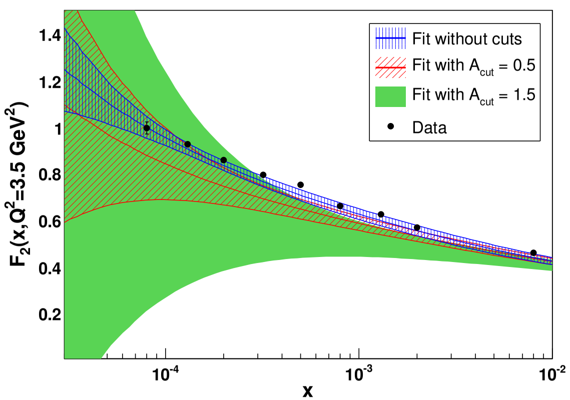


IDEA: (Gélis, 2008, ⇒ Caola, s.f. ,Rojo 2010)

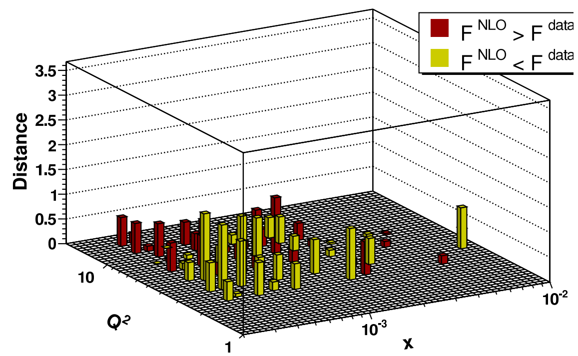
- CUT OUT DATA IN THE "DANGEROUS" (SMALL τ) REGION
- DETERMINE PDFs IN THE "SAFE" (LARGE x AND Q^2) REGION
- EVOLVE BACKWARDS AND COMPARE TO DATA

OLD HERA DATA

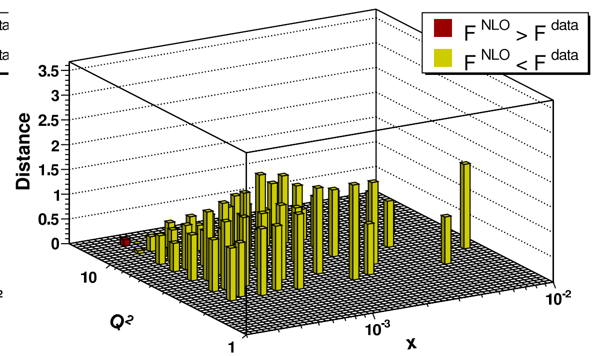
BACKWARD EV. VS DATA



DAT/TH DIST: NO CUT



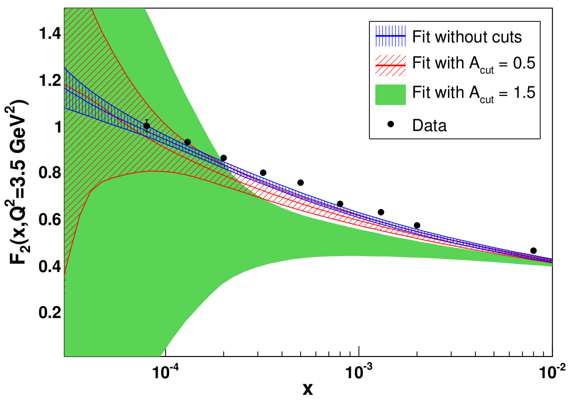
DAT/TH DIST: CUT



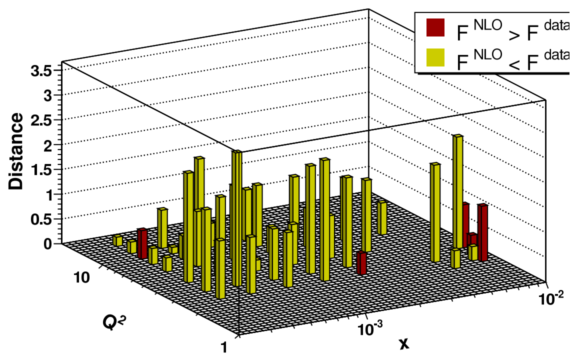
- BACKWARD EVOLVED FIT LIES SYSTEMATICALLY BELOW DATA
- DATA AT LOW x AND Q^2 SHOW LESS EVOLUTION THAN PREDICTED BY NLO DGLAP
- IF LOW x AND Q^2 DATA INCLUDED, THE FIT MANAGES TO COMPENSATE BY READJUSTING THE PDFs

NEW (COMBINED) HERA DATA

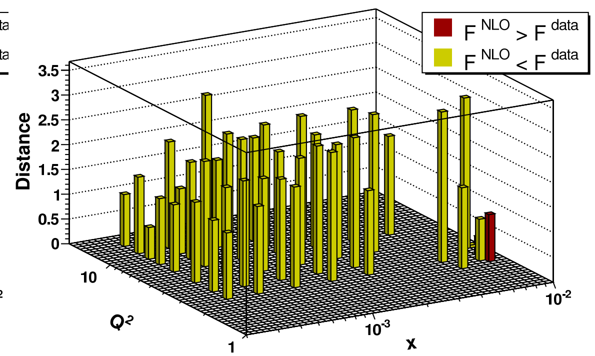
BACKWARD EV. VS DATA



DAT/TH DIST: NO CUT



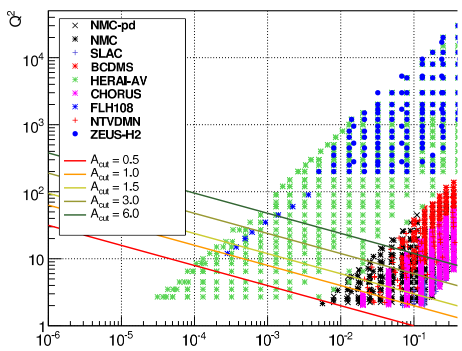
DAT/TH DIST: CUT

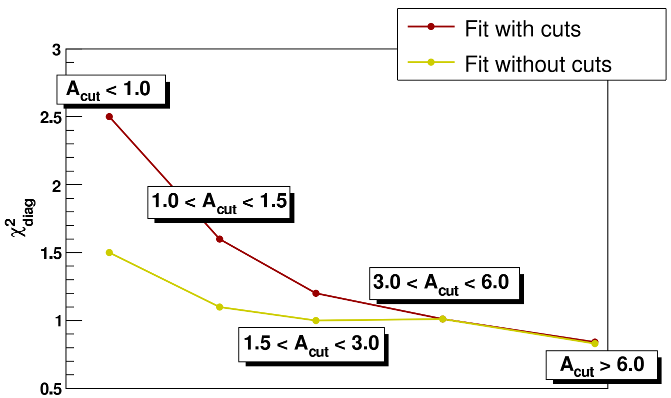


- DATA AT LOW x AND Q^2 SHOW LESS EVOLUTION THAN PREDICTED BY NLO DGLAP
- BACKWARD EVOLVED FIT LIES SYSTEMATICALLY BELOW DATA
- WITH MORE PRECISE DATA, THE FIT NO LONGER MANAGES TO COMPENSATE BY READJUSTING THE PDFS: EVEN FULL FIT LIES BELOW DATA

DETERIORATION IN FIT QUALITY:

χ^2 VS τ SLICES





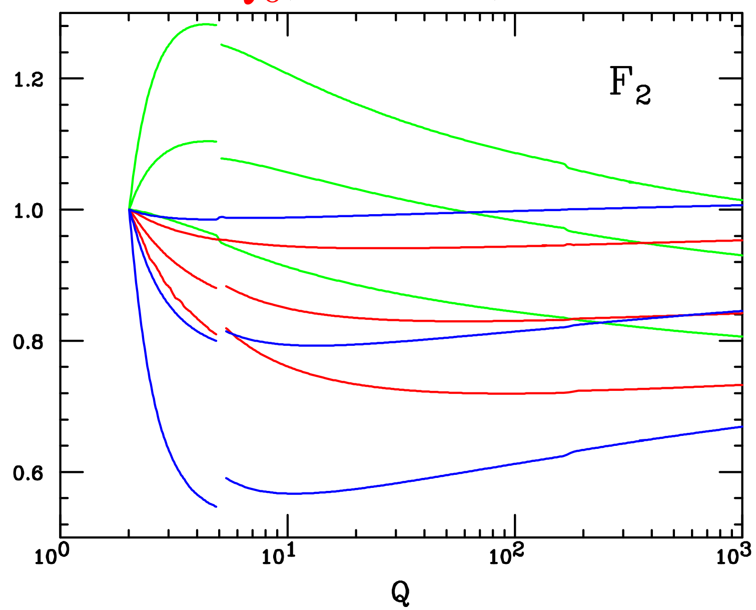
- QUALITY OF UNCUT FIT DETERIORATES IN LOW τ REGIONS
- QUALITY OF CUT FIT INCREASINGLY POOR AS τ DECREASES

DEVIATIONS FROM DGLAP:

WHERE DO THEY COME FROM?

- OBSERVED DEVIATION FROM DGLAP CANNOT BE DUE TO MISSING NNLO TERMS: DATA EVOLVE LESS THAN NLO WHILE NNLO EVOLVES MORE THAN NLO
- PERTURBATIVE RESUMMATION HAS THE RIGHT SIGN AND ROUGH SIZE
- SATURATION OR MORE GENERALLY HIGHER TWIST (POWER SUPPRESSED) EFFECTS MIGHT ALSO REDUCE EVOLUTION
- NOTE: SOUGHT-FOR EVIDENCE IS SUPPRESSED SCALE (Q^2) DEPENDENCE, REGARDLESS OF THE x GROWTH

RES. Q_0, RES. $\overline{\text{MS}}$, NNLO



(ABF, 08)

DEVIATIONS FROM DGLAP:
PHENOMENOLOGICAL IMPLICATIONS

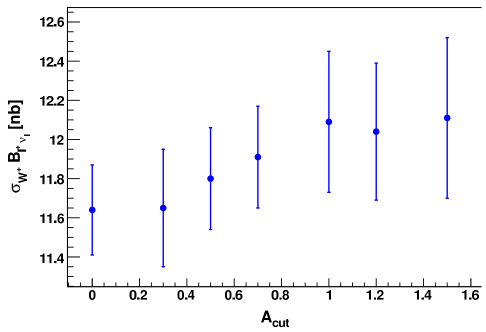
LHC STANDARD CANDLES (14 TEV)

- REMOVING DATA FROM DANGEROUS REGION LEADS TO INCREASED PDF UNCERTAINTIES
- IMPACT ON LHC STANDARD CANDLES MODERATE
- COULD HAVE SOME EFFECT ON LESS INCLUSIVE OBSERVABLES

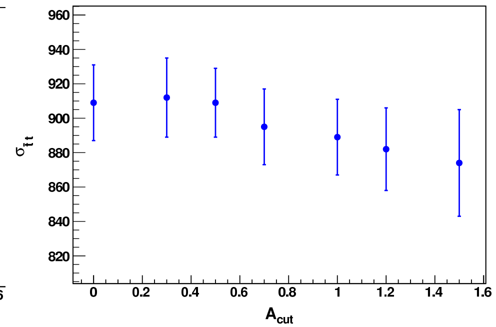
(Caola, s.f., Rojo, 10)

DEPENDENCE ON CUT

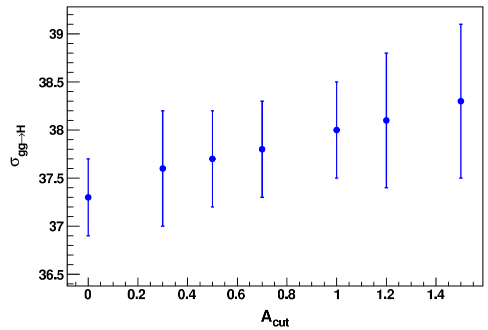
W



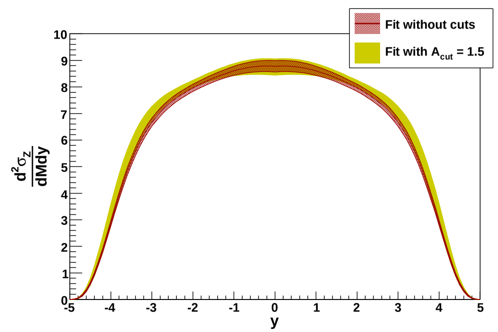
TOP



HIGGS (120 GEV)



Z RAP. DISTN.



DETERMINATION OF α_s

- α_s FROM SCALING VIOLATION MAY BE BIASED DOWNWARDS (WEAKER EVOLUTION $\to$ ARTIFICIALLY SMALLER α_s)
- BIAS STRONGER AT NNLO
- MSTW (2010) SEE A CHANGE OF 2 SIGMA FROM NLO TO NNLO $\alpha_s = 0.1202^{+0.0012}_{-0.0015}$ TO $\alpha_s = 0.1171 \pm 0.0014$