THE NNPDF APPROACH TO PARTON DETERMINATION: IDEAS AND RESULTS

STEFANO FORTE
UNIVERSITÀ DI MILANO & INFN

UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI FISICA

Istituto Nazionale di Fisica Nucleare

HESI10 WORKSHOP

KYOTO, AUGUST 26, 2010

PROLOGUE: THE DISCOVERY OF THE W

LAST TIME WE LOOKED FOR "NEW" PHYSICS AT A HADRON COLLIDER

THEORETICAL PREDICTION...

42 G. Altarelli et al. / Vector boson production

TABLE 2
Values (in nb) of the total cross sections for $W^\pm$ and Z^0 production

$\sqrt{S}$ (GeV)	$W^+ + W^-$ GHR	$W^+ + W^-$ DO1	$W^+ + W^-$ DO2	Z^0 GHR	Z^0 DO1	Z^0 DO2	$\frac{\sigma(W^+ + W^-)}{\sigma(Z^0)}$ GHR	$\frac{\sigma(W^+ + W^-)}{\sigma(Z^0)}$ DO1	$\frac{\sigma(W^+ + W^-)}{\sigma(Z^0)}$ DO2
540	4.2	4.3	4.1	1.3	1.3	1.2	3.1	3.4	3.5
700	6.2	6.3	6.1	2.0	1.9	1.8	3.1	3.3	3.4
1000	9.5	9.5	9.6	3.1	3.0	2.9	3.1	3.2	3.3
1300	12.5	12.5	12.9	4.0	3.9	3.9	3.1	3.2	3.3
1600	15.5	15.6	16.5	5.0	4.8	5.0	3.1	3.2	3.3

ALTARELLI, ELLIS, GRECO, MARTINELLI, 1984

...AND EXPERIMENTAL DISCOVERY

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN-EP/85-108
11 July 1985

W PRODUCTION PROPERTIES AT THE CERN SPS COLLIDER

UA1 Collaboration, CERN, Geneva, Switzerland

Aachen[1] - Amsterdam (NIKHEF)[2] - Annecy (LAPP)[3] - Birmingham[4] - CERN[5] - Harvard[6] - Helsinki[7] - Kiel[8] - London (Imperial College[9] and Queen Mary College[10]) - Padua[11] - Paris (Coll. de France)[12] - Riverside[13] - Rome[14] - Rutherford Appleton Lab.[15] - Saclay (CEN)[16] - Victoria[17] - Vienna[18] - Wisconsin[19] Collaboration

The corresponding experimental result for the 1984 data at $\sqrt{s}$ = 630 GeV is

$$(\sigma \cdot B)_W = 0.63 \pm 0.05\ (\pm 0.09)\ \text{nb}.$$

This is in agreement with the theoretical expectation [14] of $0.47^{+0.14}_{-0.08}$ nb. We note that the 15%

- AGREEMENT AND UNCERTAINTIES AT 20% CONSIDERED TO BE SATISFACTORY
- RESULTS FROM DIFFERENT PDF SETS DIFFER BY AT LEAST 5%
- NO WAY TO ESTIMATE PDF UNCERTAINTIES

PDFs: THEN AND NOW

PDFs IN 1984

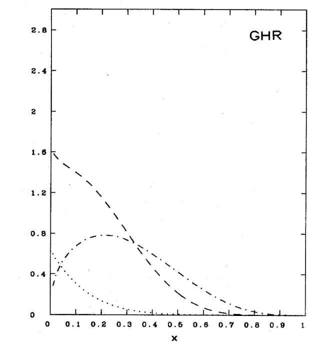


FIG. 25. Parton distributions of Glück, Hoffmann, and Reya (1982), at $Q^2=5$ GeV^2: valence quark distribution $x[u_v(x)+d_v(x)]$ (dotted-dashed line), $xG(x)$ (dashed line), and q_v (dotted line).

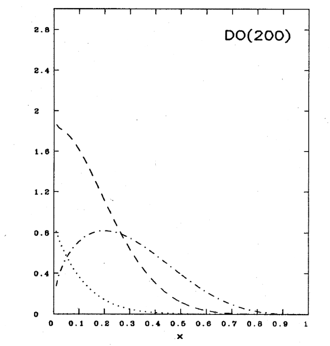


FIG. 27. "Soft-gluon" ($\Lambda=200$ MeV) parton distributions of Duke and Owens (1984) at $Q^2=5$ GeV^2: valence quark distribution $x[u_v(x)+d_v(x)]$ (dotted-dashed line), $xG(x)$ (dashed line), and $q_v(x)$ (dotted line).

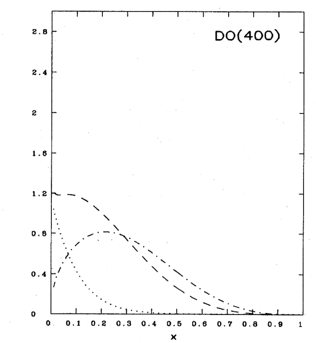
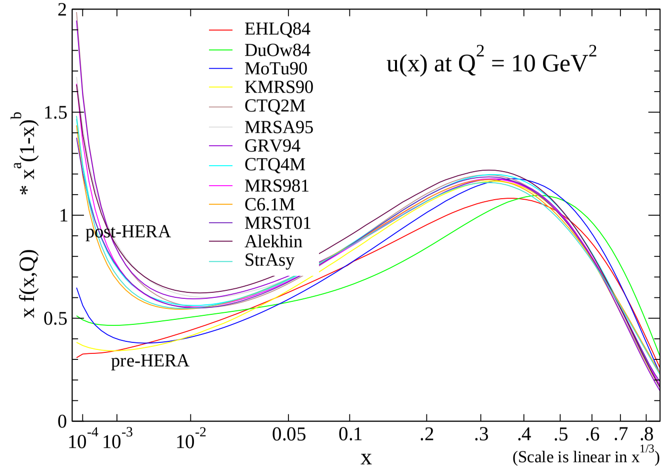


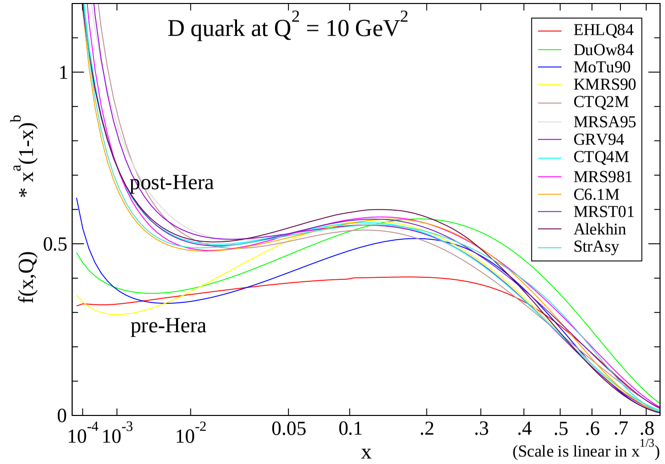
FIG. 26. "Hard-gluon" ($\Lambda=400$ MeV) parton distributions of Duke and Owens (1984) at $Q^2=5$ GeV^2: valence quark distribution $x[u_v(x)+d_v(x)]$ (dotted-dashed line), $xG(x)$ (dashed line), and $q_v(x)$ (dotted line).

Rev. Mod. Phys., Vol. 56, No. 4, October 1984

GHR VS DUKE-OWENS

20 YEARS OF VALENCE PDFs





W.-K. TUNG, 2004

- HADRON COLLIDERS CIRCA 1985 $\Rightarrow$ QUALITATIVE QCD: DISCOVERY PHYSICS
- DIS AT NMC AND HERA 1995-2005 $\Rightarrow$ QUANTITATIVE QCD: PRECISION PHYSICS
- HADRON COLLIDERS CIRCA 2010 $\Rightarrow$ PRECISION QCD $\leftrightarrow$ NEW PHYSICS

SUMMARY

- THE STATE OF THE ART

  PDFS FOR THE LHC

- IDEAS

  THE NNPDF APPROACH

- ISSUES AND SOLUTIONS

  THE PROBLEM OF PDF UNCERTAINTIES

- RESULTS

  NNPDF2.0, THE FIRST GLOBAL NLO PDF SET

- APPLICATIONS

  SENSITIVITY TO DISCOVERY

- PROGRESS

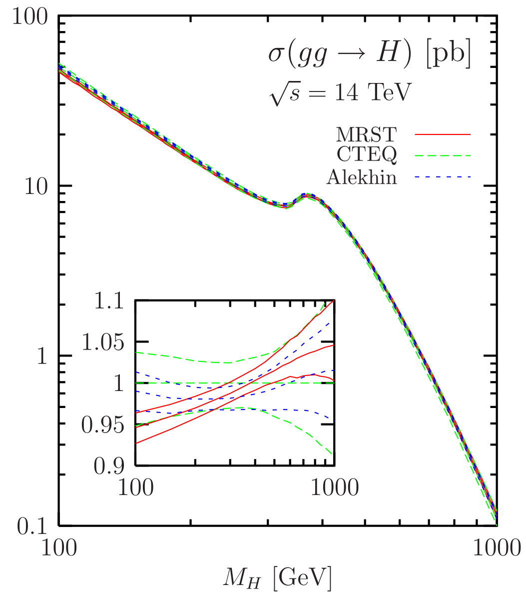
  THEORETICAL UNCERTAINTIES

THE STATE OF THE ART

PDFS: TOWARDS LHC PHYSICS

THE HIGGS CROSS SECTION

FIRST PDFS WITH UNCERTAINTIES



(Djouadi, Ferrag, 2004)

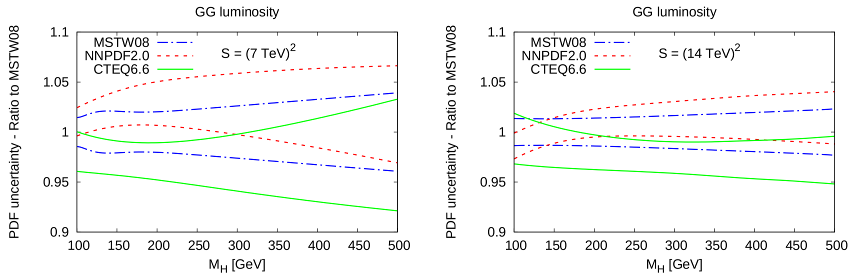
PDFS WITH ERROR
(2002-2003)

CTEQ, MRST (global); Alekhin (DIS)

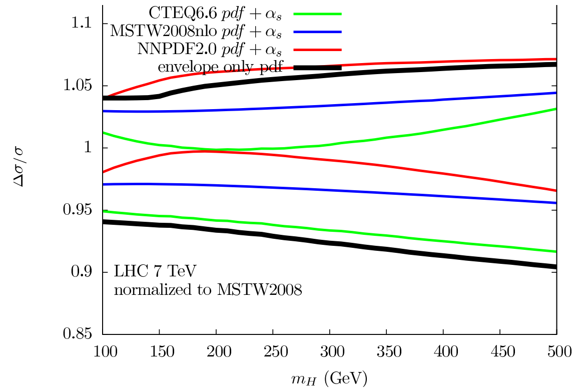
- WIDELY DIFFERENT UNCERTAINTY ESTIMATES
- UNSATISFACTORY AGREEMENT WITHIN UNCERTAINTIES
- $\Delta\chi^2 = 100$ (CTEQ); 50 (MRST); 1 (ALEKHIN)

CURRENT GLOBAL PDF SETS

gluon luminosities



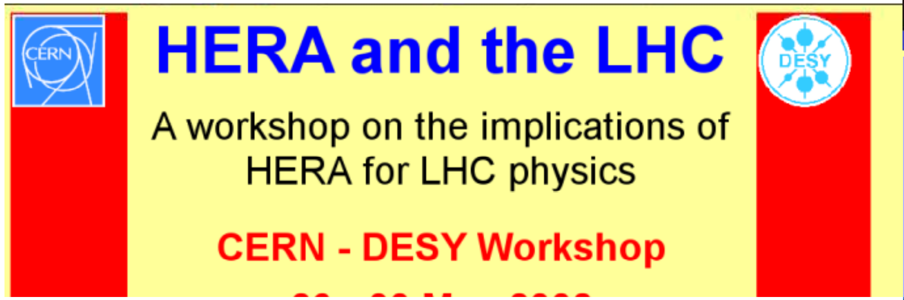
gg→H cross section



(Demartin et al., 2010)

- THREE GLOBAL (DIS+HADRONIC) PDF SETS AVAILABLE
- REASONABLE AGREEMENT OF CENTRAL VALUES & UNCERTAINTIES

AN ONGOING EFFORT

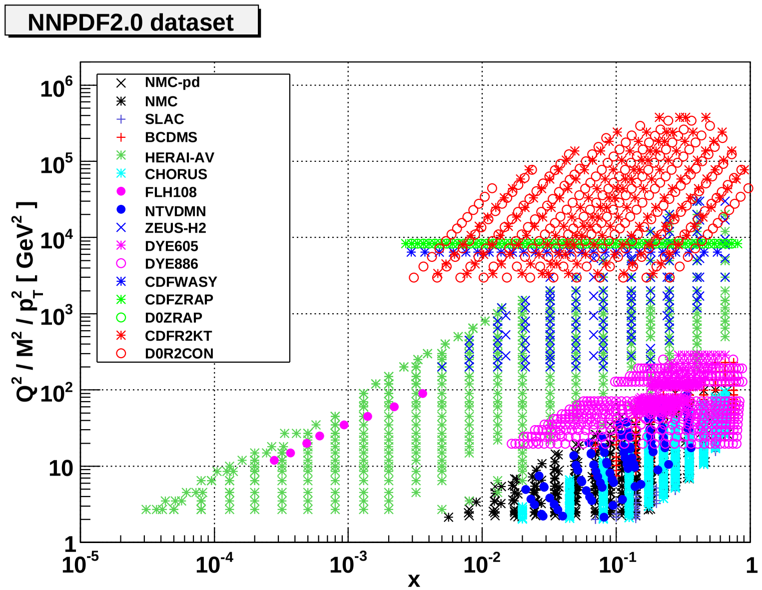


- HERALHC (2004-2008): A WORKSHOP TO TRANSFER KNOWN-HOW FROM HERA TO THE LHC COMMUNITY
  - two sessions (2004-2005; 2006-2007), four plenary meetings, five mid-term and working group meetings
  - two CERN/DESY reports (yellow books): **HEP-PH/0601012-HEP-PH/0601013**; **ARXIV:0903.3861**.
- PDF4LHC (2008 - ONGOING): A PERMANENT WORKING GROUP TO PROVIDE GUIDANCE ON PDF TO LHC EXPERIMENTS AND PHENOMENOLOGY
  - quarterly meetings, 10 since inception in February 2008
  - website `http://www.hep.ucl.ac.uk/pdf4lhc/` and wiki `https://wiki.terascale.de/index.php?title=PDF4LHC_WIKI` resources available
  - fruitful interactions with the LHC Higgs Cross Section WG

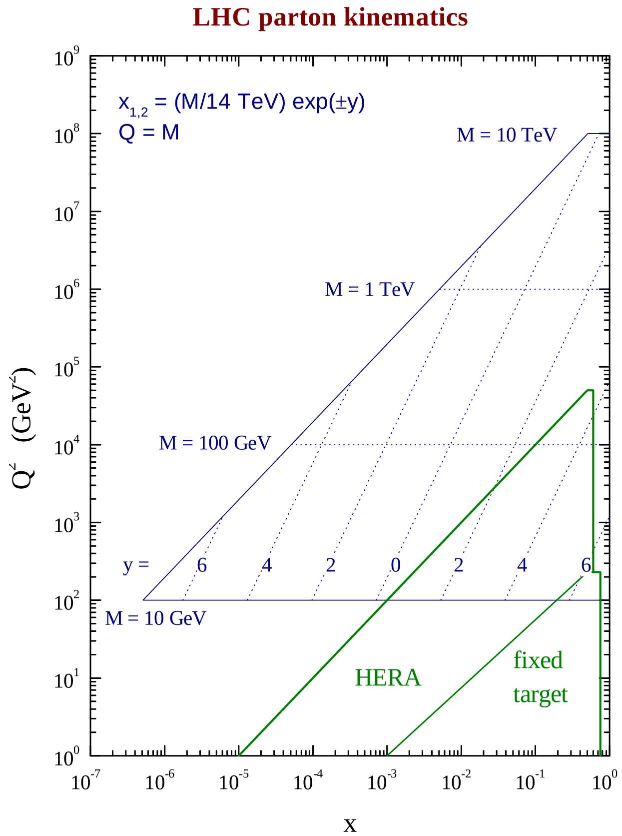
CURRENT PDF SETS

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

data for a global fit (NNPDF2.0)



LHC kinematics



- 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,...)

FEATURES, TRADEOFFS AND CHOICES

- DATASET:CTEQ, MSTW, NNPDF FIXED TARGET AND COLLIDER, eP AND $\bar{P}P$ DATA; ABKM, GJR GLOBAL DIS+ FIXET-TARGET DY; HERAPDF: HERA ONLY
- STATISTICAL TREATMENT: CTEQ HESSIAN WITH TOLERANCE; MSTW HESSIAN WITH DYNAMICAL TOLERANCE; HERAPDF, ABKM, GJR, STANDARD HESSIAN; NNPDF MONTE CARLO (ALSO STUDIED BY HERAPDF)
- PARTON PARAMETRIZATION:CTEQ, MSTW, HERAPDF $x^\alpha(1-x)^\beta\times$ POLYNOMIALS; GJR: DITTO + VALENCELIKE ASSUMPTION; NNPDF NEURAL NETS; CHEBYSHEV POLYNOMIALS STUDIED BY HERAPDF;
- HEAVY QUARKS:CTEQ: GM-VFN (SACOT-χ SCHEME); MSTW: GM-VFN (ACOT+TR' SCHEME); NNPDF: ZM-VFN, ORELIM: GM-VFN (FONLL-A SCHEME); ABKM: FFN ($N_f = 3, 4$ MATCHED WITH BMSN SCHEME); GJR: FFN ($N_f = 3$)
- PERTURBATIVE ORDER:CTEQ: NLO, BUT DY LO WITH K-FACTORS; MSTW: NNLO, BUT DY LO WITH K-FACTORS; NNPDF FULL NLO; ABKM, GJR: FULL NNLO
- α_s VALUE:CTEQ, NNPDF: EXTERNAL PARAMETER, SEVERAL VALUES AVAILABLE; MSTW: FITTED, BUT ALSO VARIABLE AS EXT.PARAMETER; ABKM, GJR: FITTED, NOT VARIABLE AS EXT. PARAMETER

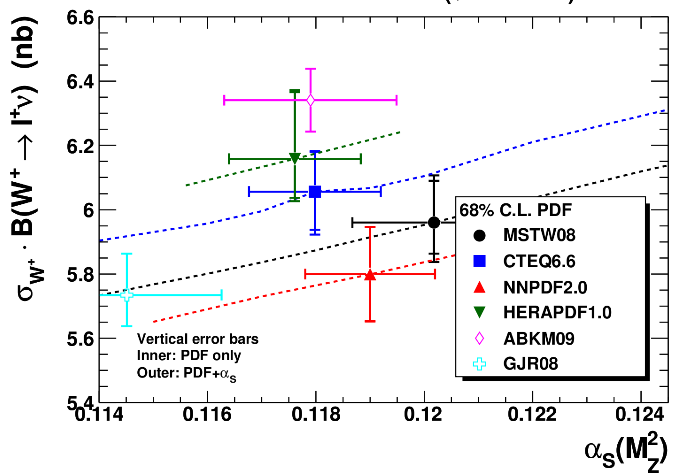
THE PDF4LHC RECOMMENDATION

- AT NLO, ENVELOPE OF CTEQ, MSTW, NNPDF
- AT NNLO, MSTW WITH UNCERTAINTY RESCALED MY MSTWNLO/NLO ENVELOPE

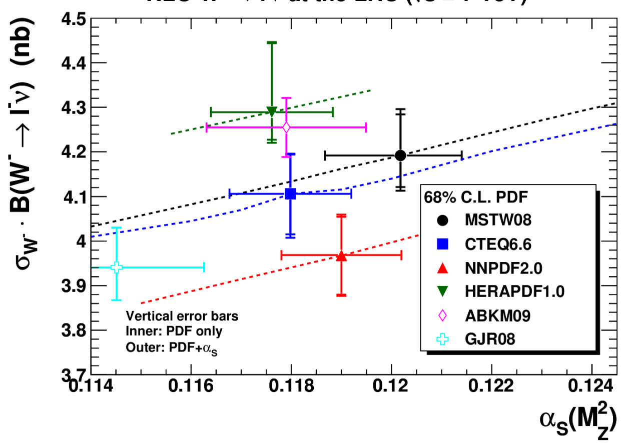
WHERE DO WE STAND?

LHC STANDARD CANDLES

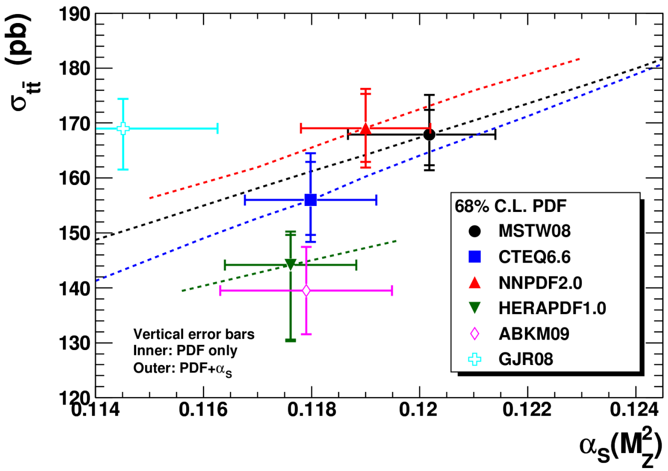
W^+



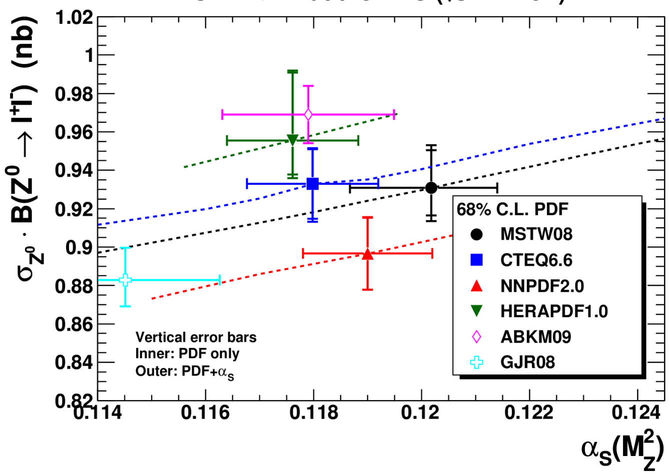
W^-



TOP



Z



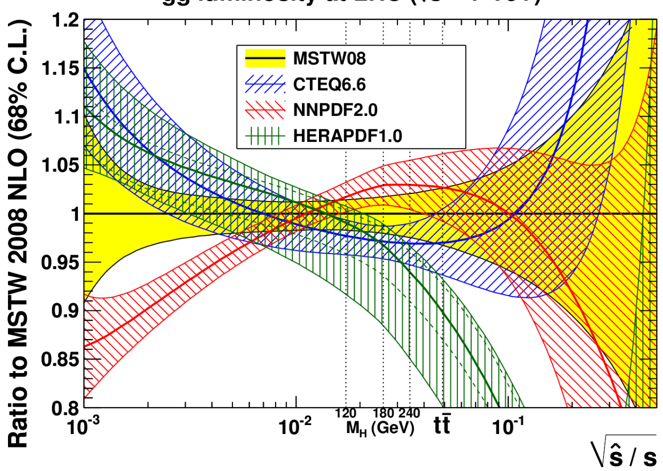
(G. Watt, 2010)

- GLOBAL FITS IN GOOD MUTUAL AGREEMENT
- CHOICE OF α_s VALUE IMPORTANT

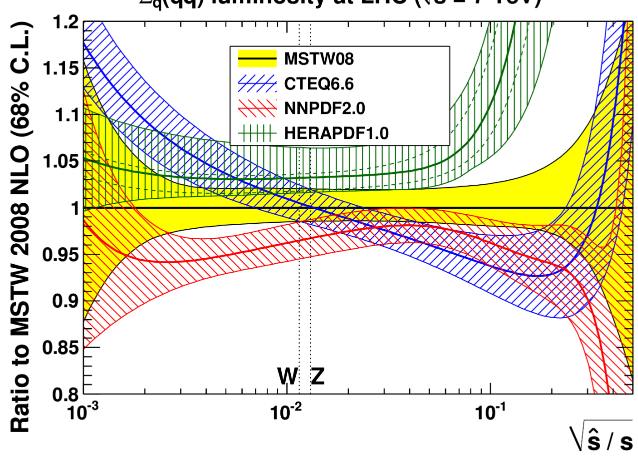
WHERE DO WE STAND?

PARTON LUMINOSITIES

GLUON-GLUON

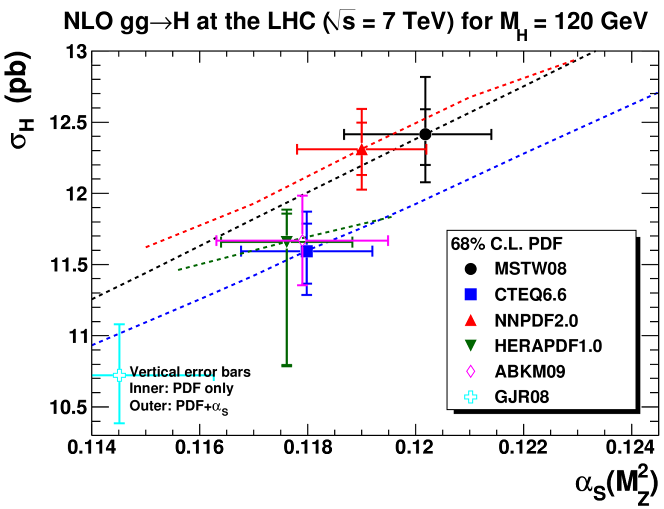


QUARK-QUARK

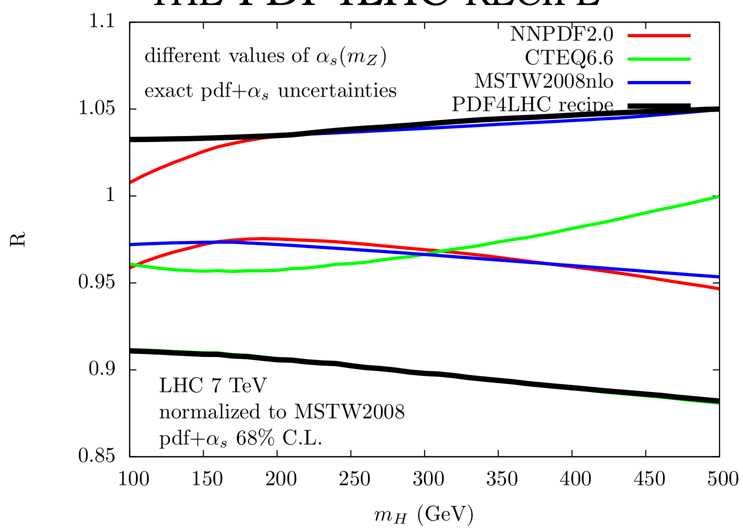


THE HIGGS CROSS SECTION

$m_H = 120$ GeV: DIFFERENT GROUPS



THE PDF4LHC RECIPE



(G. Watt, 2010)

(A. Vicini, 2010)

ENVELOPE TAKES CARE OF POORLY UNDERSTOOD DISAGREEMENTS

THE NNPDF APPROACH

THE HESSIAN "STANDARD" APPROACH

(MSTW, CTEQ):

FUNCTIONAL PARTON FITTING

- CHOOSE A FIXED FUNCTIONAL FORM:
  - MSTW: 20 PARMS.

    $$xq(x,Q_0^2) = A(1-x)^{\eta}(1+\epsilon x^{0.5}+\gamma x)x^{\delta}, (5\ \text{indep. fns.}); x[\bar{u}-\bar{d}](x,Q_0^2) = A(1-x)^{\eta}(1+\gamma x+\delta x^2)x^{\delta};$$

    $$x[s-\bar{s}](x,Q_0^2) = A_-(1-x)^{\eta_s}x^{\delta-}(1-x/x_0); xg(x,Q_0^2) = A_g(1-x)^{\eta_g}(1+\epsilon_g x^{0.5}+\gamma_g x)x^{\delta_g}+A_{g'}(1-x)$$

  - CTEQ/TEA: 22$\rightarrow$ 26 PARMS.

    $$x\,f(x,Q_0) = a_0\,x^{a_1}\,(1-x)^{a_2}\ \exp\left(a_3 x + a_4 x^2 + a_5\sqrt{x} + a_6 x^{-a_7}\right)\ (7\ \text{indep. functions})$$

    a_6, a_7 ONLY USED FOR GLUON

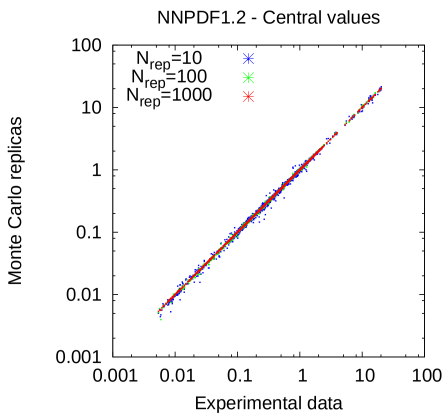
  - BASIS FUNCTIONS: $u_v \equiv u-\bar{u}$, $d_v \equiv d-\bar{d}$, $\bar{u}\pm\bar{d}$ (MSTW) OR $\bar{u}$, $\bar{d}$ (CTEQ), $s^{\pm} \equiv s\pm\bar{s}$ (CTEQ $\bar{s}+s$ ONLY), g.
- EVOLVE TO DESIRED SCALE & COMPUTE PHYSICAL OBSERVABLES
- DETERMINE BEST-FIT VALUES OF PARAMETERS
- DETERMINE ERROR BY PROPAGATION OF ERROR ON PARMS. ('HESSIAN METHOD') PARM. SCANS ALSO POSSIBLE ('LAGR. MULTIPLIER METHOD')

THE NNPDF APPROACH:
THE NEURAL MONTE CARLO

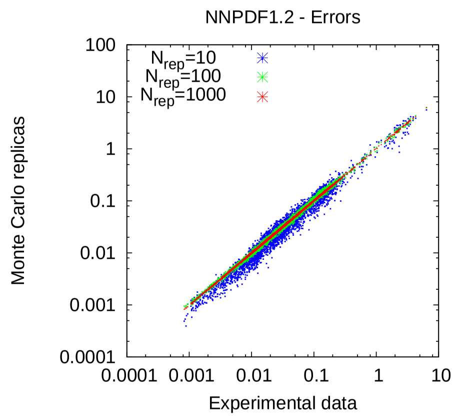
BASIC IDEA: MONTE CARLO SAMPLING
OF THE PROBABILITY MEASURE IN THE (FUNCTION) SPACE OF PDFs

- 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

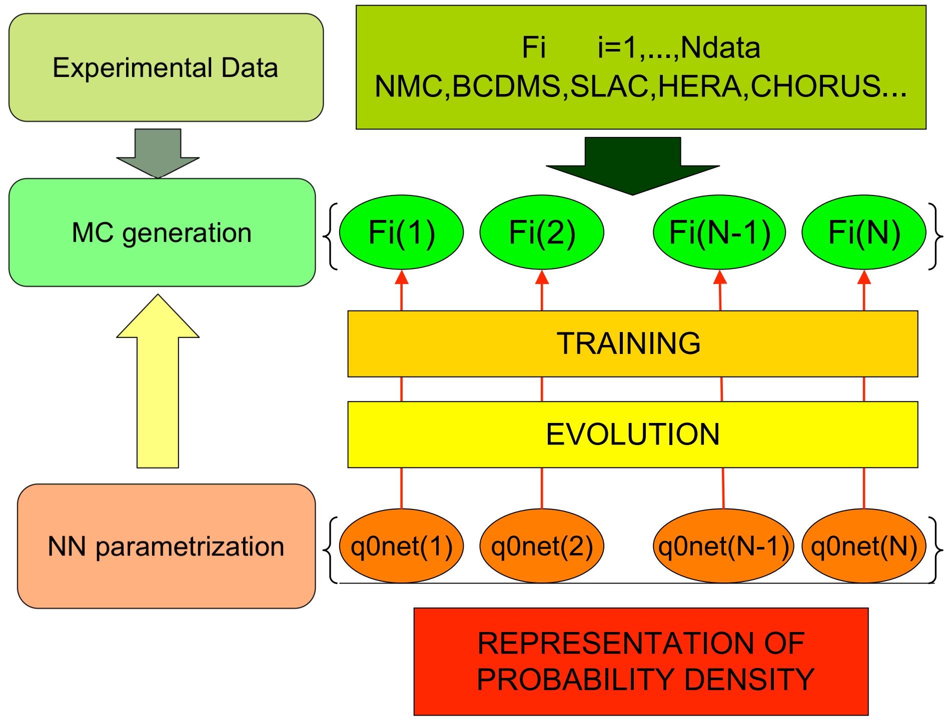
replica averages vs. central values



replica standard dev. vs. uncertainties

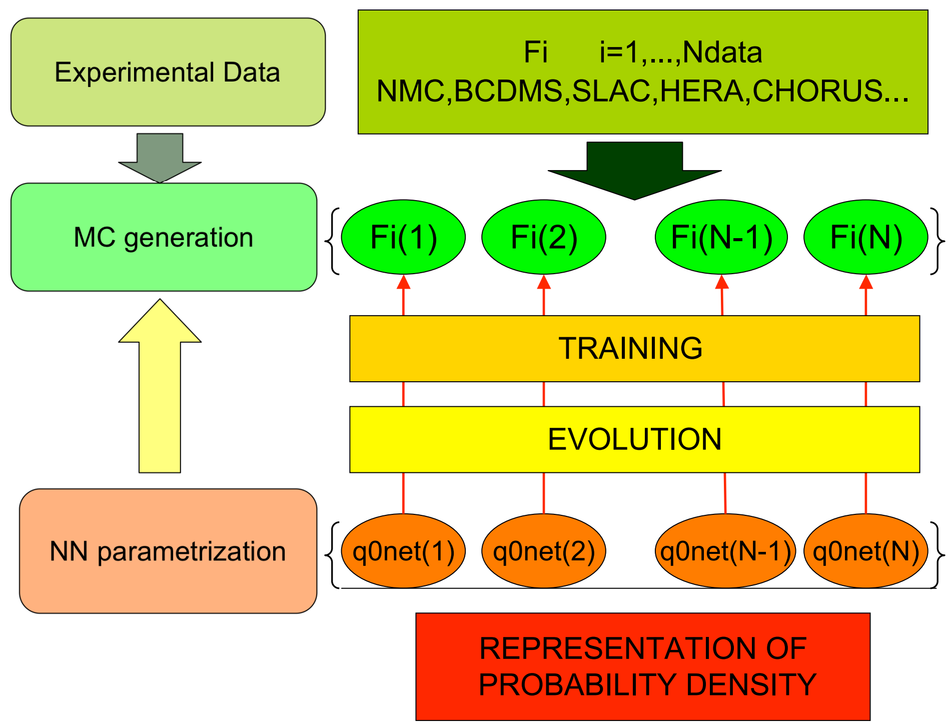


10 REPLICAS ENOUGH FOR CENTRAL VALS, 100 FOR UNCERTAINTIES, 1000 FOR CORRELNS



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 $\rightarrow$ 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")

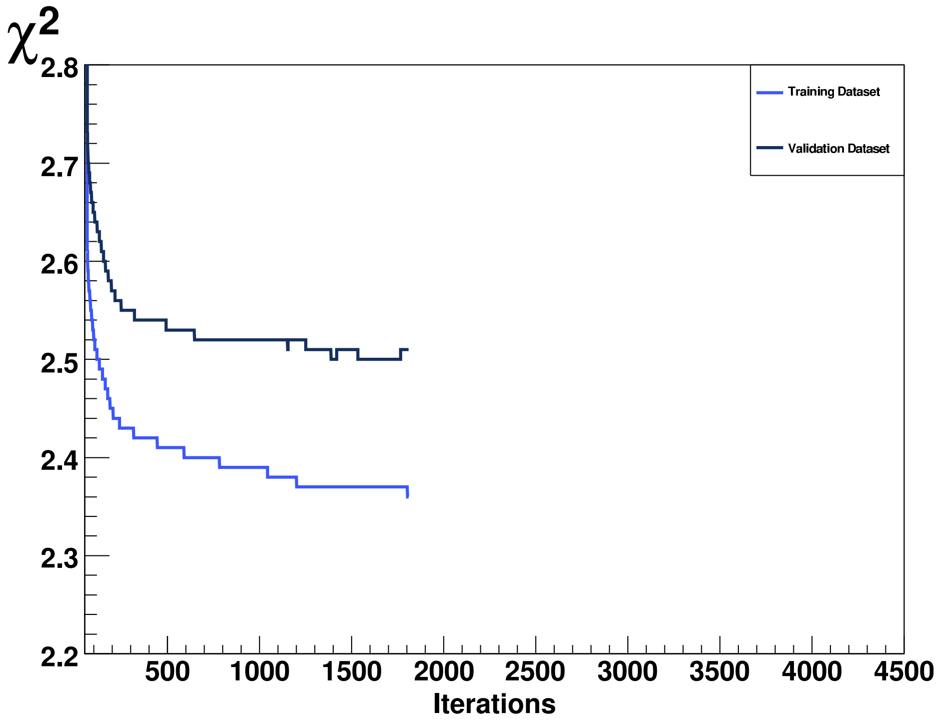


DATA MONTE CARLO $\Rightarrow$ PDF MONTE CARLO
CROSS-VALIDATION

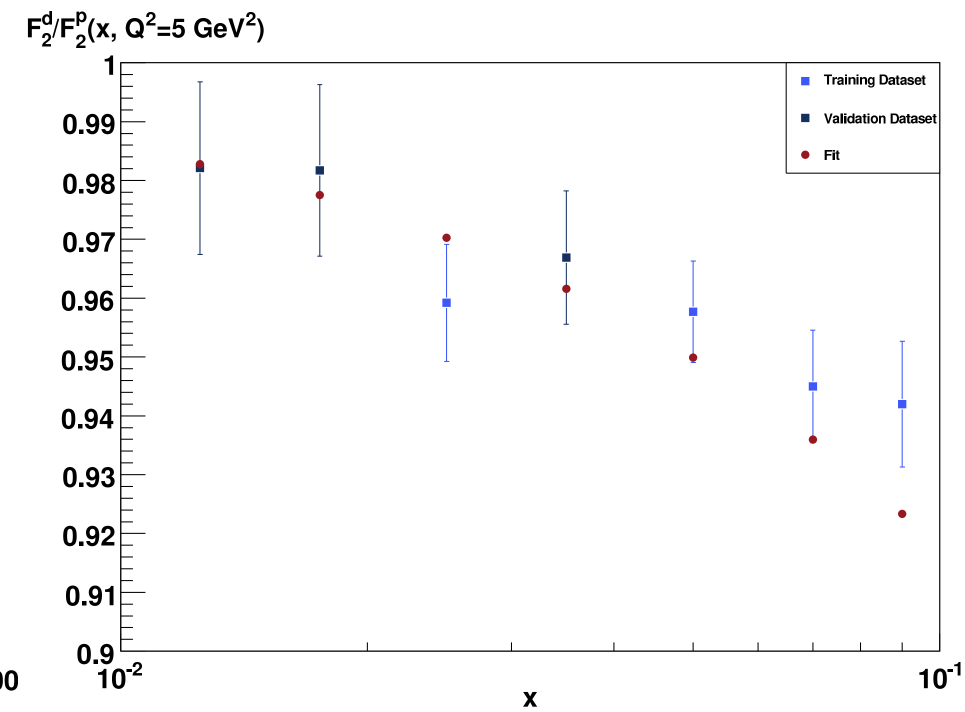
- REPLICAS ARE FITTED TO A DATA SUBSET
- A DIFFERENT SUBSET OF DATA USE FOR EACH REPLICA
- OPTIMAL FIT WHEN FIT TO VALIDATION (CONTROL) DATA STOPS IMPROVING
-

OPTIMAL FITTING

χ^2



FIT TO DATA

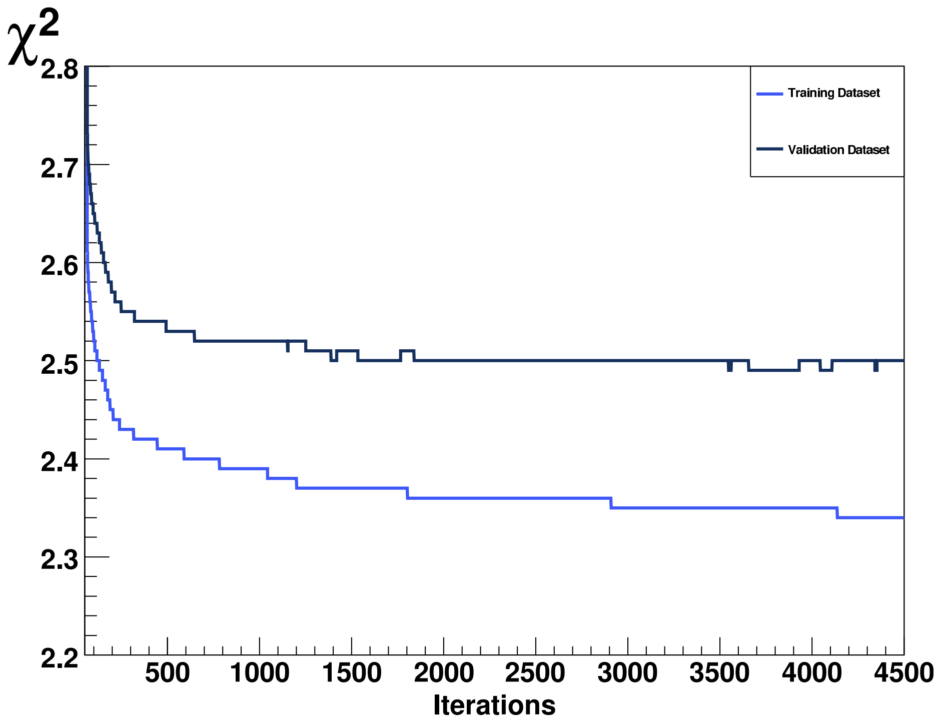


DATA MONTE CARLO ⇒ PDF MONTE CARLO
CROSS-VALIDATION

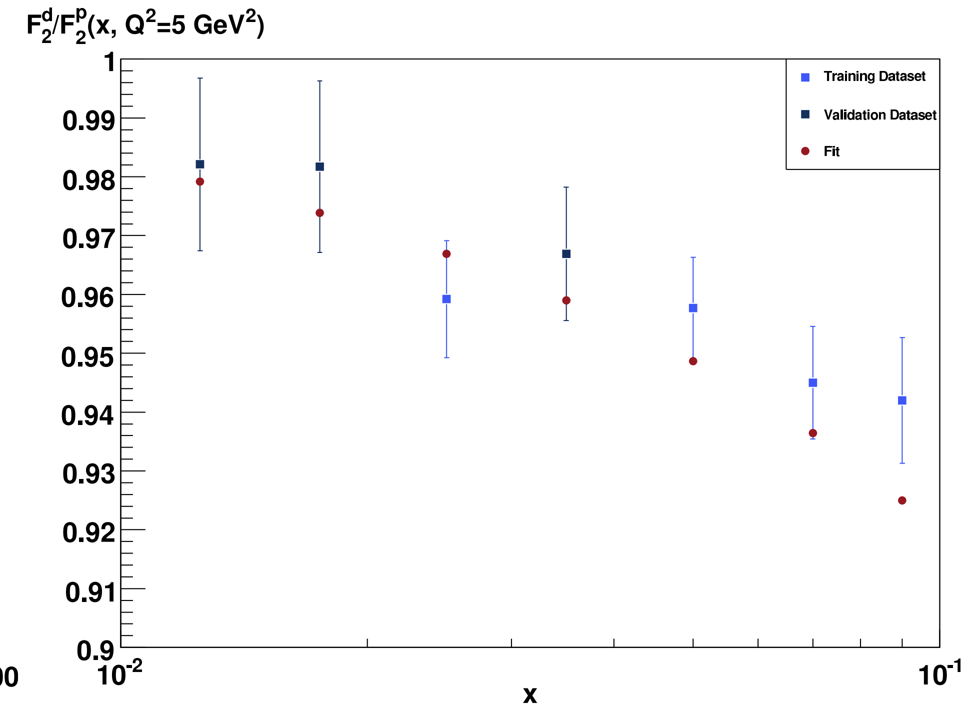
- REPLICAS ARE FITTED TO A DATA SUBSET
- A DIFFERENT SUBSET OF DATA USE FOR EACH REPLICA
- OPTIMAL FIT WHEN FIT TO VALIDATION (CONTROL) DATA STOPS IMPROVING
- THE BEST FIT IS NOT AT THE MINIMUM OF THE χ^2

OVERFITTING

χ^2



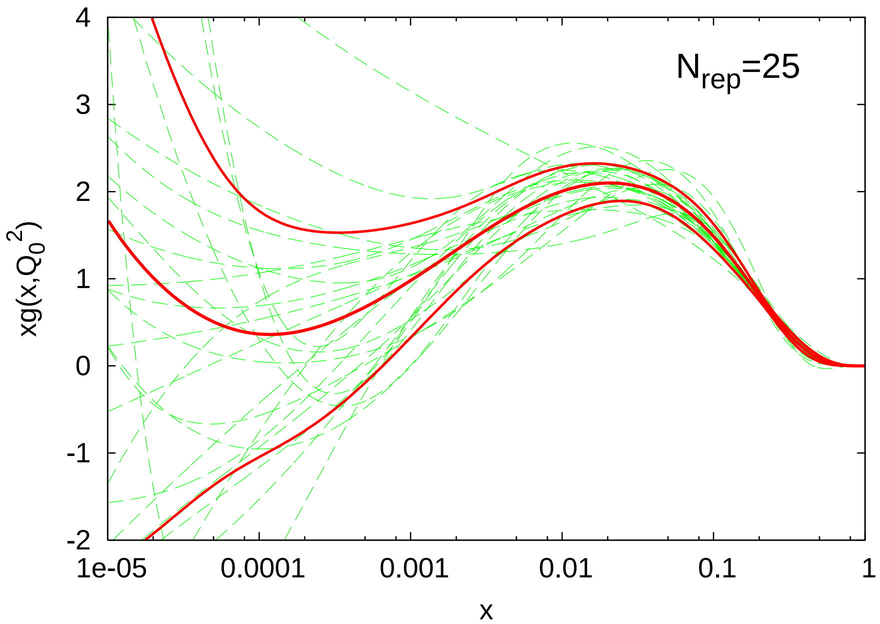
FIT TO DATA

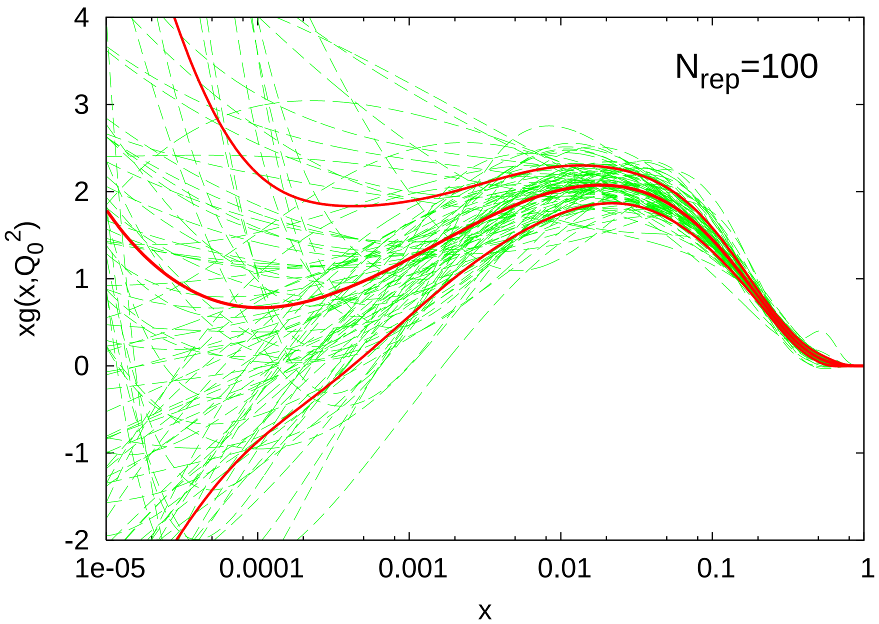


WHAT IS A ONE-σ UNCERTAINTY?

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

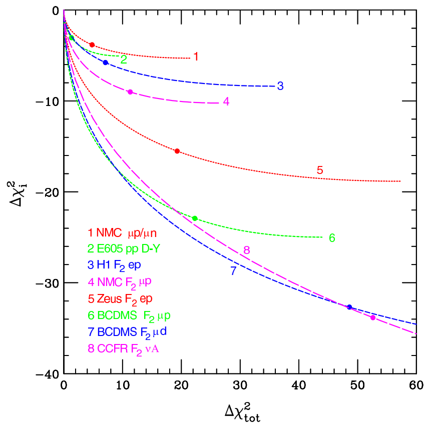
THE PROBLEM OF PDF UNCERTAINTIES

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

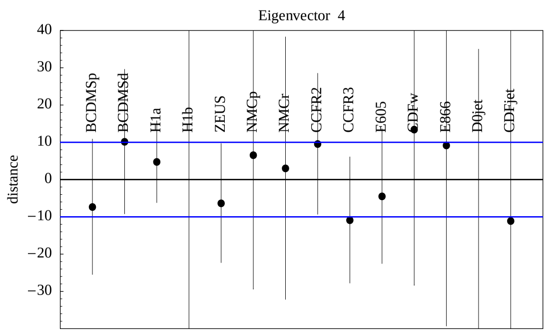
WHAT IS A ONE-σ UNCERTAINTY?

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

CTEQ TOLERANCE CRITERION &

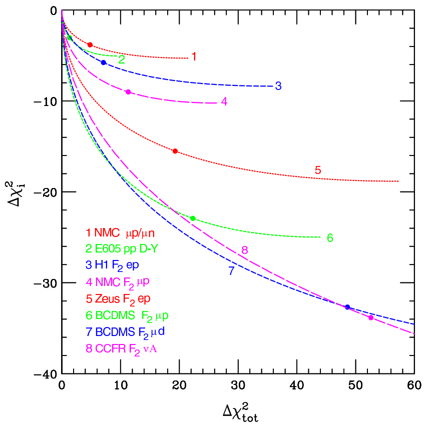
- STANDARD $\Delta\chi^2 = 1$ BANDS TOO NARROW $\Rightarrow$ LARGE DISCREPANCIES FOR INDIVIDUAL EXPERIMENTS
- TOLERANCE $\Rightarrow$ ENVELOPE OF UNCERTAINTIES OF EXPERIMENTS

CTEQ TOLERANCE PLOT FOR 4TH EIGENVEC.



MINIMUM χ^2_i

VS GLOBAL χ^2



Collins, Pumplin
2001

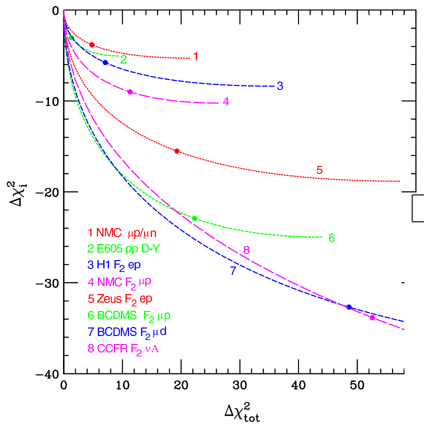
WHAT IS A ONE-σ UNCERTAINTY?

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

CTEQ TOLERANCE CRITERION & MSTW DYNAMICAL TOLERANCE

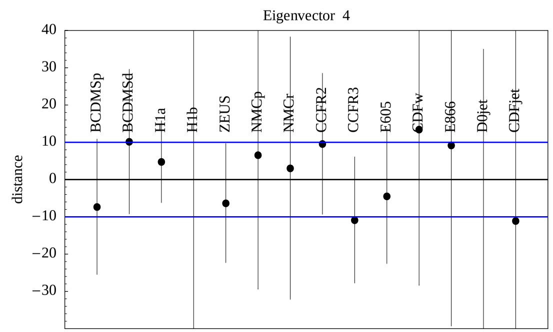
- STANDARD $\Delta\chi^2 = 1$ BANDS TOO NARROW $\Rightarrow$ LARGE DISCREPANCIES FOR INDIVIDUAL EXPERIMENTS
- TOLERANCE $\Rightarrow$ ENVELOPE OF UNCERTAINTIES OF EXPERIMENTS
- DYNAMICAL $\Rightarrow$ SEPARATELY DETERMINED FOR EACH HESSIAN EIGENVECTOR

MINIMUM χ_i^2 VS GLOBAL χ^2

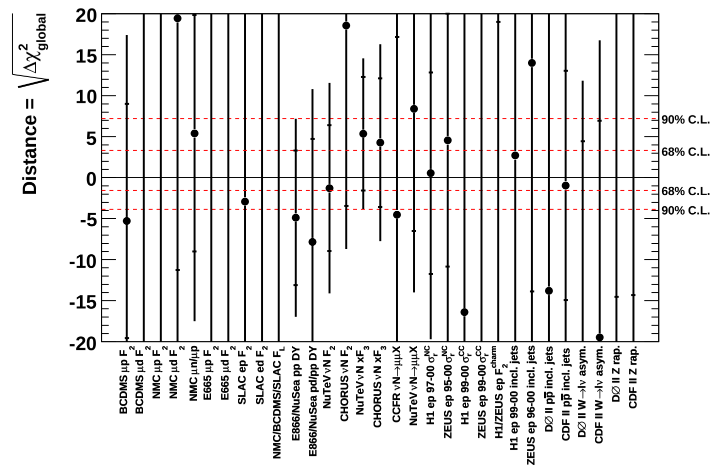


Collins, Pumplin 2001

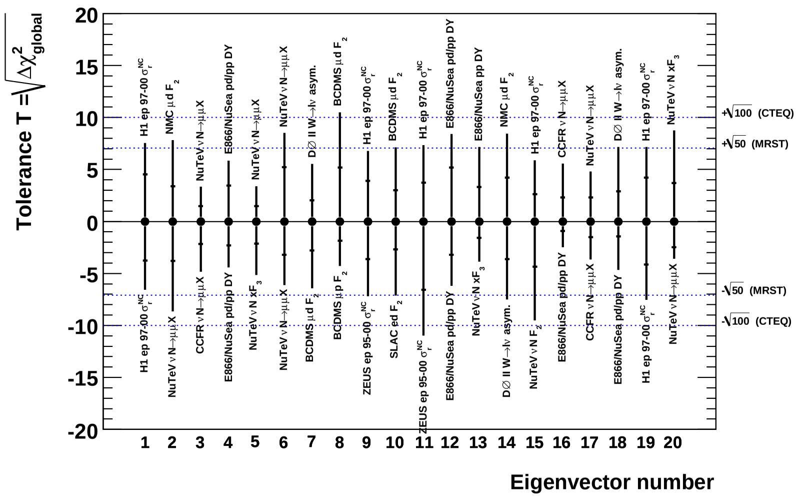
CTEQ TOLERANCE PLOT FOR 4TH EIGENVEC.



MSTW TOLERANCE PLOT FOR 13TH EIGENVEC.



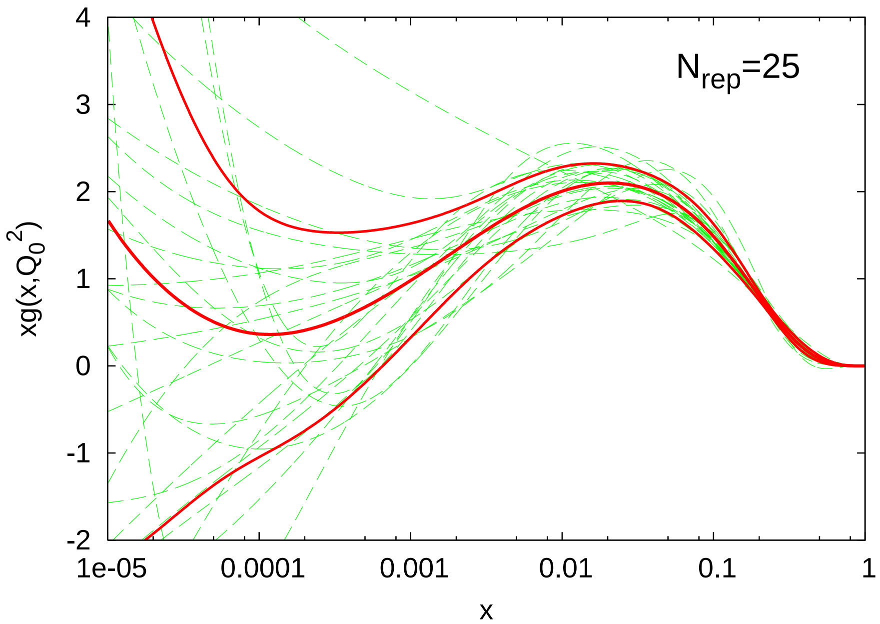
GLOBAL MSTW TOLERANCE

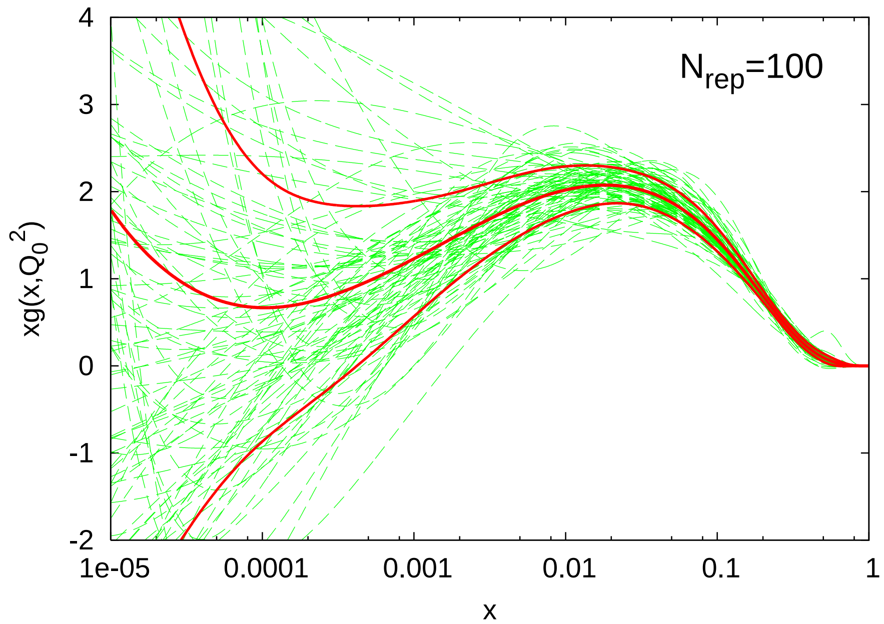


WHAT IS A ONE-σ UNCERTAINTY?

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

Example: the gluon distribution in the NNPDF1.0 set





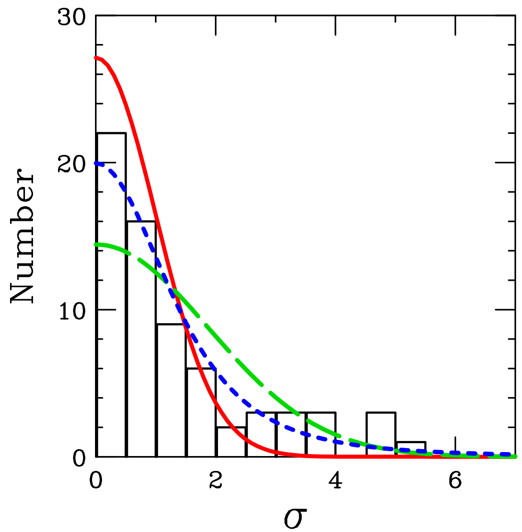
- ENSEMBLE OF REPLICAS $\leftrightarrow$ PROBABILITY DISTRIBUTION OF PDFS
- EXPECTED CENTRAL VALUE $\leftrightarrow$ MEAN; UNCERTAINTY $\leftrightarrow$ STANDARD DEVIATION
- ANY FEATURES OF DISTRIBUTION CAN BE DETERMINED (C.L. INTERVALS, CORRELATIONS...)

WHERE IS THE UNCERTAINTY COMING FROM?

WHY DOES ONE NEED LARGE TOLERANCES?

DATA INCOMPATIBILITY(Pumplin, 2009)

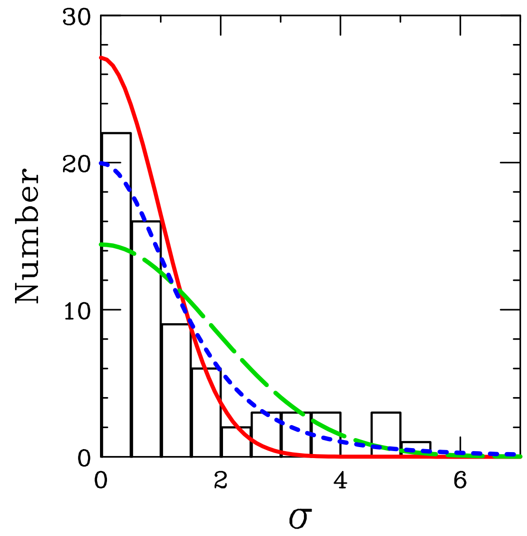
- CAN "REDIAGONALIZE": DIAGONALIZE SIMULTANEOUSLY χ^2 FOR TOTAL AND i–TH EXPT $\Rightarrow$ COMPATIBILITY OF EACH EXPT WITH GLOBAL FIT
- STUDY DISTRIBUTION OF DISCREPANCIES
- APPROX. GAUSSIAN WITH UNCERTAINTIES RESCALED BY 2 $\Rightarrow$ $\Delta\chi^2 \sim 10$ FOR 90%C.L.



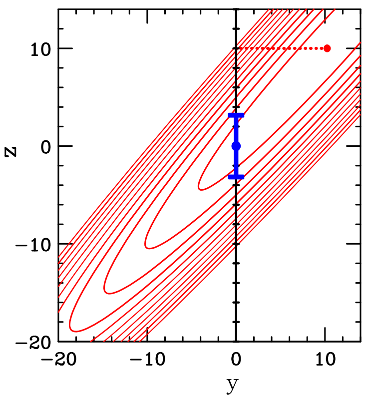
WHERE IS THE UNCERTAINTY COMING FROM?
WHY DOES ONE NEED LARGE TOLERANCES?

DATA INCOMPATIBILITY (Pumplin, 2009)

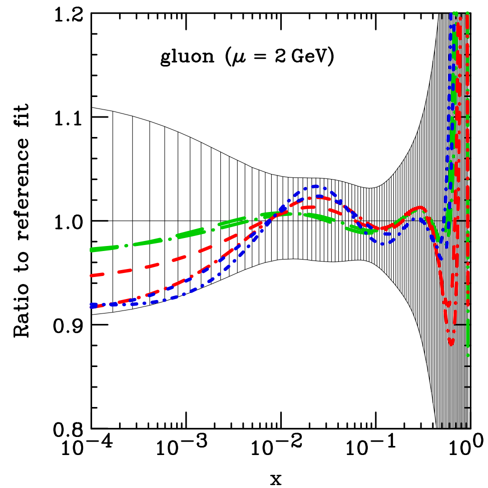
- CAN "REDIAGONALIZE": DIAGONALIZE SIMULTANEOUSLY χ^2 FOR TOTAL AND i–TH EXPT $\Rightarrow$ COMPATIBILITY OF EACH EXPT WITH GLOBAL FIT
- STUDY DISTRIBUTION OF DISCREPANCIES
- APPROX. GAUSSIAN WITH UNCERTAINTIES RESCALED BY 2 $\Rightarrow$ $\Delta\chi^2 \sim 10$ FOR 90%C.L.



FUNCTIONAL BIAS (Pumplin, 2009)



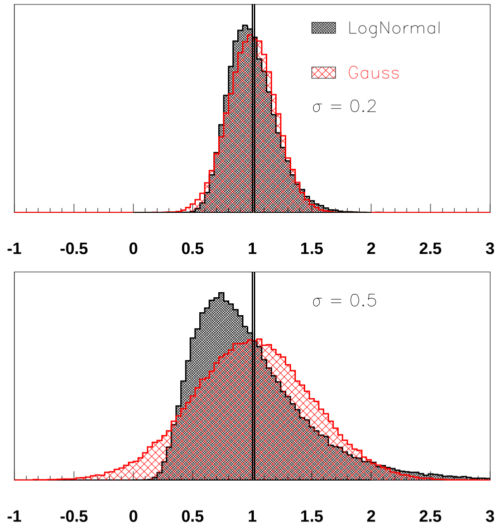
- IF PARM. NOT GENERAL ENOUGH, GLOBAL MIN. IS NOT TRUE MIN.
- ONE-σ VARIATION ABOUT FAKE MIN CORRESP. TO LARGE χ^2 VARIATION
- USE OF CHEBYSHEV POLYNOMIALS SUGGESTS "MOST GENERAL" PARM. WITHIN $\Delta\chi^2 = 10$ OF CTEQ6.6 PARM.



PARAMETRIZATION UNCERTAINTIES?

NONGAUSSIAN BEHAVIOUR?

LOGNORMAL VS. GAUSSIAN

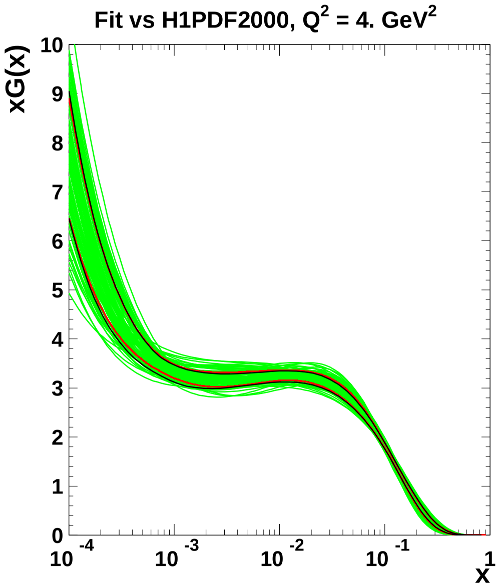


THE HERALHC BENCHMARK

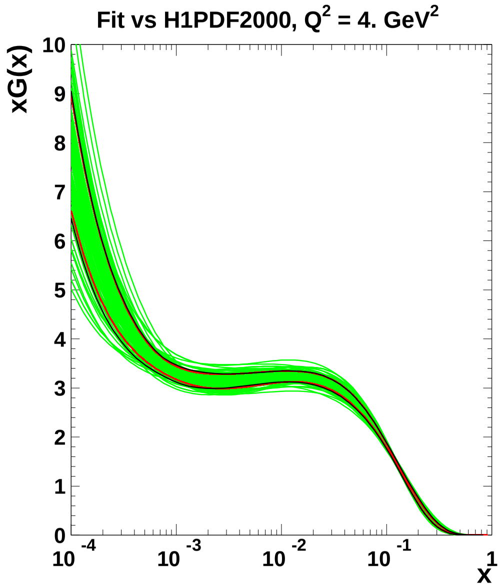
(Feltesse, Glazov, Radescu + NNPDF 2008)

- TRY EXPERIMENTAL SYSTEMATICS GIVEN BY EITHER GAUSSIAN OR LOGNORMAL DISTRIBUTION
- REPEAT (BENCHMARK) HERAPDF, WITH MONTECARLO LOGNORMAL OR GAUSSIAN, IN EITHER CASE DETERMINE UNCERTAINTY EITHER WITH HESSIAN OR MONTECARLO

LOGNORMAL: HESS. VS MC



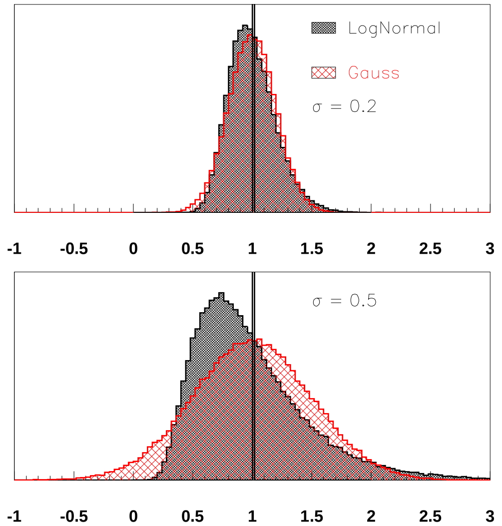
GAUSSIAN: HESS. VS MC



- NO DIFFERENCE BETWEEN LOGNORMAL, GAUSSIAN, MC, HESSIAN

PARAMETRIZATION UNCERTAINTIES?
NONGAUSSIAN BEHAVIOUR?

LOGNORMAL VS. GAUSSIAN

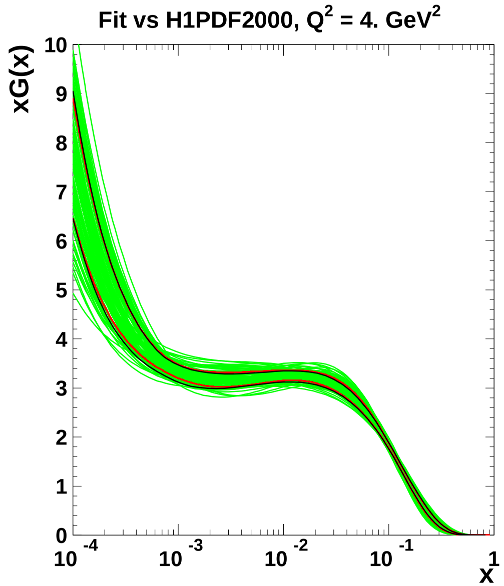


THE HERALHC BENCHMARK

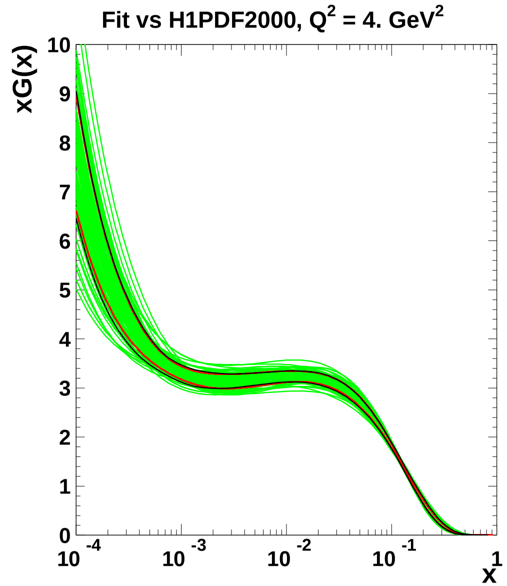
(Feltesse, Glazov, Radescu + NNPDF 2008)

- TRY EXPERIMENTAL SYSTEMATICS GIVEN BY EITHER GAUSSIAN OR LOGNORMAL DISTRIBUTION
- REPEAT (BENCHMARK) HERAPDF, WITH MONTECARLO LOGNORMAL OR GAUSSIAN, IN EITHER CASE DETERMINE UNCERTAINTY EITHER WITH HESSIAN OR MONTECARLO
- COMPARE TO NNPDF FIT TO SAME DATA

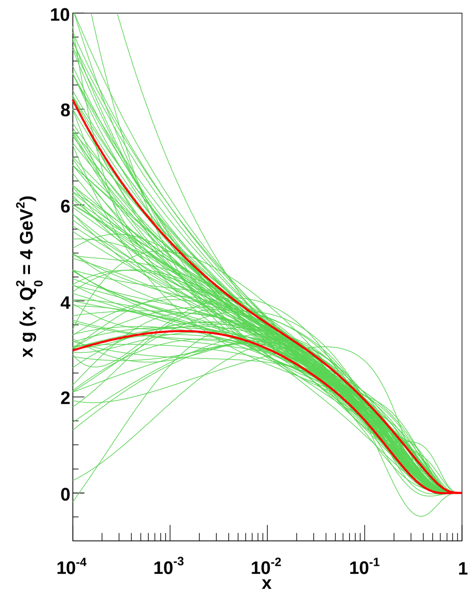
LOGNORMAL: HESS. VS MC



GAUSSIAN: HESS. VS MC



NNPDF

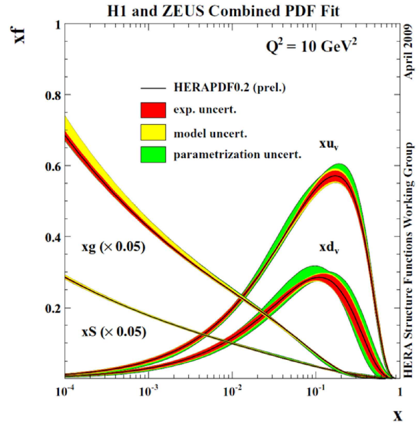


- NO DIFFERENCE BETWEEN LOGNORMAL, GAUSSIAN, MC, HESSIAN
- SIZABLE DIFFERENCE WR TO FLEXIBLE NNPDF PARAMETRIZATION

PARAMETRIZATION UNCERTAINTIES?

EXPLORING THE SPACE OF PARAMETERS: HESSIAN APPROACH

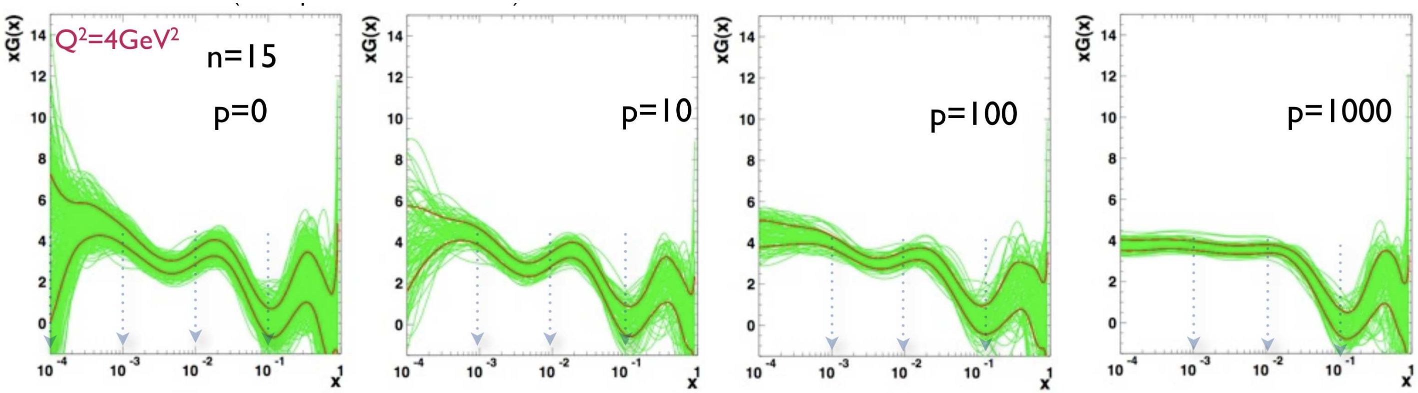
- IN HESSIAN APPROACH CAN VARY THE FUNCTIONAL FORM, ASSUMPTIONS, STARTING SCALE
- DONE IN THE HERAPDF1.0 FIT: VARIATION OF STRANGENESS FRACTION, LARGE x BEHAVIOUR, HIGHER ORDER POLYNOMIAL TERMS
- NO TOLERANCE ($\Delta\chi^2 = 1$), UNCERTAINTY DOUBLED



- Red: experimental uncertainties
- Yellow: model uncertainties
- Green: pdf parametrization uncertainties

ORTHOGONAL POLYNOMIALS

- OLD IDEA (PARISI, SOURLAS, 1978; ZOMER 1996): EXPAND PDFS OVER BASIS OF ORTHOGONAL POLYNOMIALS
- GLAZOV, RADESCU, 2009: COUPLED TO MONTE CARLO METHOD
- LENGTH PENALTY TO STABILIZE THE FIT



(Glazov, Radescu, 2009)

PARAMETRIZATION UNCERTAINTIES?

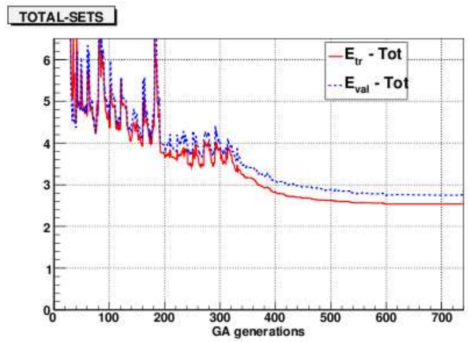
EXPLORING THE SPACE OF PARAMETERS: NNPDF APPROACH

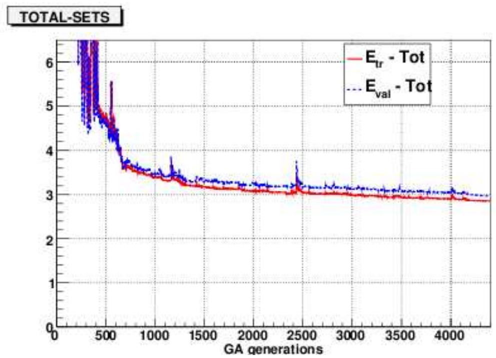
CENTRAL VALUES: VARYING PARTITION VS FIXED PARTITION

	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

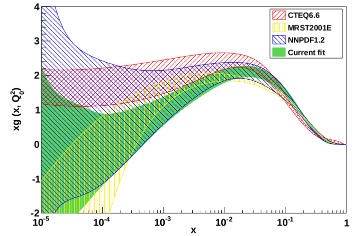
fixed partition results obtained averaging over 5 different choices of partition (100 replicas each); more partitions needed for accurate results

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

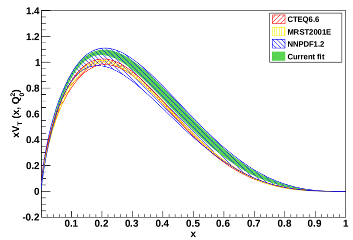




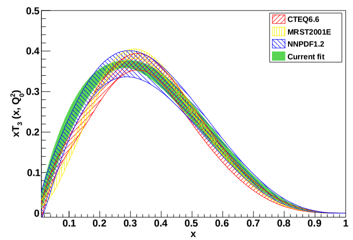
GLUE



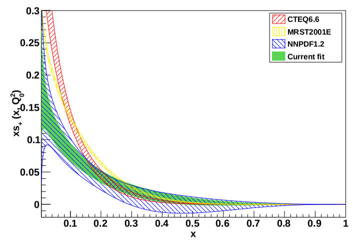
VALENCE



TRIPLET



STRANGE



PARAMETRIZATION UNCERTAINTIES?

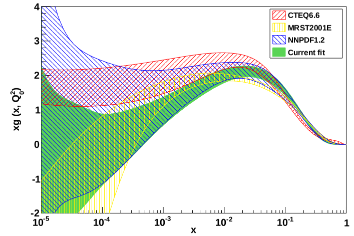
EXPLORING THE SPACE OF PARAMETERS: NNPDF APPROACH

CENTRAL VALUES: VARYING PARTITION VS FIXED PARTITION

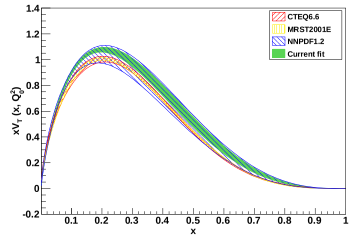
	REPLICAS	CENTRAL VALUE	FIXED PARTITION
χ^2	1.32	1.32	$\sim$1.3
$\langle\chi^2\rangle_{\rm rep}$	2.79 ± 0.24	1.65 ± 0.20	$\sim 1.6 \pm 0.2$
$\langle\sigma^{\rm dat}\rangle$	0.039	0.035	$\sim$0.03

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

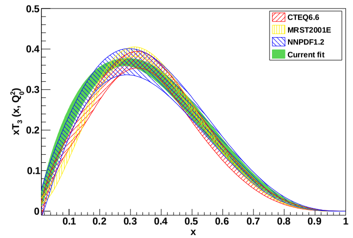
GLUE



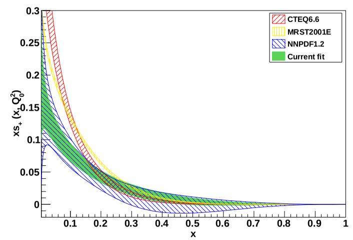
VALENCE



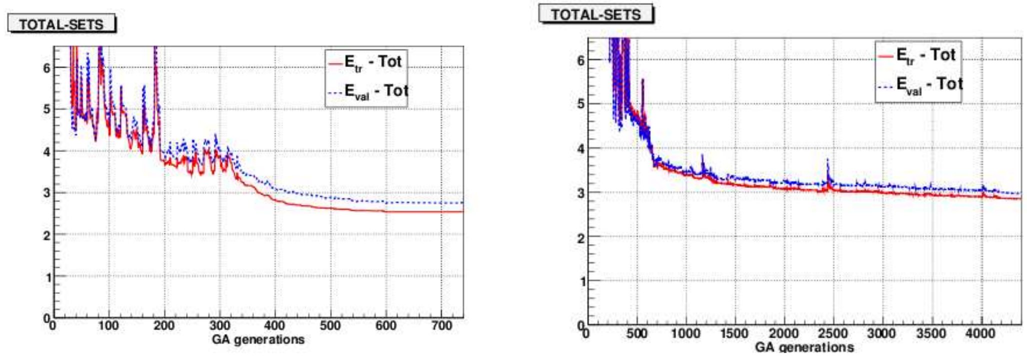
TRIPLET



STRANGE



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

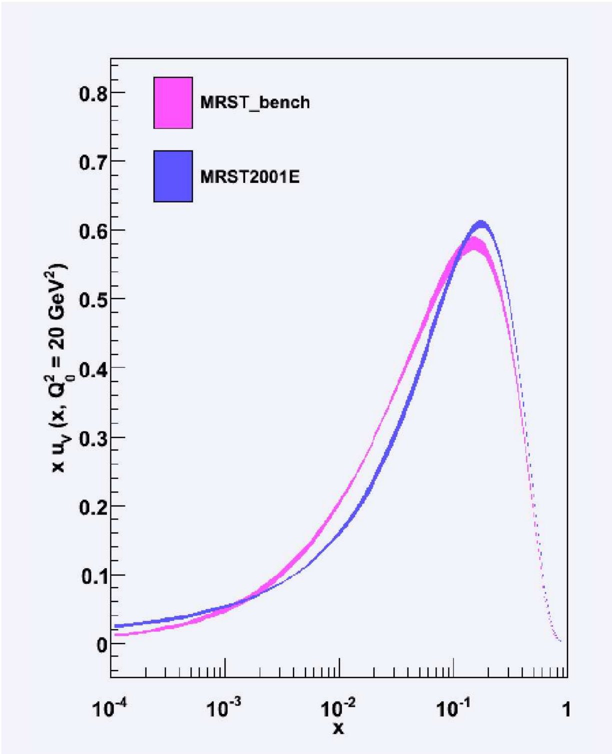


FUNCTIONAL UNCERTAINTY

- 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.

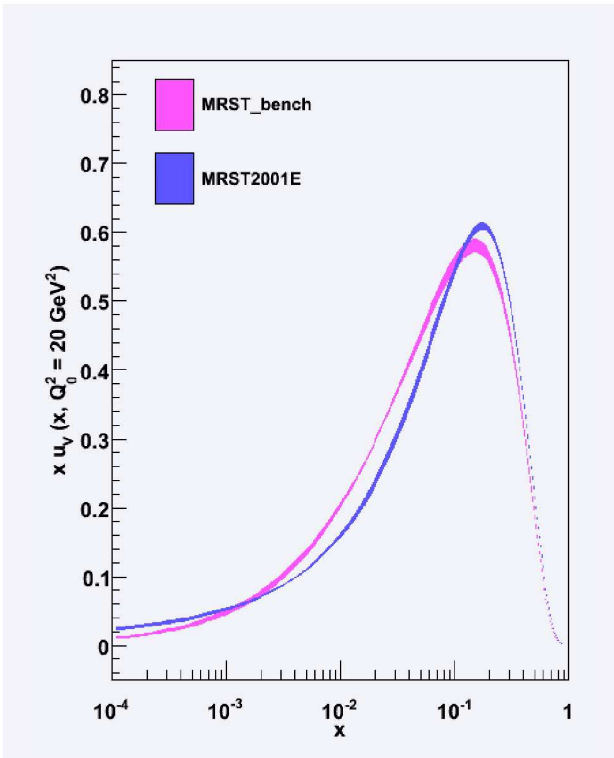
THE PROBLEM OF BENCHMARK FITS... (HERALHC 2005-2008)

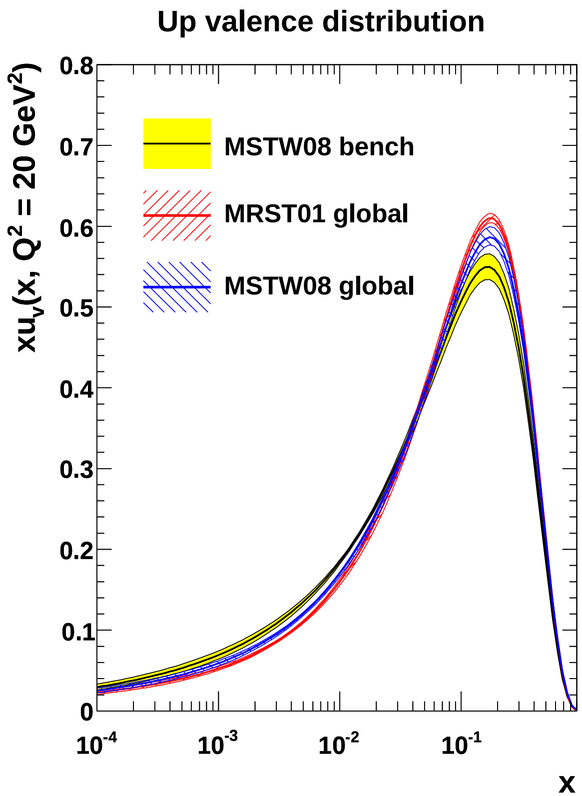
- PERFORM A MRST (MRSTBENCH) FIT TO A CONSISTENT SUBSET OF DATA, USE $\Delta\chi^2 = 1$
  $\Rightarrow$ RESULTS NOT CONSISTENT, UNCERTAINTY DOES NOT GROW AS DATASET DECREASES



THE PROBLEM OF BENCHMARK FITS... (HERALHC 2005-2008)

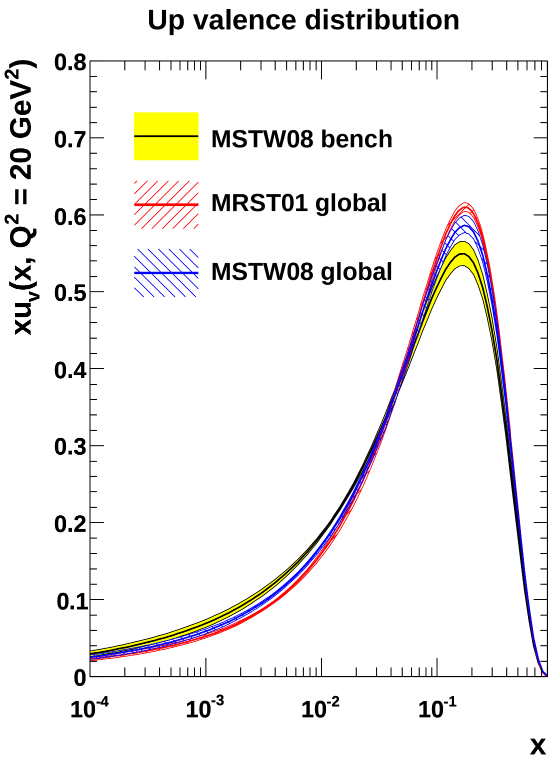
- PERFORM A MRST (MRSTBENCH) FIT TO A CONSISTENT SUBSET OF DATA, USE $\Delta\chi^2 = 1$
  $\Rightarrow$ RESULTS NOT CONSISTENT, UNCERTAINTY DOES NOT GROW AS DATASET DECREASES
- ...BUT MRST WAS DONE WITH TOLERANCE 50: REPEAT WITH DYNAMICAL TOLERANCE (MSTW08BENCH)
- IMPROVEMENT, BUT PROBLEM NOT SOLVED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



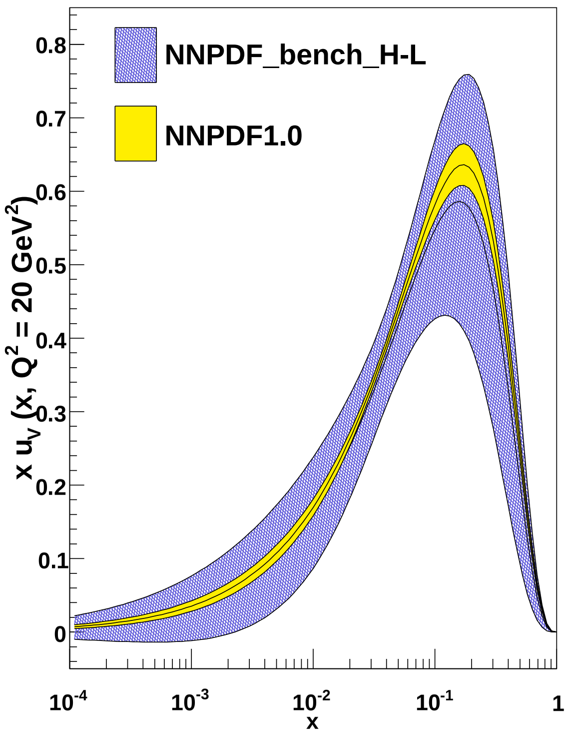


...AND THE NNPDF SOLUTION (HERALHC 2008)

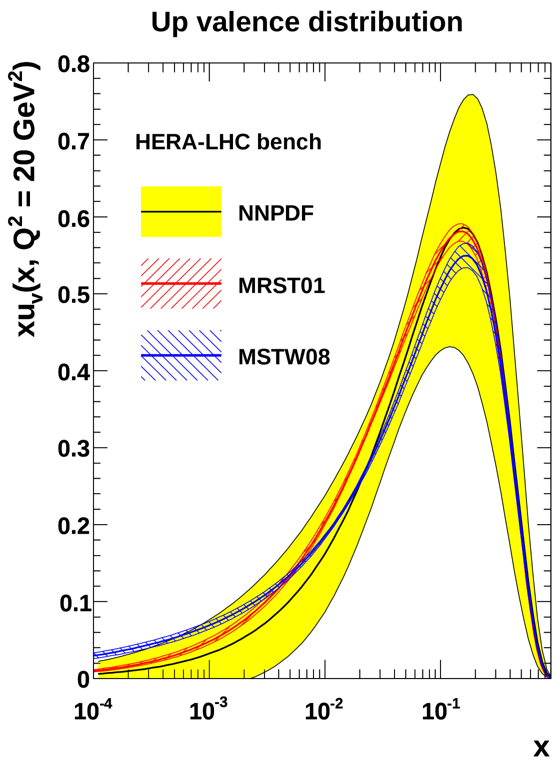
MRST/MSTW: BENCH VS REF



NNPDF: BENCH VS REF



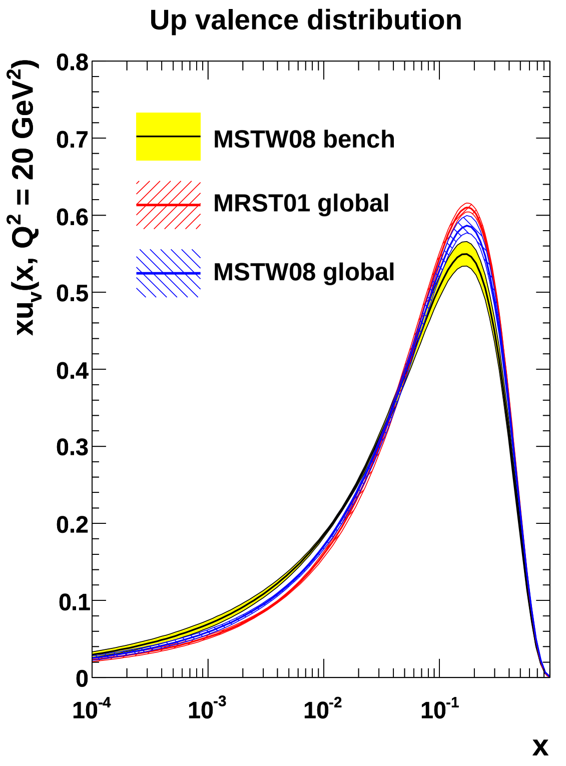
NNPDF BENCH VS MRST/MSTW BENCH



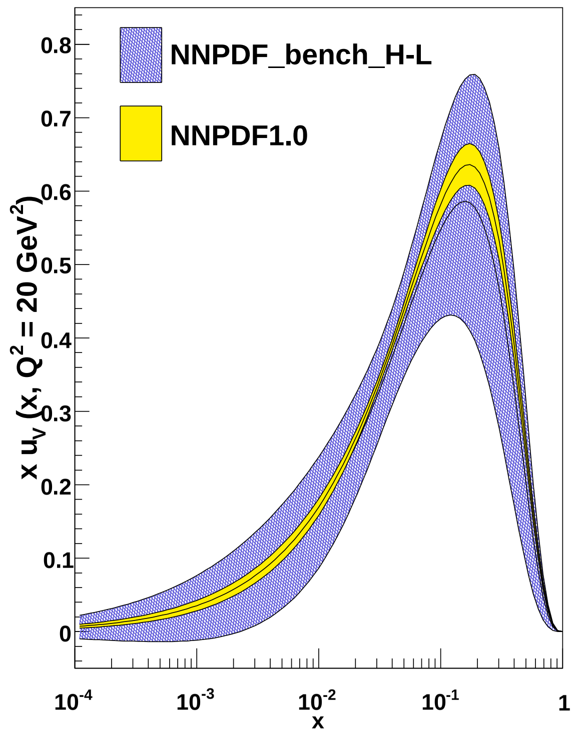
- SINGLE PARAMETRIZATION AND STAT. TREATMENT CAN ACCOMMODATE DIFFERENT DATASETS

…AND THE NNPDF SOLUTION (HERALHC 2008)

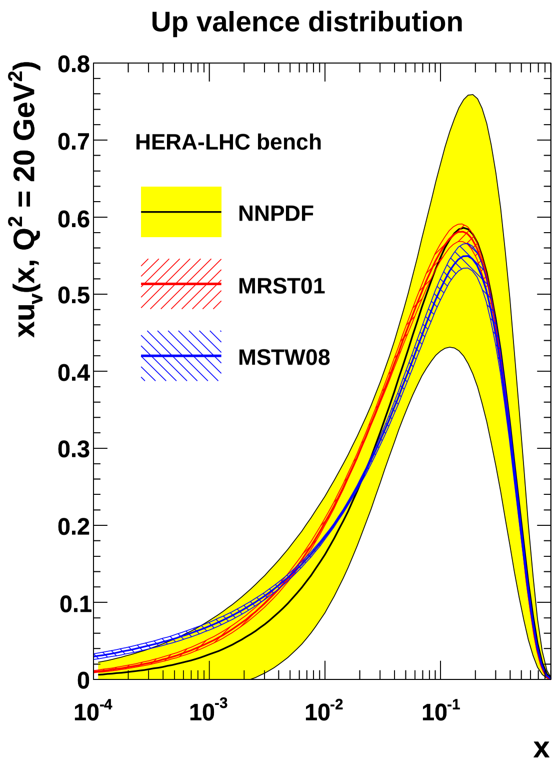
MRST/MSTW: BENCH VS REF



NNPDF: BENCH VS REF



NNPDF BENCH VS MRST/MSTW BENCH



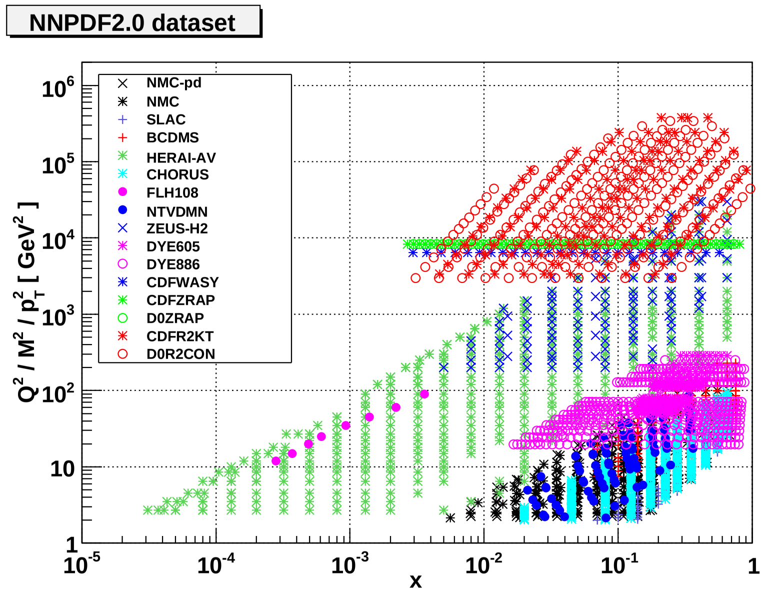
- SINGLE PARAMETRIZATION AND STAT. TREATMENT CAN ACCOMMODATE DIFFERENT DATASETS
- IMPACT OF DATA CAN BE STUDIED INDEPENDENT OF THEORETICAL FRAMEWORK

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,...)

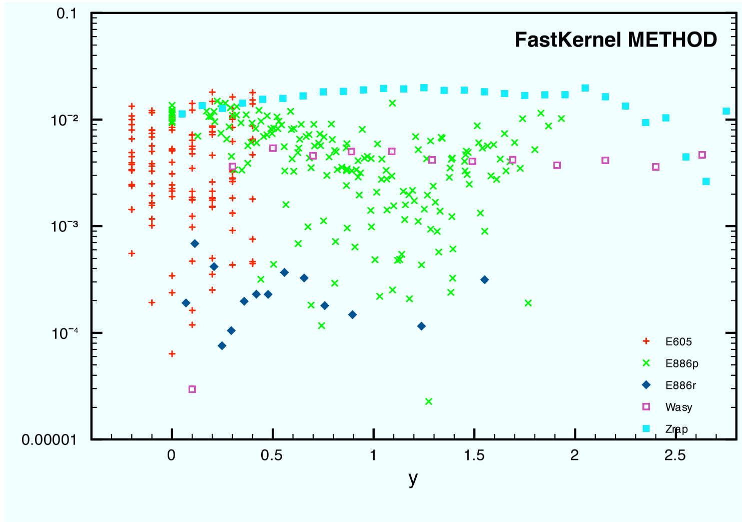
THE PROBLEM OF NLO FITTING

- 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

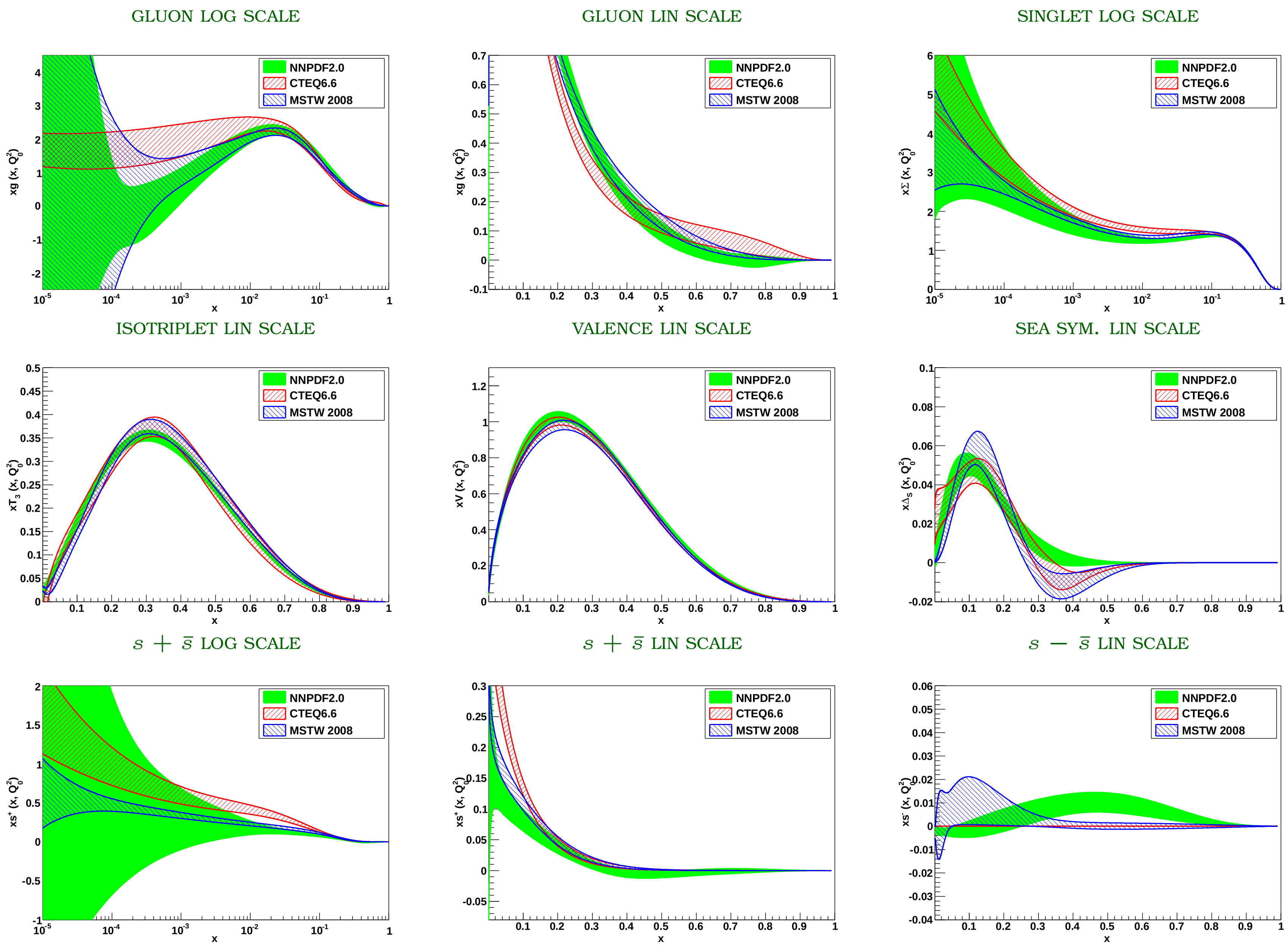
NNPDF2.0: THE FIRST GLOBAL NLO FIT

- 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



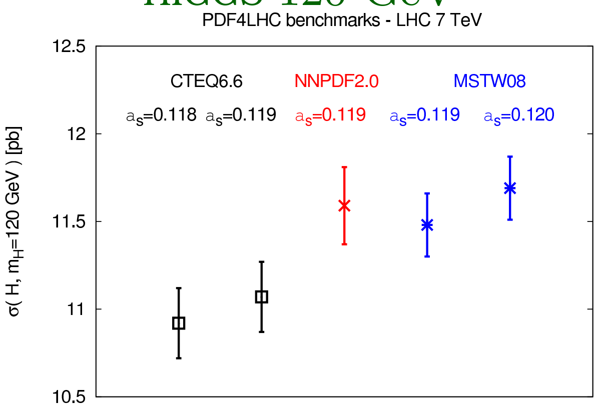
NNPDF2.0 PDFS VS. OTHER GLOBAL SETS



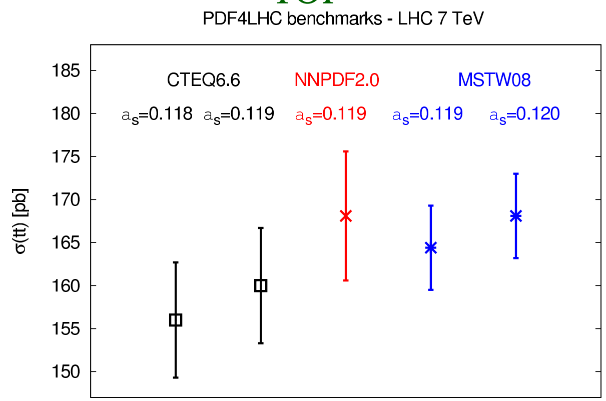
NNPDF2.0 STANDARD CANDLES
VS. OTHER GLOBAL SETS

- 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

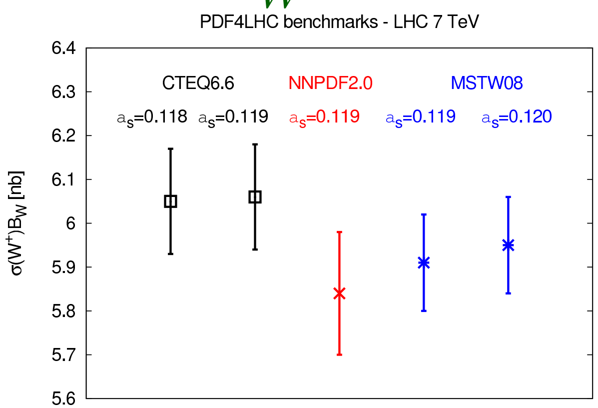
HIGGS 120 GeV



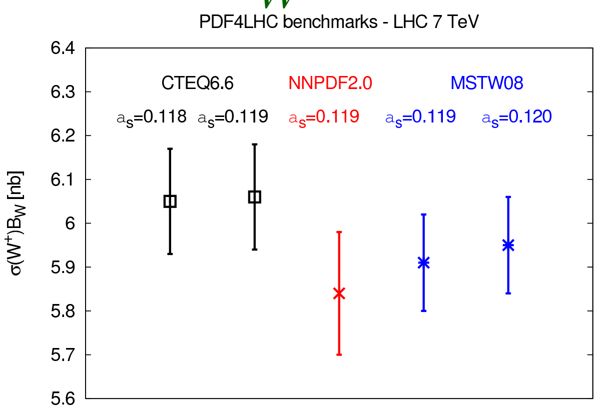
TOP



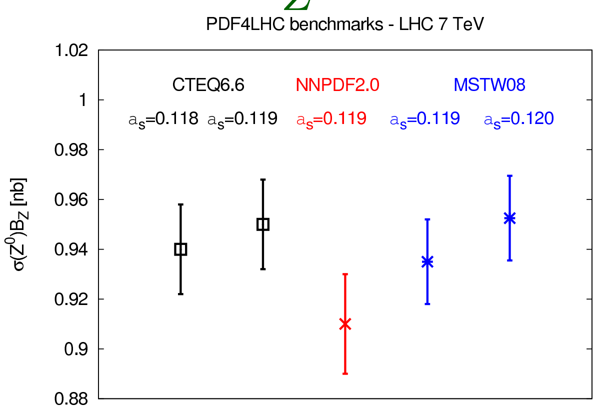
W^+



W^-



Z



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

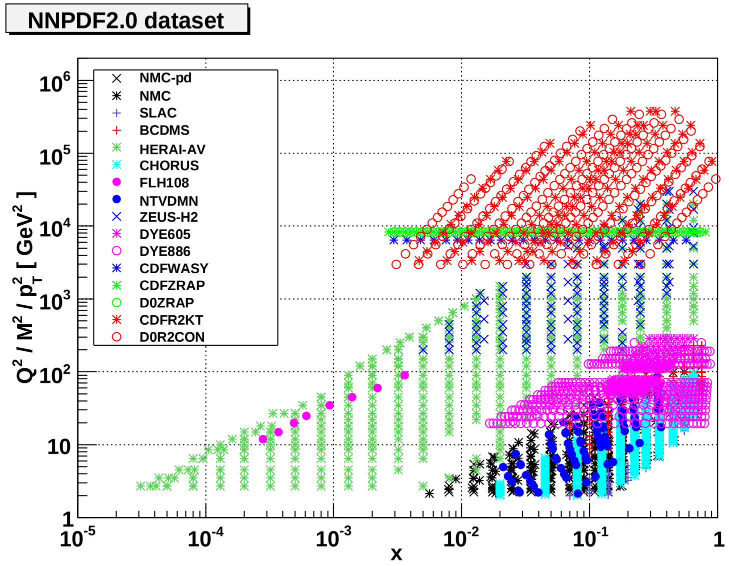
A USEFUL TOOL:
DISTANCE BETWEEN PDF SETS

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

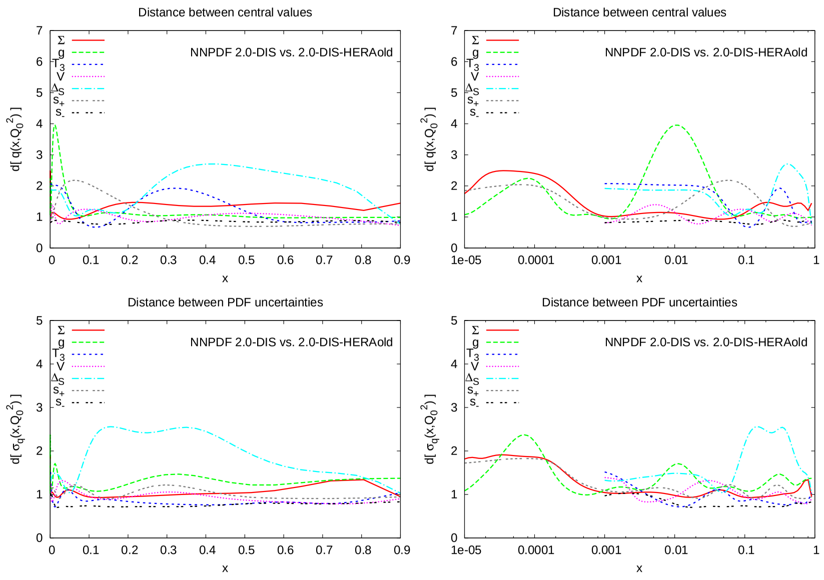
- MEAN PDF EXTRACTED FROM MONTECARLO SET OF SIZE $N_{\rm rep}$ FLUCTUATES WITH STANDARD DEVIATION $\frac{\sigma}{N_{\rm rep}}$
- IF TWO MC SETS EXTRACTED FROM SAME UNDERLYING DISTRIBUTION, THEN $d \sim 1$

FLEXIBILITY ⇒ 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

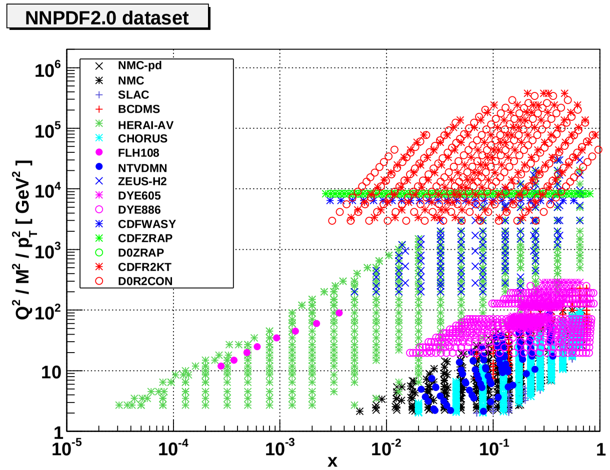


- PDFs UNCHANGED, MARGINAL IMPROVEMENT IN GLUON
- QUALITY OF FIT TO FIXED-TARGET DIS UNCHANGED
- SLIGHT DETERIORATION IN QUALITY OF FIT TO HERA DATA
- ⇒ COULD BE EVIDENCE FOR (MARGINALLY) INADEQUATE THEORY

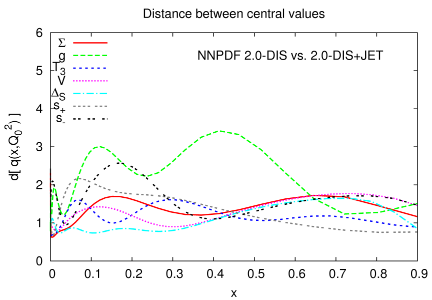


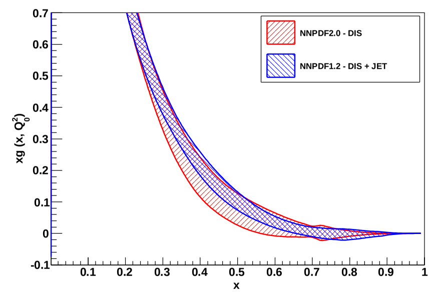
FLEXIBILITY $\Rightarrow$ CONSISTENCY: DIS DATA VS. JET DATA

	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



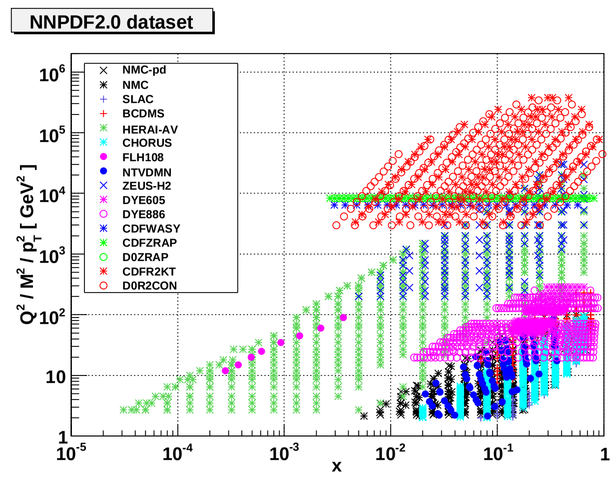
- HIGH E_T JET DATA WELL REPRODUCED EVEN WHEN NOT FITTED $\Rightarrow$ LARGE x GLUON WELL DETERMINED BY SCALING VIOLATIONS!
- SIGNIFICANT IMPROVEMENT IN LARGE x GLUON ACCURACY
- OTHER PDFS UNCHANGED



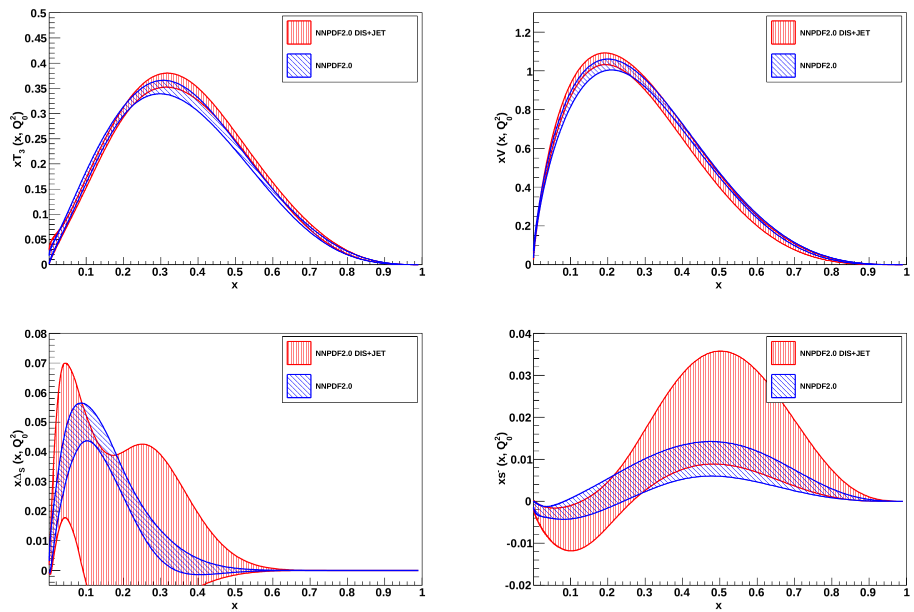


CONSISTENCY ⇒ ACCURACY:
DIS+JETS VS. DRELL-YAN (AND W, Z) DATA

	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**



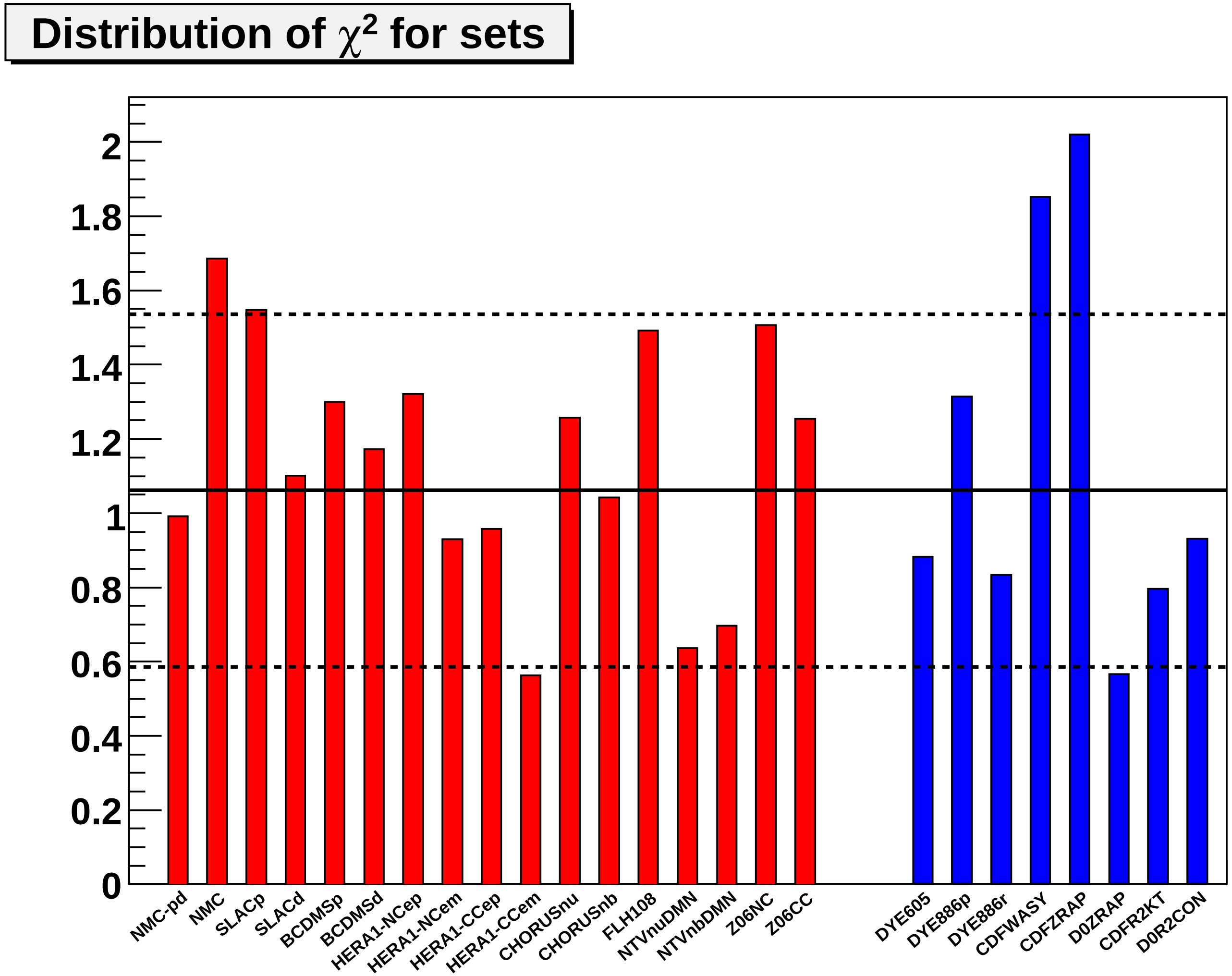
- VERY SUBSTANTIAL IMPROVEMENT IN FIT QUALITY WHEN DATA INCLUDED ⇒ SOME PDF COMBINATIONS POORLY DETERMINED WITHOUT THESE DATA
- HUGE IMPROVEMENT IN SEA ASYM $\bar u - \bar d$ & STRANGENESS $s - \bar s$
- SIGNIFICANT IMPROVEMENT IN TOTAL VALENCE $(\sum_i (q_i - \bar q_i))$ & ISOTRIPLET $(u + \bar u -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



DATA INCOMPATIBILITY?

FIT QUALITY: DIS DATA AND HADRONIC DATA

NNPDF2.0



- NO OBVIOUS MUTUAL TENSION BETWEEN DIS AND HADRONIC DATA
- CLEAR SIGN OF INTERNAL DATA INCONSISTENCIES (NMC DIS DATA, CDF Z AND W RAPIDITY DISTRIBUTIONS)

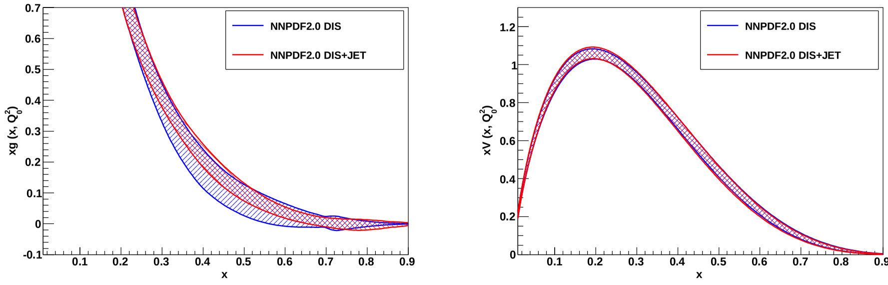
DATA INCOMPATIBILITY?

DIS VS. HADRONIC DATA

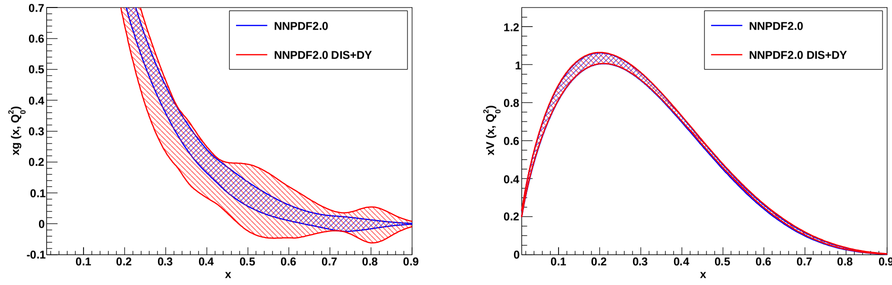
A SENSITIVE TEST: IS THE IMPACT OF A DATASET INDEP. OF THE DATA IT IS ADDED TO?

ADDING JET DATA. . .

. . . TO DIS DATA

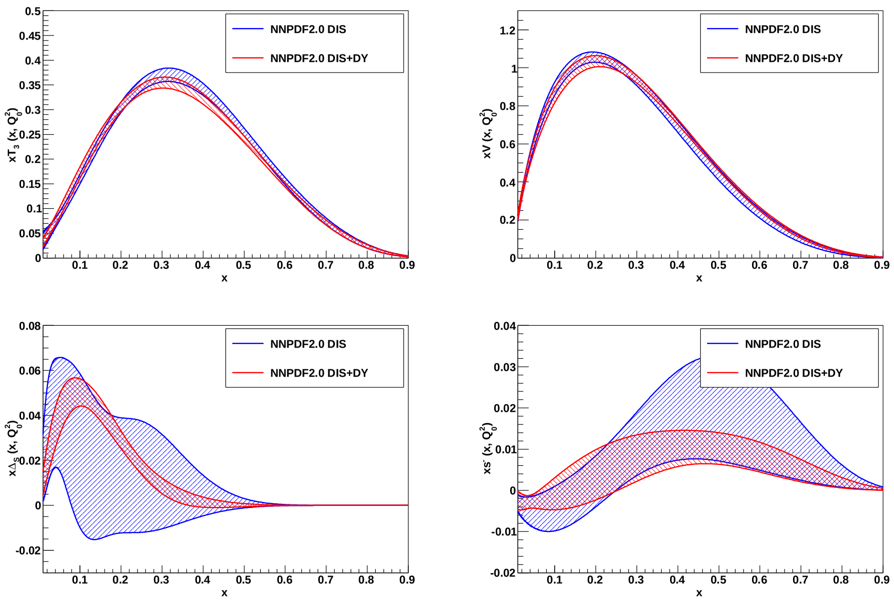


. . . TO DIS+DY DATA

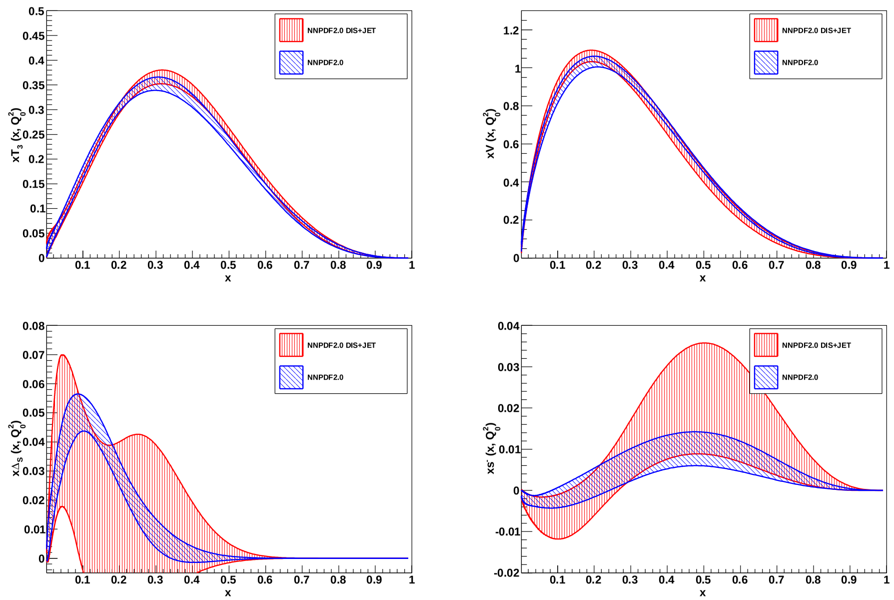


ADDING DRELL-YAN DATA. . .

. . . TO DIS DATA



. . . TO DIS+JET DATA



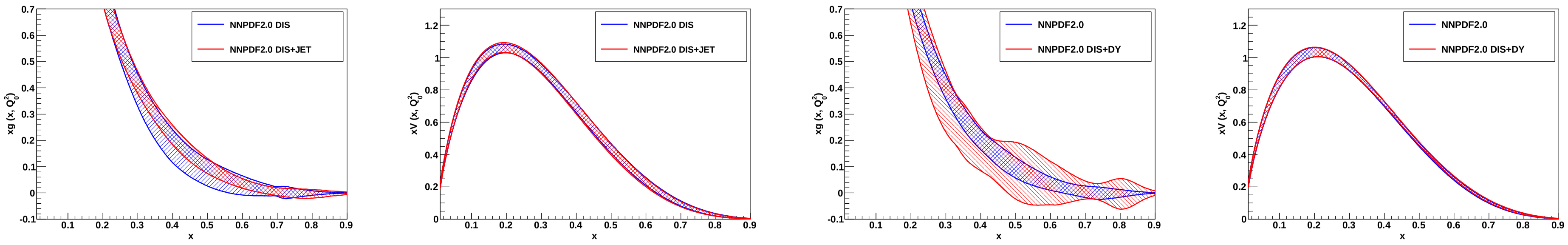
DATA INCOMPATIBILITY?

DIS VS. HADRONIC DATA

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

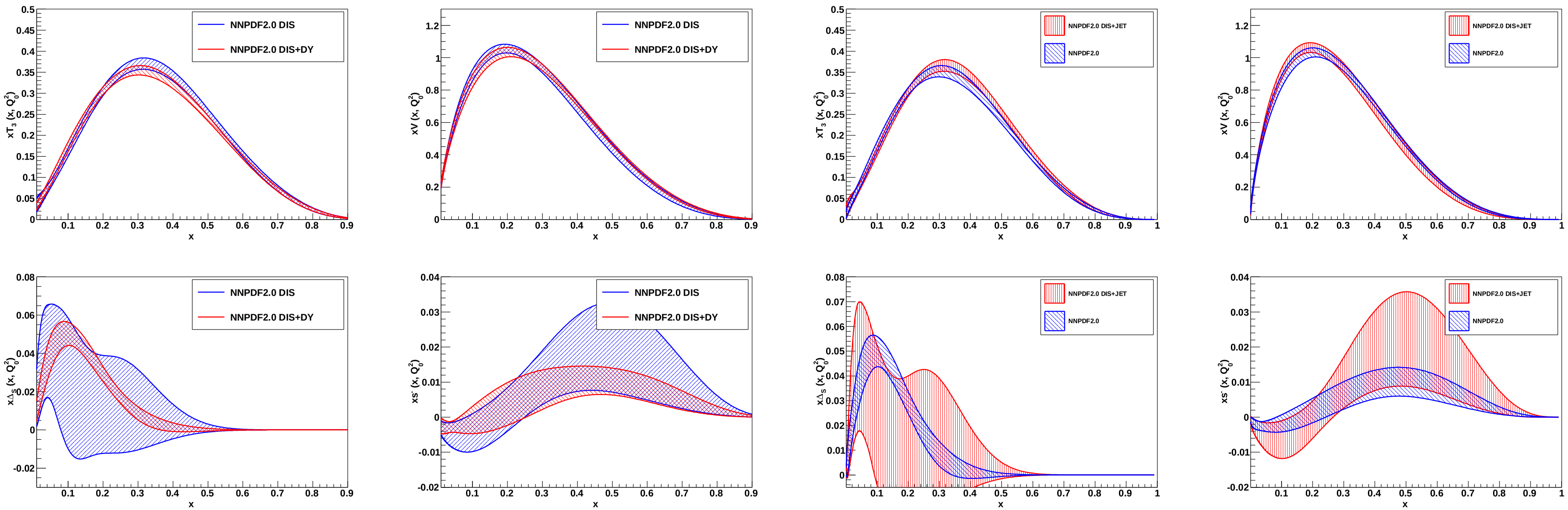
ADDING JET DATA. . .

. . . TO DIS DATA | . . . TO DIS+DY DATA



ADDING DRELL-YAN DATA. . .

. . . TO DIS DATA | . . . TO DIS+JET DATA



FITS COMMUTE $\Rightarrow$ GOOD COMPATIBILITY!

SENSITIVITY TO DISCOVERY

A DISCOVERY THAT WASN'T
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

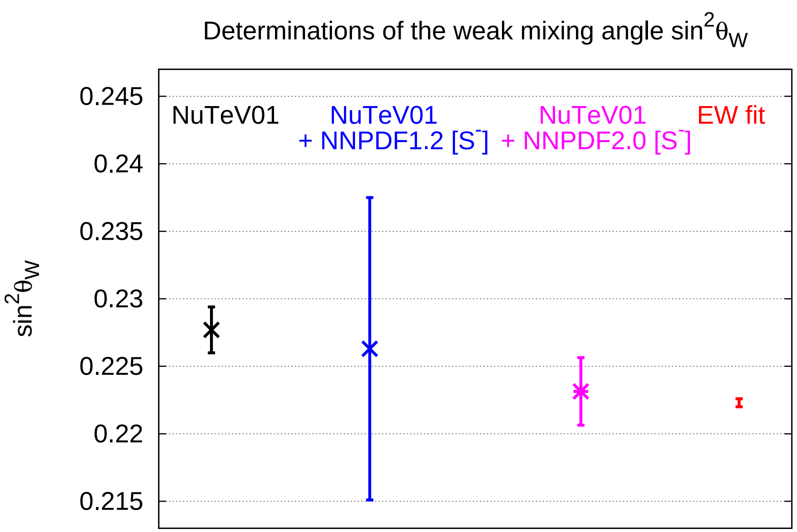
A DISCOVERY THAT WASN'T
THE NUTEV ANOMALY...

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

- 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 $\Rightarrow$ THREE σ DISCREPANCY WITH GLOBAL FIT

...IS GONE

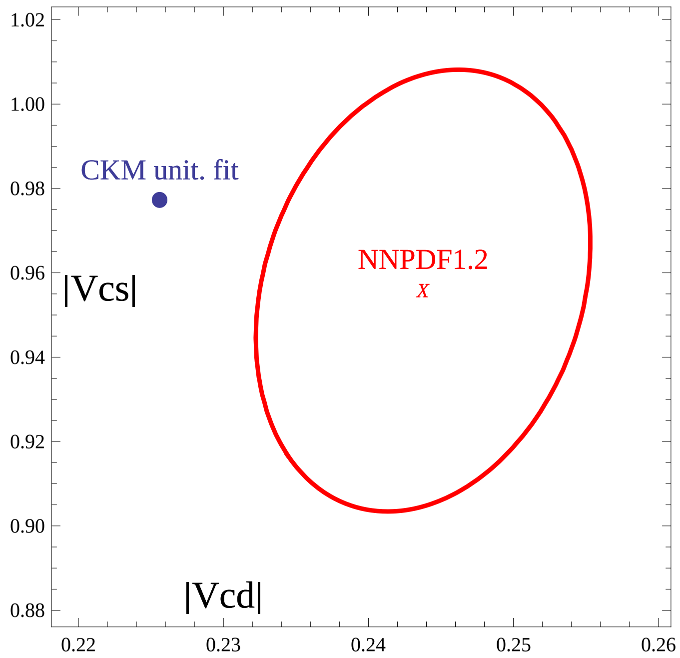
- NNPDF: s, $\bar s$ LIKE ANY OTHER PDF $\to$ 37 PARMS. EACH (CTEQ6.6: $s=\bar s$, TWO PARMS; MSTW08, S4S4, $\bar s$ TWO PARMS EACH)
- IF STRANGENESS UNCERTAINTY KEPT INTO ACCOUNT (DIS ONLY FIT: NNPDF1.2) $\Rightarrow$ EFFECT LOSES STAT. SIGNIFICANCE
- IF HADRONIC DATA INCLUDED (NNPDF2.0 GLOBAL FIT)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



(NNPDF2.0, 2010)

CKM MATRIX ELEMENTS

CAN BE DETERMINED WITH SURPRIZING ACCURACY

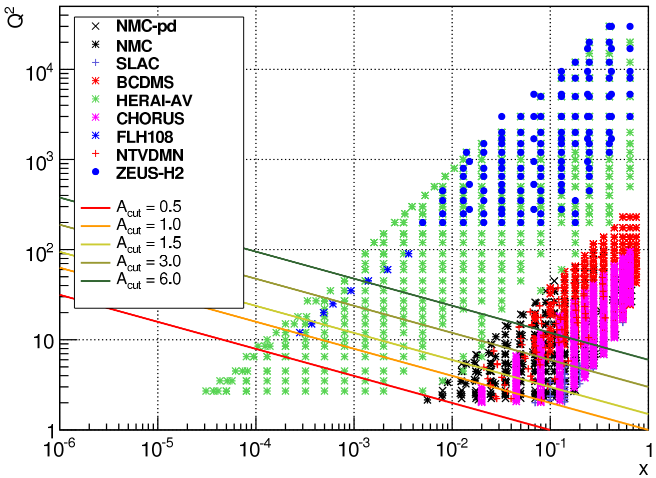


ANALYSIS	DETERMINATION	$\lvert V_{cs} \rvert$
NNPDF1.2	DIRECT FROM GLOBAL PDF ANALYSIS	$0.96 \pm 0.07^{\mathrm{tot}}$
CDHS	LO FROM $\nu N \to \mu^+ \mu^- X$	≥ 0.59 (90% C.L.)
CCFR	NLO FROM $\nu N \to \mu^+ \mu^- X$	≥ 0.74 (90% C.L.)
PDG08	AVERAGE FROM D DECAYS	1.04 ± 0.06
HOCKER	AVERAGE FROM $\nu N \to \mu^+ \mu^- X$	1.04 ± 0.16
DELPHI	DIRECT FROM $W^+ \to c\bar{s}$ DECAYS	$0.94^{+0.32}_{-0.26} \pm 0.13$
PDG08	CKM UNITARITY FIT	0.97334 ± 0.00023

ANALYSIS	DETERMINATION	$\lvert V_{cd} \rvert$
NNPDF1.2	DIRECT FROM GLOBAL PDF ANALYSIS	$0.244 \pm 0.019^{\mathrm{tot}}$
CDHS	LO FROM $\nu N \to \mu^+ \mu^- X$	0.24 ± 0.03
CCFR	NLO FROM $\nu N \to \mu^+ \mu^- X$	$0.232^{+0.017}_{-0.019}$
PDG08	AVERAGE FROM $\nu N \to \mu^+ \mu^- X$	0.230 ± 0.011
PDG08	AVERAGE FROM $D \to K/\pi l \nu$ DECAYS	0.218 ± 0.023
PDG08	CKM UNITARITY FIT	0.2256 ± 0.0010

- FIRST DETERMINATION OF CKM MATRIX ELEMENTS FROM DIS
- NNPDF1.2: ONLY DIS DATA
- MORE ACCURATE (V_{cs}) OR COMPETITIVE (V_{cd}) THAN OTHER DIRECT DETERMINATIONS

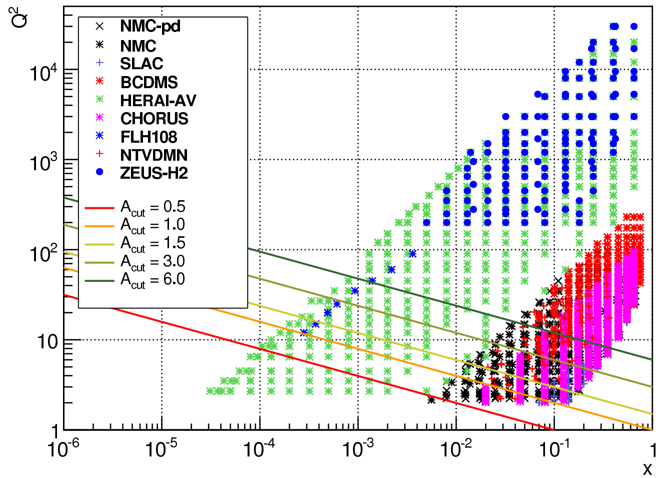
BEYOND DGLAP?
DISCOVERING A NEW QCD EFFECT IN HERA DATA



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?
DISCOVERING A NEW QCD EFFECT IN HERA DATA

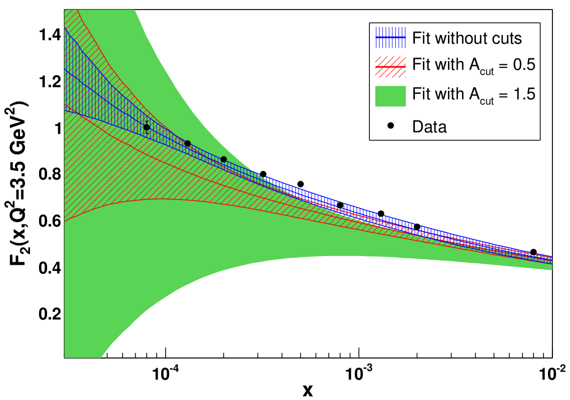


IDEA: (Gélis, 2008, $\Rightarrow$ Caola, s.f. ,Rojo 2010)

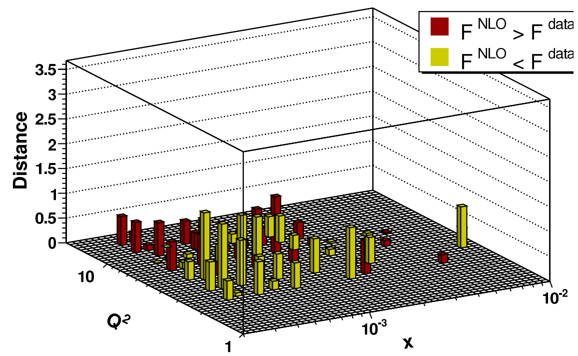
- CUT OUT DATA IN THE "DANGEROUS" (SMALL τ) REGION
- DETERMINE PDFs IN THE "SAFE" (LARGE x AND Q^2) REGION
- EVOLVE BACKWARDS AND COMPARE TO DATA

OLD HERA DATA

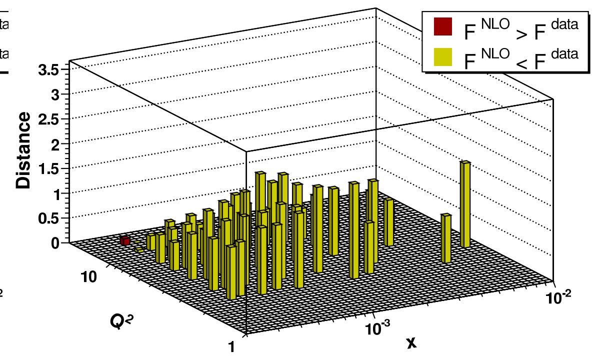
BACKWARD EV. VS DATA



DAT/TH DIST: NO CUT



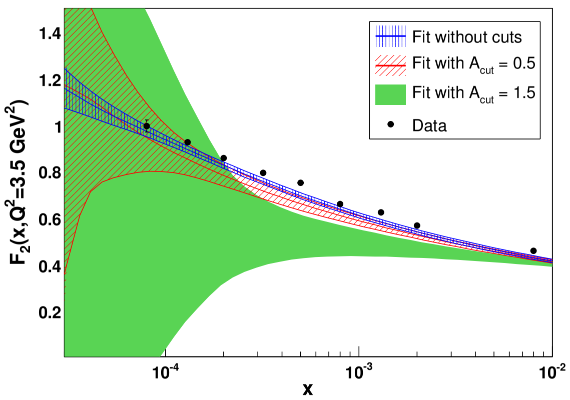
DAT/TH DIST: CUT



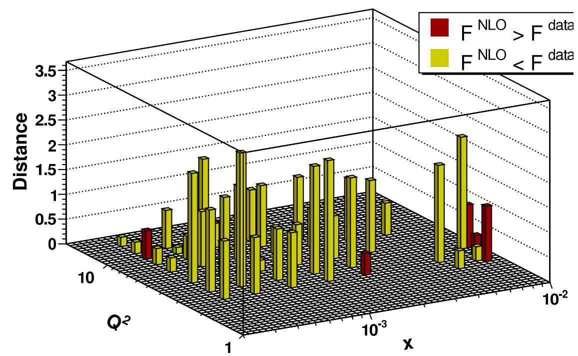
- BACKWARD EVOLVED FIT LIES SYSTEMATICALLY BELOW DATA
- DATA AT LOW x AND Q^2 SHOW LESS EVOLUTION THAN PREDICTED BY NLO DGLAP
- IF LOW x AND Q^2 DATA INCLUDED, THE FIT COMPENSATES READJUSTING PDFs

NEW (COMBINED) HERA DATA

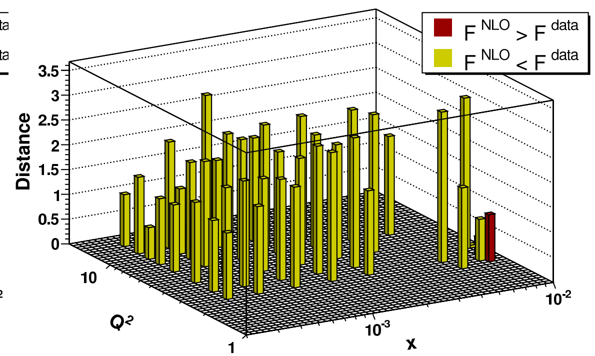
BACKWARD EV. VS DATA



DAT/TH DIST: NO CUT



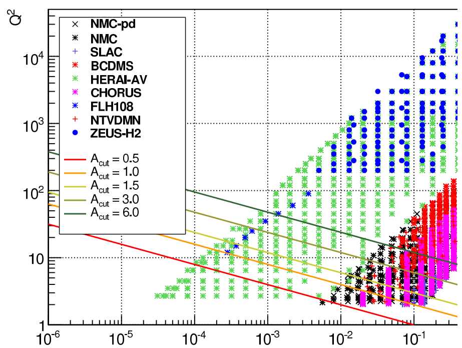
DAT/TH DIST: CUT

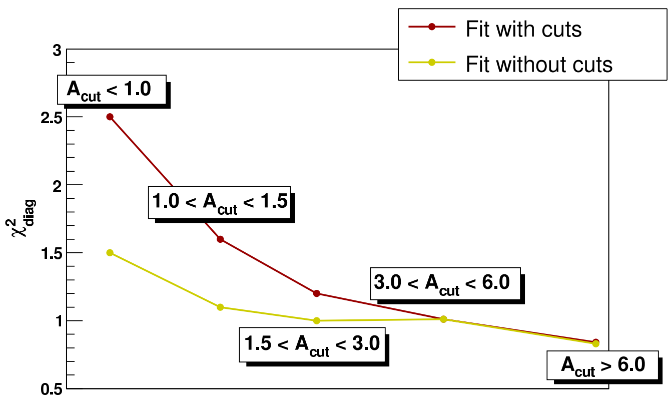


- DATA AT LOW x AND Q^2 SHOW LESS EVOLUTION THAN PREDICTED BY NLO DGLAP
- BACKWARD EVOLVED FIT LIES SYSTEMATICALLY BELOW DATA
- WITH MORE PRECISE DATA, THE FIT NO LONGER MANAGES TO COMPENSATE BY READJUSTING THE PDFS: EVEN FULL FIT LIES BELOW DATA

DETERIORATION IN FIT QUALITY:

$\langle d^2 \rangle$ VS τ SLICES



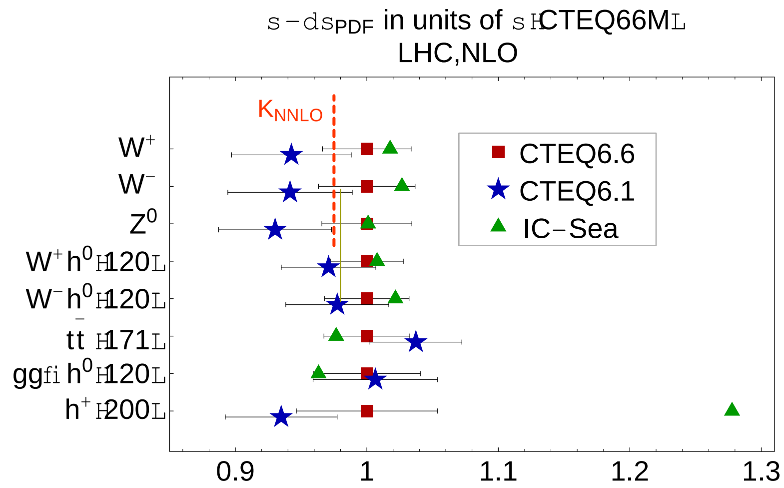


- QUALITY OF UNCUT FIT DETERIORATES IN LOW τ REGIONS
- QUALITY OF CUT FIT INCREASINGLY POOR AS τ DECREASES
- NOTE PDF UNCERTAINTY INCLUDED IN DISTANCE

WHERE ARE WE GOING?

THE PROBLEM OF HEAVY QUARKS

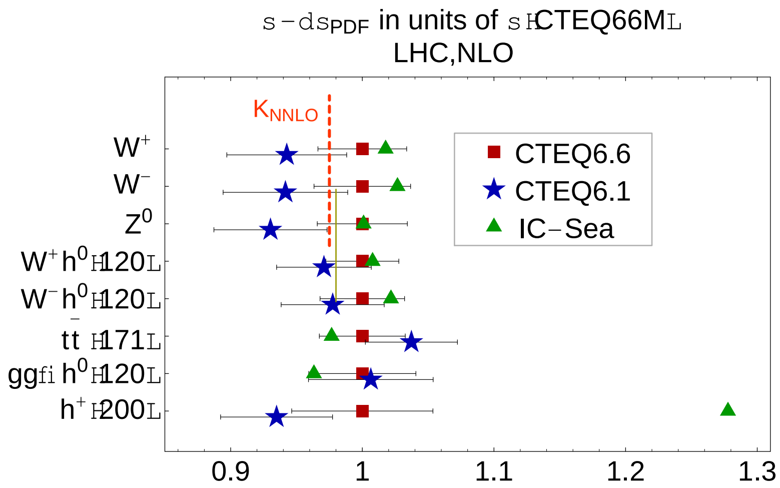
- MANY FITS (CTEQ<6, NNPDF, ALEKHIN<09) TREAT CHARM AS MASSLESS ABOVE THRESHOLD ⇒ “ZMVFN” SCHEME
- TR/TR’ PROCEDURE IMPLEMENTED SINCE ’98 IN MRST
- WHEN CTEQ IMPLEMENTED ACOT IN 2008, SURPRISING CHANGE CTEQ61→CTEQ6.5 IN σ_W, & AGREEMENT WITH MRST SPOILED (LATER RESTORED)



(Nadolsky et al., 2008)

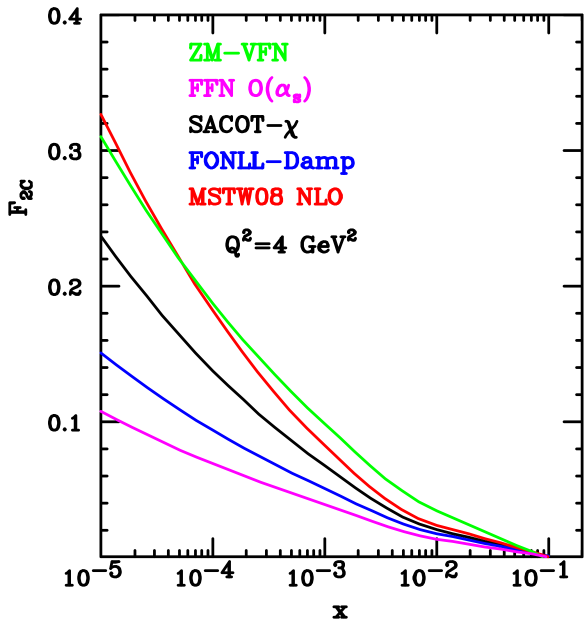
THE PROBLEM OF HEAVY QUARKS

- MANY FITS (CTEQ<6, NNPDF, ALEKHIN<09) TREAT CHARM AS MASSLESS ABOVE THRESHOLD ⇒ "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

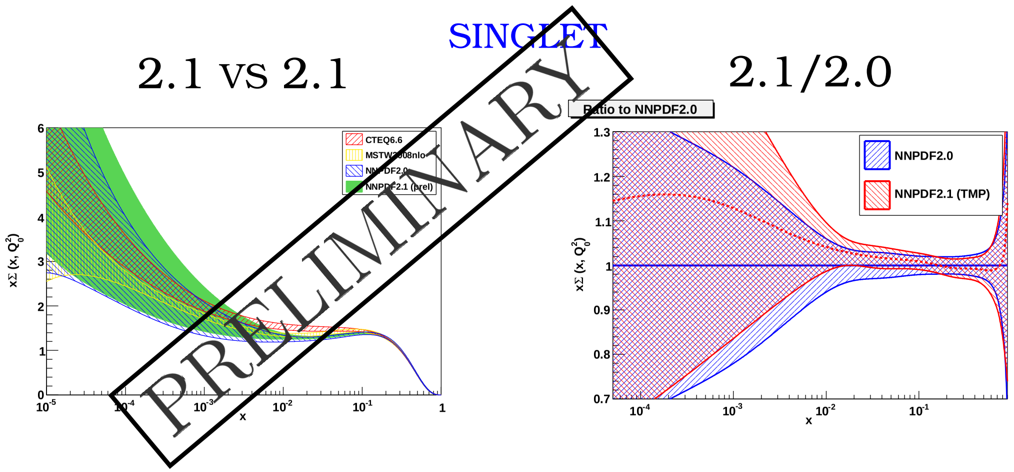
FLEXIBILITY ⇒ STABILITY (PDFs)

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 → 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



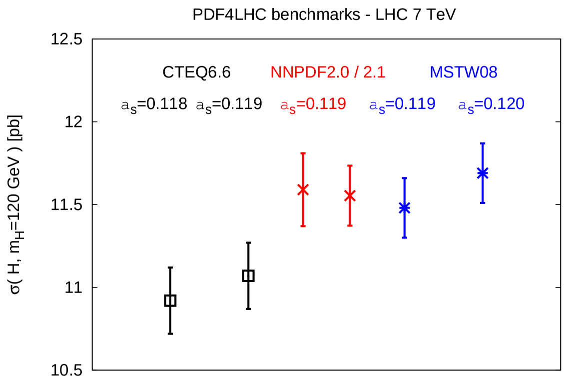
HEAVY QUARKS IN NNPDF

FLEXIBILITY ⇒ STABILITY (STANDARD CANDLES)

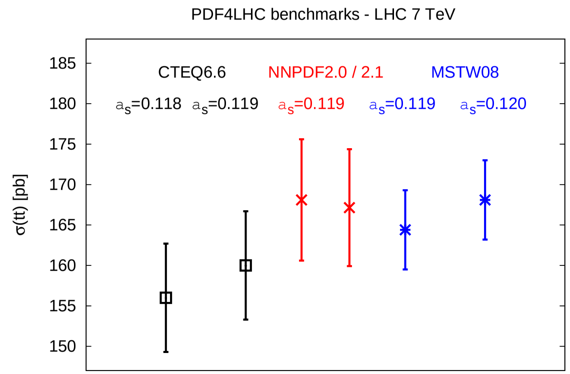
LHC 7 TeV:

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

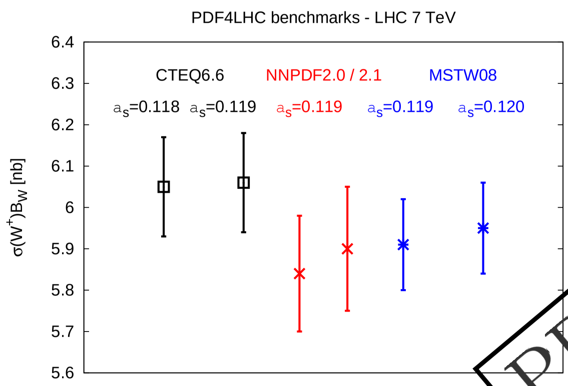
HIGGS 120 GeV



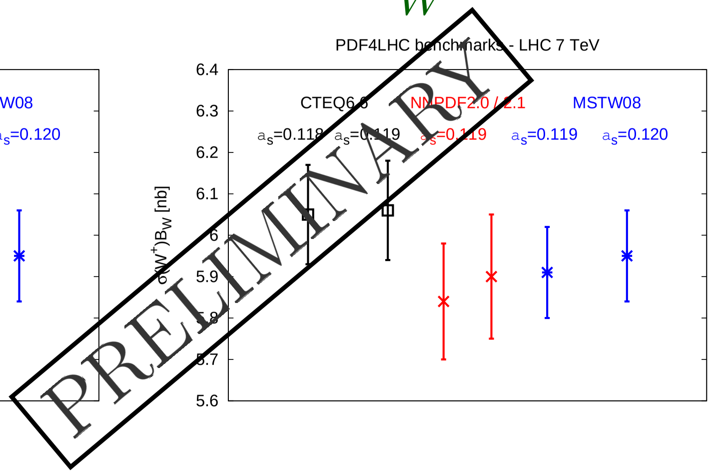
TOP



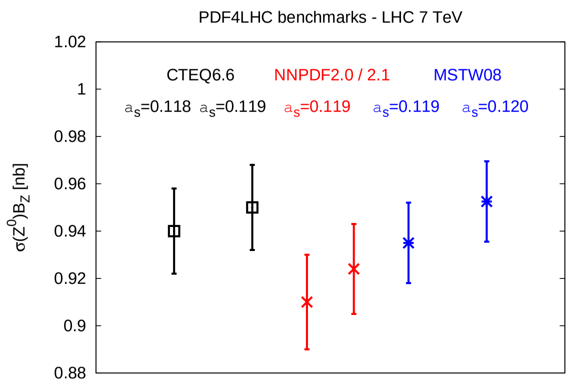
W^+



W^-



Z



LEADING ORDER PDFs

- LO PDFs NEEDED FOR MONTE CARLOS
- MSTW+CTEQ LO $\rightarrow$ 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 $\to$ 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

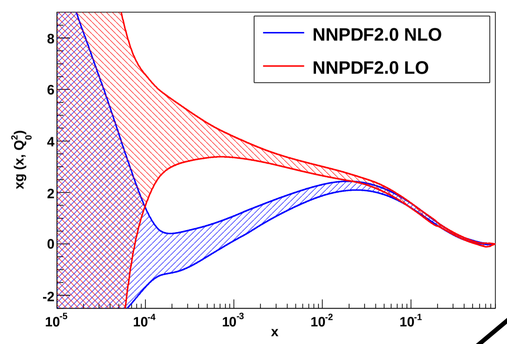
FLEXIBILITY $\Rightarrow$ STABILITY (PDFs)

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

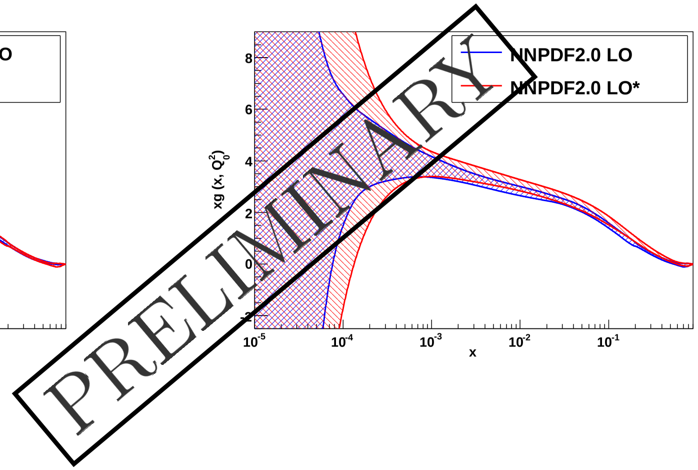
- LO FIT MARGINALLY WORSE THAN NLO
- NLO α_s IN LO FIT (LOMOD) $\to$ NO EFFECT
- MOMENTUM SUM RULE OK AT 1.5σ IN LO FIT
- LO* VS. LO/LOMOD $\Rightarrow$ SIZABLY LARGER UNCERTAINTY

THE GLUON

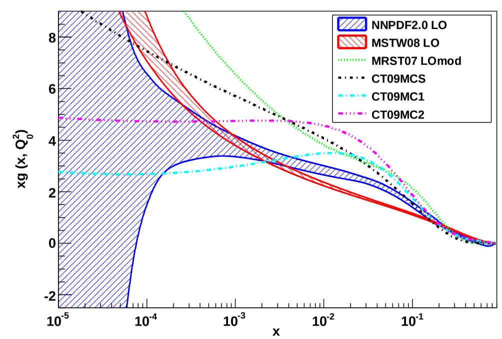
LO VS. NLO



LO VS. LO*



NNPDF VS. MSTW/TEA



LEADING ORDER PDFs

- LO PDFs NEEDED FOR MONTE CARLOS
- MSTW+CTEQ LO $\rightarrow$ 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

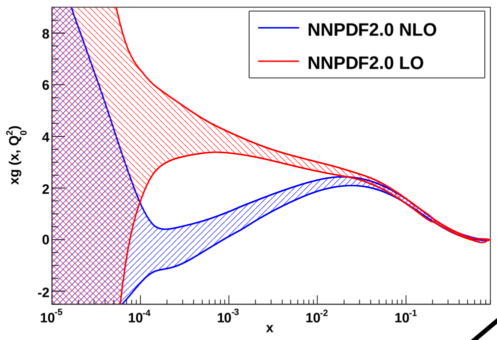
FLEXIBILITY $\Rightarrow$ STABILITY (PDFs)

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

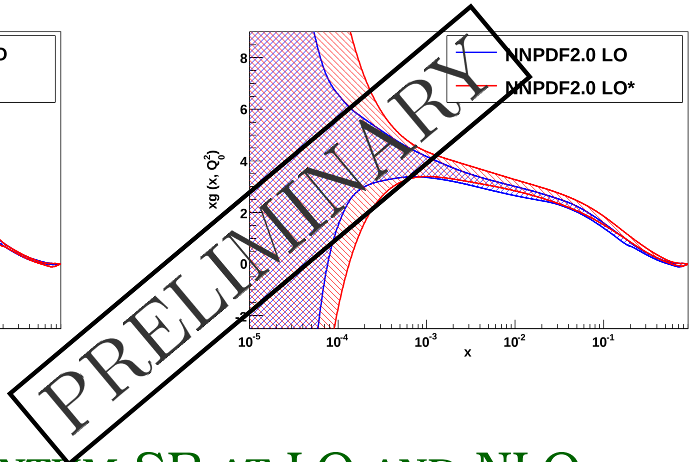
- LO FIT MARGINALLY WORSE THAN NLO
- NLO α_s IN LO FIT (LOMOD) $\rightarrow$ NO EFFECT
- MOMENTUM SUM RULE OK AT 1.5σ IN LO FIT
- LO* VS. LO/LOMOD $\Rightarrow$ SIZABLY LARGER UNCERTAINTY

THE GLUON

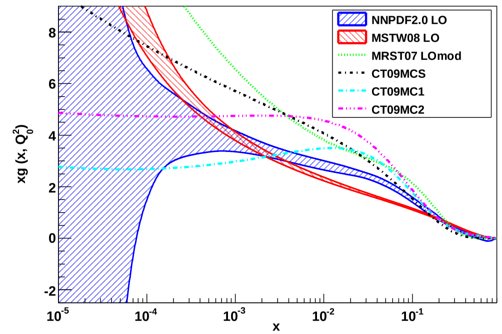
LO VS. NLO



LO VS. LO*



NNPDF VS. MSTW/TEA



NO NEED FOR LO*

DATA TEST THE MOMENTUM SR AT LO AND NLO

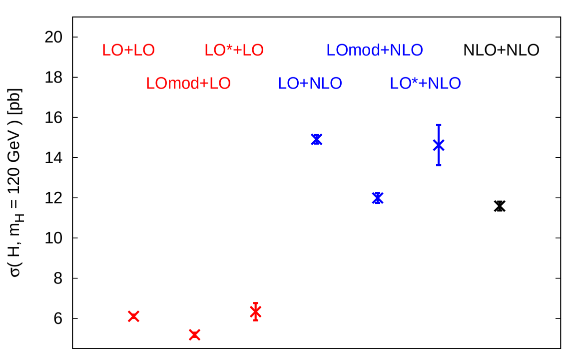
NNPDF2.0 AT LEADING ORDER

FLEXIBILITY ⇒ STABILITY (STANDARD CANDLES)

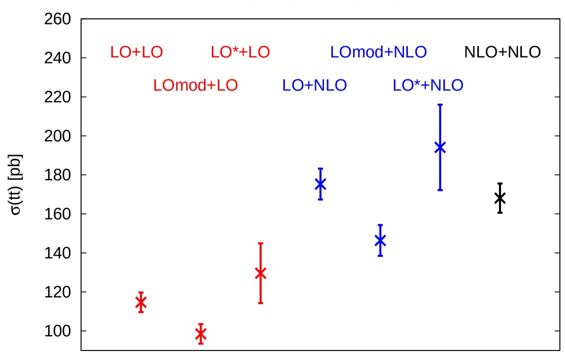
LHC 7 TeV:

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

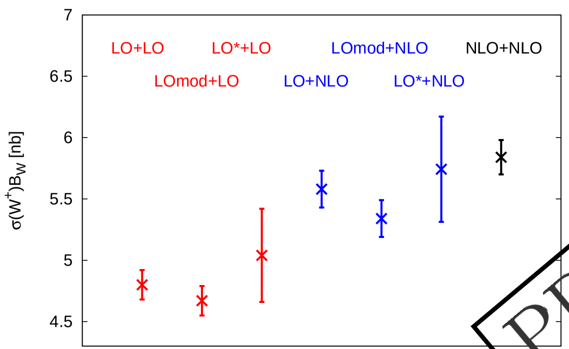
HIGGS 120 GeV



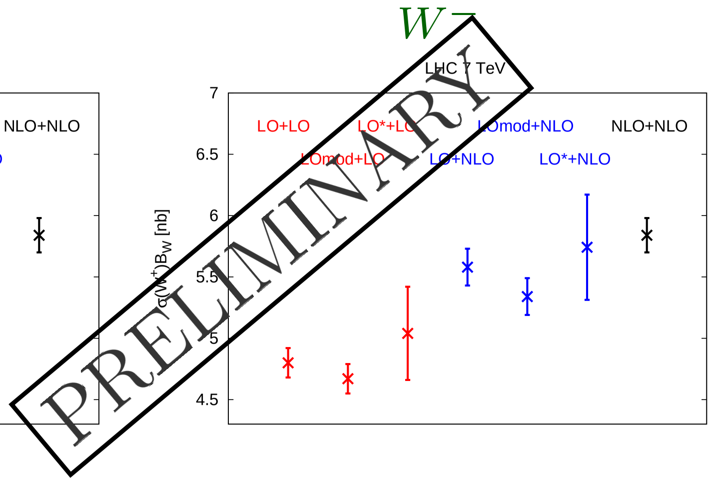
TOP



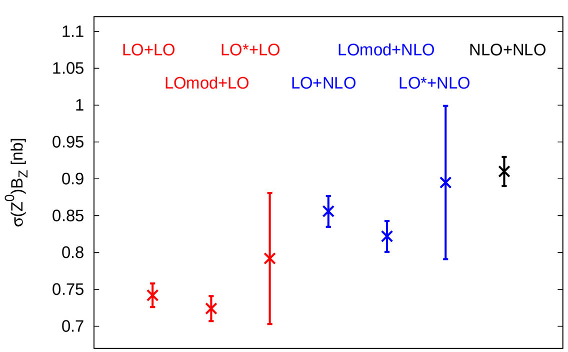
W^+



W^-

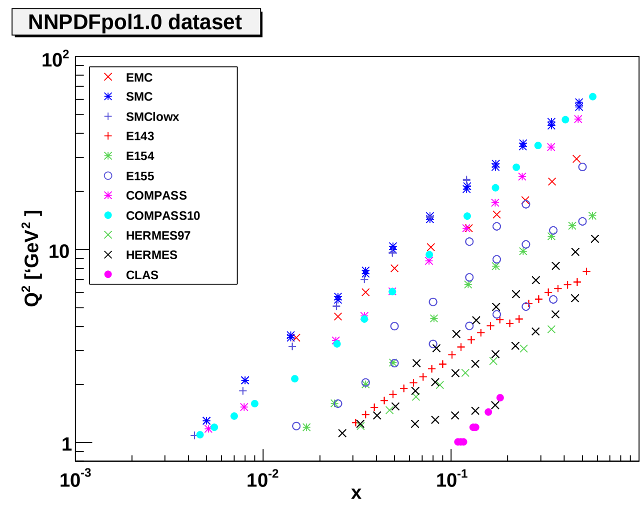


Z

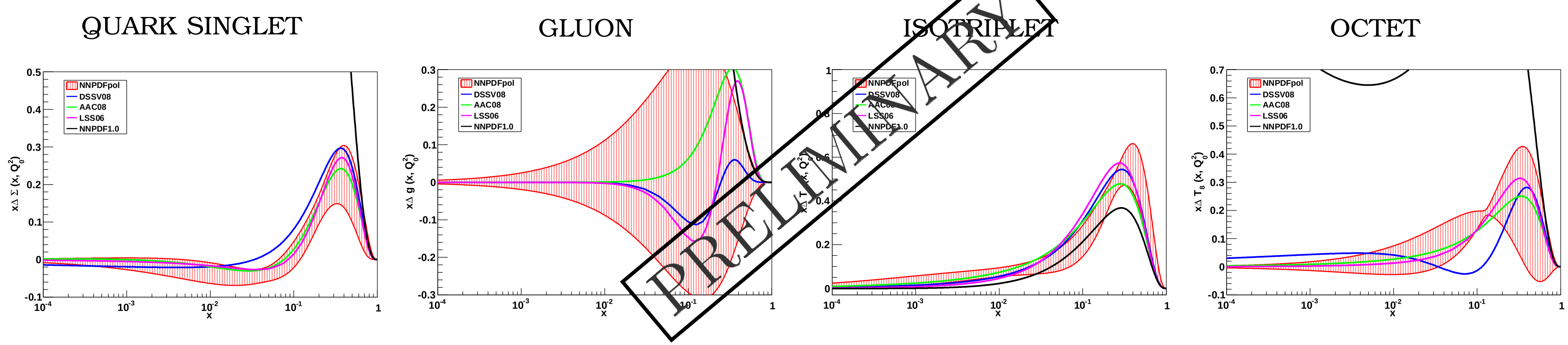


POLARIZED PDFS

NNPDF1.0POL



- ALL INCLUSIVE DIS DATA INCLUDED ($N_{\rm dat} \approx 250$); RHIC+COMPASS FORTCOM.
- FOUR INDEPENDENT PDFs PARAMETRIZED (4 $\otimes$ 37 PARMS.)
- FASTKERNEL IMPLEMENTED
- PRELIMINARY: NO STOPPING YET (FIXED-LENGTH TRAINING)



- GOOD AGREEMENT WITH OTHER FITS, GLUON UNCERTAINTY MUCH LARGER
- MORE WORK NEEDED FOR PRECISION STUDIES

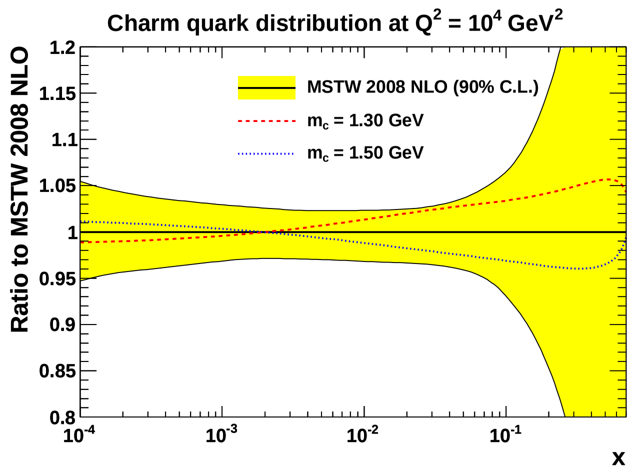
WHAT NEXT?
THEORETICAL UNCERTAINTIES

α_s UNCERTAINTIES: CAN BE COMBINED IN QUADRATURE WITH PDF UNCERTAINTIES (Lai et al, CTEQ, 2010); AVAILABLE FOR CTEQ, HERAPDF, MSTW, NNPDF

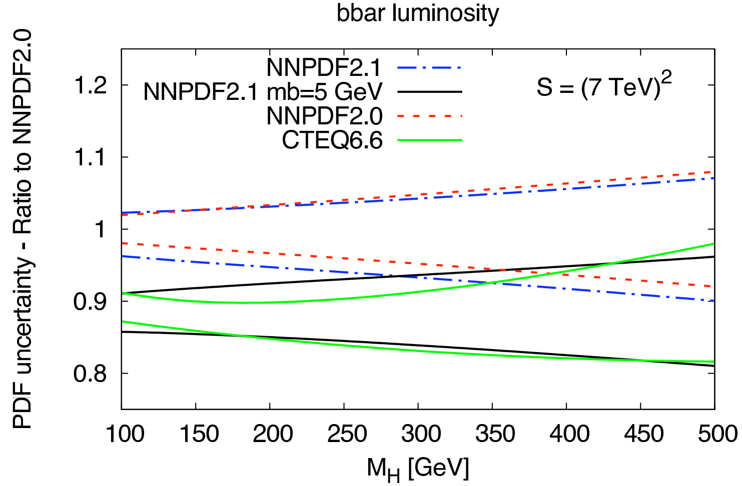
IN PROGRESS:

DEPENDENCE ON c & b MASS

m_c: MSTW08



m_b: NNPDF, CTEQ



NEEDED:

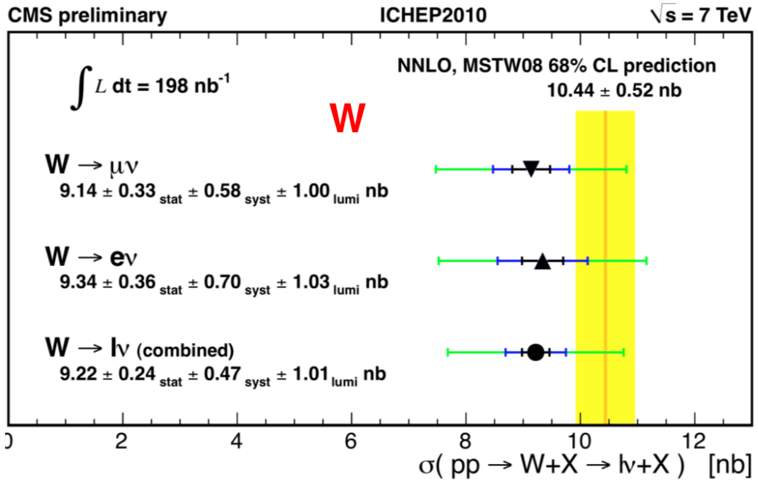
- ORDERS: RENORMALIZATION AND FACTORIZATION SCALE VARIATION
- HEAVY QUARKS: MATCHING SCHEME VARIATION
- RESUMMATION: SMALL x REGION AT HERA, LARGE τ REGION FOR FIXED-TARGET DY

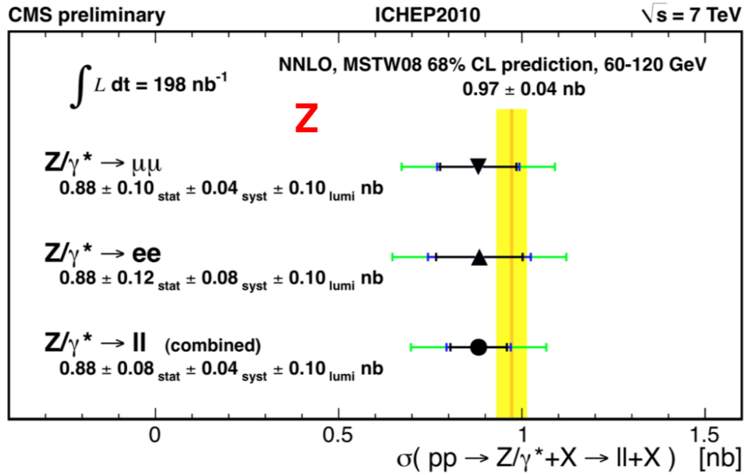
THE NNPDF TIMELINE

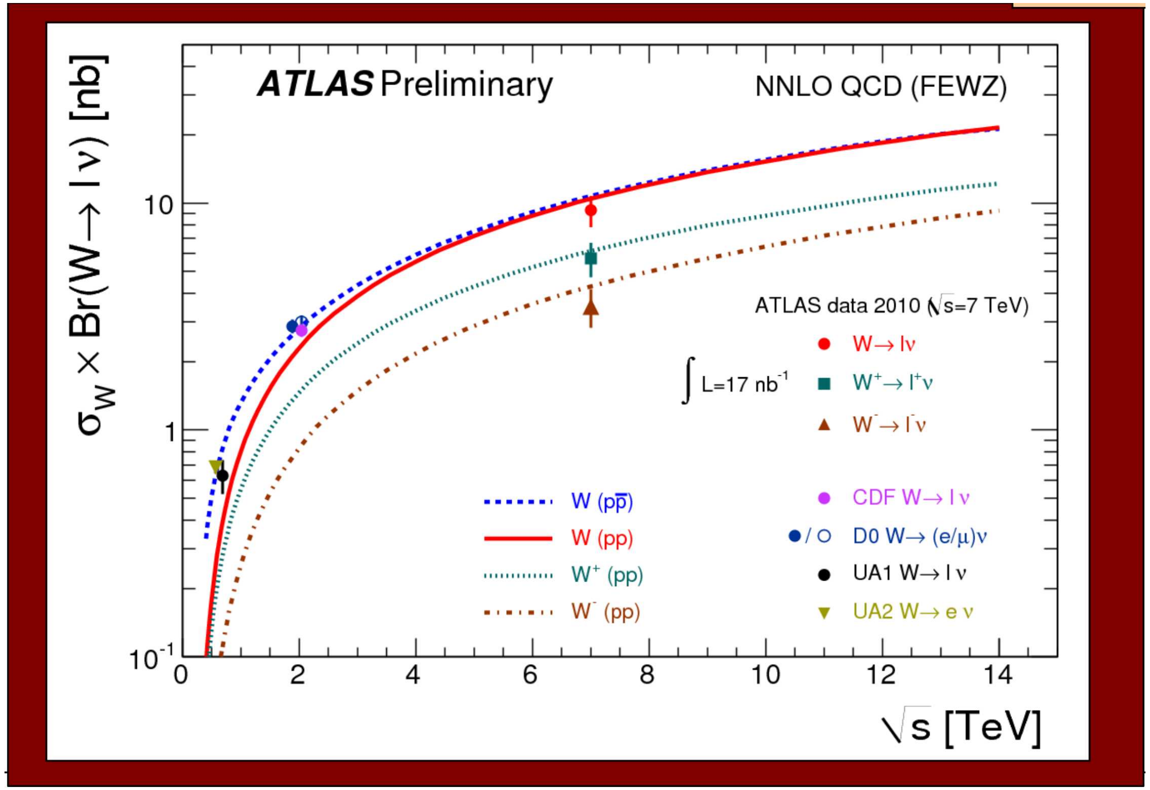
- 2002 FIT OF NUCLEON STRUCTURE FUNCTION TO TWO DATASETS
- 2005 FIT OF NUCLEON STRUCTURE FUNCTION TO MANY DATASETS
- 2007 FIT OF ONE SINGLE (ISOTRIPLET) PARTON DISTRIBUTION TO TWO DATASETS
- 2008 FIRST FULL PARTON SET: NNPDF1.0 (WORLD DIS DATA)
- 2009 PARTON SET WITH ACCURATE STRANGENESS: NNPDF1.2 (DIS+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

CONCLUSION

THIS IS JUST THE BEGINNING!



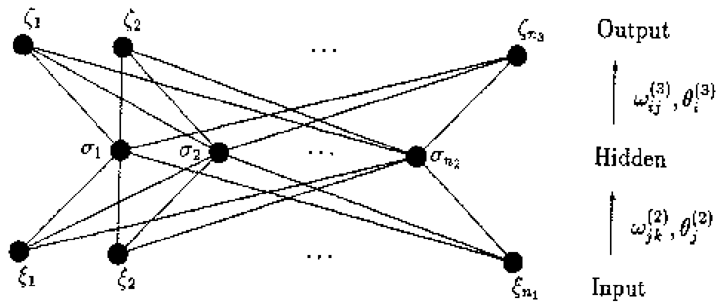


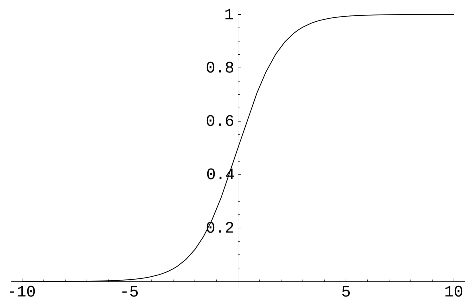


EXTRAS

NEURAL NETWORKS

A NONLINEAR FUNCTIONAL FORM





MULTILAYER FEED-FORWARD NETWORKS

- Each neuron receives input from neurons in preceding layer and feeds output to neurons in subsequent layer
- Activation determined by weights and thresholds

  $$\xi_i = g\left(\sum_j \omega_{ij}\xi_j - \theta_i\right)$$

- Sigmoid activation function

  $$g(x) = \frac{1}{1+e^{-\beta x}}$$

EXAMPLE: A 1-2-1 NN

$$f(x) = \frac{1}{1+e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)} - x\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)} - x\omega_{21}^{(1)}}}}}$$

THANKS TO NONLINEAR BEHAVIOUR,
ANY FUNCTION CAN BE REPRESENTED BY A SUFFICIENTLY BIG NEURAL NETWORK

HESSIAN ERROR ESTIMATES

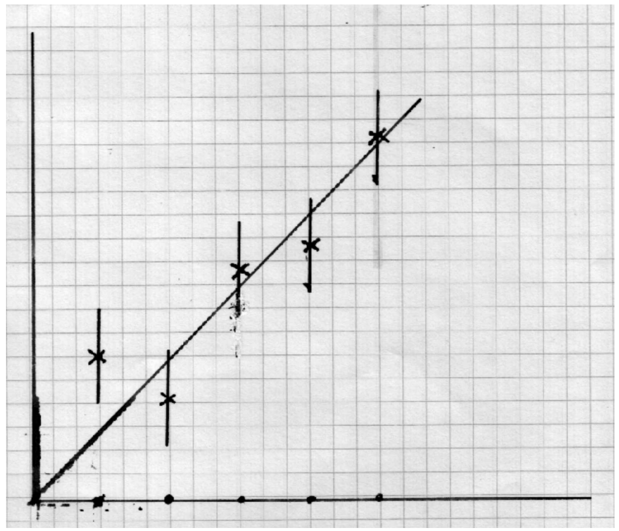
HYPOTESIS TESTING VS. PARAMETER FITTING

"PARADOX"

- THE STANDARD DEVIATION OF χ^2 FOR $N_{\rm dat}$ DATA $\sigma_{\chi^2} = \sqrt{2N_{\rm dat}}$
  HYPOTESIS-TESTING RANGE: COMPARE $\Delta\chi^2 = \chi^2 - \langle\chi^2\rangle$ TO $\sigma^2_{\chi^2}$.
  IF TOO LARGE, SOMETHING WRONG WITH THEORY (OR DATA)
- BUT THE ONE-σ RANGE FOR A PARM. OF THE THEORY IS THE CURVE $\chi^2 - \chi^2_{\rm min} = 1$
  PARAMETER-FITTING RANGE: UNIT DEVIATION FROM THE PARAMETRIC MINIMUM $\chi^2_{\rm min}$

WHY?

- CONSIDER DEVIATIONS Δ_i FROM LINEAR FIT $y = x + k$; DETERMINE INTERCEPT k AS FREE PARAMETER
- IF STANDARD DEVIATION FOR EACH Δ_i IS σ_Δ, THEN AVERAGE SQUARE DEVIATION IN UNITS OF σ_Δ FOR $N_{\rm dat}$ DATA: $\sigma_{\chi^2} = N_{\rm dat}$
- BEST-FIT INCTERCEPT: $k = \langle\Delta_i\rangle$
- UNCERTAINTY ON IT: $\sigma_k = \frac{\sigma_\Delta}{N_{\rm dat}}$
- IF $\Delta k = \sigma_k$, THEN $\Delta\chi^2 = 1$

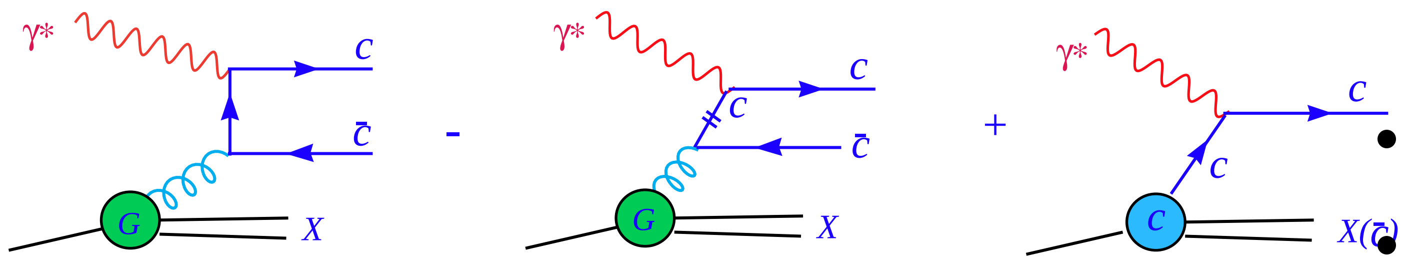


HEAVY QUARKS IN DIS

- IN $\overline{\text{MS}}$ SCHEME, $n_f = 6$ IN LOOPS, α_s RUNNING AND DGLAP EQNS AT ALL SCALES
  $\Rightarrow$ FOR $Q^2 >> m_q^2$, NEGLECT THE MASS
- WHEN $Q^2 << m_q^2$ A DECOUPLING SCHEME MORE CONVENIENT: LOOPS SUBTRACTED AT ZERO MOMENTUM, $n_f = n_l$ IN α_s RUNNING AND DGLAP EQNS
  $\Rightarrow$ FOR $Q^2 >> m_q^2$, NEGLECT THE HEAVY QUARK
- WHAT HAPPENS WHEN $Q^2 \approx m_q^2$?

MATCHED SCHEMES: ACOT (Aivazis, Collins, Olness, Tung, 1988, 1994)

$m_c \neq 0$, LO charm radiation — $m_c = 0$, LO charm radiation + $m_c = 0$ charm pdf



- USE $\overline{\text{MS}}$ FOR $Q^2 > m_q^2$ WITH FULL MASS DEP. RETAINED
- KEEP ALL FLAVOURS IN RUNNING, DGLAP
- SUBTRACT DOUBLE COUNTING

SIMPLIFIED SACOT:

EVEN IN MASSIVE CONTRIBUTION, NEGLECT m_q IN FINAL STATE

MATCHED SCHEMES: TR (Thorne, Roberts, 1998, 2008)

- SWITCH OFF HQ FOR $Q^2 < m_q^2$
- USE MASSLESS APPROX FOR $Q^2 > m_q^2$
- ADD MASSIVE TERMS AND ENFORCE CONTINUITY AT THRESHOLD VIA SUBL. TERMS

MATCHED SCHEMES: FONLL

(Cacciari, Greco, Nason, 1998; for DIS s.f., Laenen, Rojo, Nason, 2010)

- USE $\overline{\mathrm{MS}}$ (MASSLESS) PARTONS
- COMPUTE MASSIVE CONTRIBUTIONS IN THE DECOUPLING SCHEME, BUT EXPRESS EVERYTHING IN TERMS OF $\overline{\mathrm{MS}}$ PARTONS
- ADD MASSIVE EXPRESSION TO THE MASSLESS ONE, SUBTRACT DOUBLE COUNTING (TRIVIAL, AS EVERYTHING EXPRESSED IN SAME SCHEME)

$$F^{(n_l)}(x,Q^2) = x\int_x^1 \frac{dy}{y}\sum_{i=q,\bar{q},g} B_i\left(\frac{x}{y},\frac{Q^2}{m^2},\alpha_s^{(n_l+1)}(Q^2)\right) f_i^{(n_l+1)}(y,Q^2),$$

$$F^{(n_l,0)}(x,Q^2) \equiv x\int_x^1 \frac{dy}{y}\sum_{i=q,\bar{q},g} B_i^{(0)}\left(\frac{x}{y},\frac{Q^2}{m^2},\alpha_s^{(n_l+1)}(Q^2)\right) f_i^{(n_l+1)}(y,Q^2); \quad \lim_{m\to 0}\left[B_i(x,\frac{Q^2}{m^2}) - B_i^{(0)}\left(x,\frac{Q^2}{m^2}\right)\right] = 0$$

$$F^{\mathrm{FONLL}}(x,Q^2) \equiv F^{(d)}(x,Q^2) + F^{(n_l)}(x,Q^2); \qquad F^{(d)}(x,Q^2) \equiv \left[F^{(n_l+1)}(x,Q^2) - F^{(n_l,0)}(x,Q^2)\right]$$

THE PROBLEM OF DAMPING TERMS

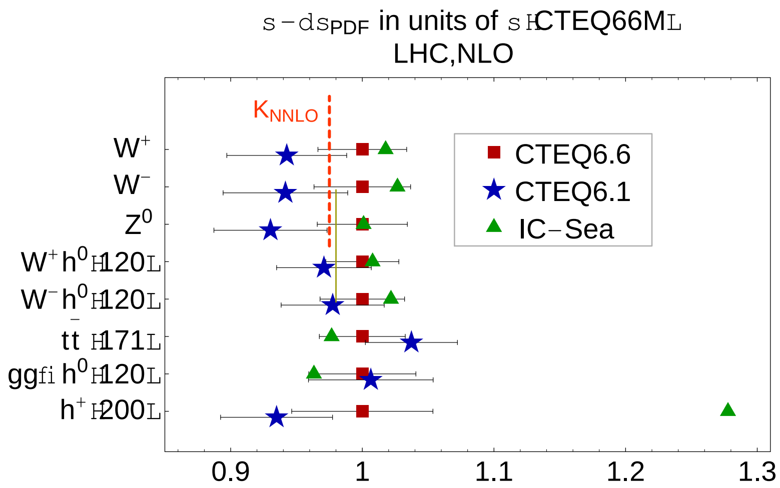
- IN ANY SCHEME, DGLAP RESUMMATION PRODUCES TERMS $\sim \alpha_s(Q^2)\ln\frac{Q^2}{m_q^2}$ TO ALL ORDERS IN $\alpha_s(Q^2)$
- MASS CORRECTIONS TO SUCH TERMS ARE PROVIDED AT LOW ORDERS IN $\alpha_s(Q^2)$, BUT NOT TO HIGHER ORDERS
- THESE TERMS ARE COMPLETELY INACCURATE WHEN Q^2 IS JUST ABOVE m_q^2 AND CAN BE NON–NEGLIGIBLE IN PRACTICE
- SOLUTION: KILL THESE TERMS WITH A SUITABLE DAMPING PRESCRIPTION

THE IMPACT ON PHENOMENOLOGY

- MANY FITS (CTEQ<6, NNPDF, ALEKHIN<09) TREAT CHARM AS MASSLESS ABOVE THRESHOLD $\Rightarrow$ "ZMVFN" SCHEME
- TR/TR' PROCEDURE IMPLEMENTED SINCE '98 IN MRST

THE IMPACT ON PHENOMENOLOGY

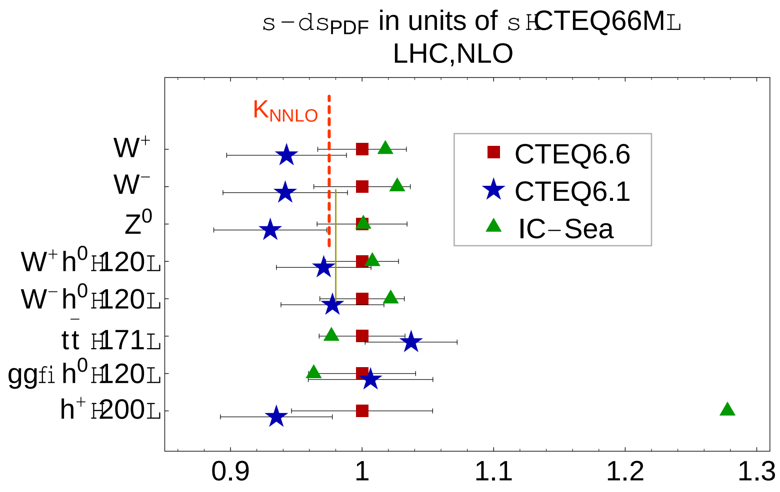
- MANY FITS (CTEQ<6, NNPDF, ALEKHIN<09) TREAT CHARM AS MASSLESS ABOVE THRESHOLD $\Rightarrow$ “ZMVFN” SCHEME
- TR/TR' PROCEDURE IMPLEMENTED SINCE '98 IN MRST
- WHEN CTEQ IMPLEMENTED ACOT IN 2008, SURPRISING CHANGE CTEQ61$\rightarrow$CTEQ6.5 IN σ_W, & AGREEMENT WITH MRST SPOILED (LATER RESTORED)



(Nadolsky et al., 2008)

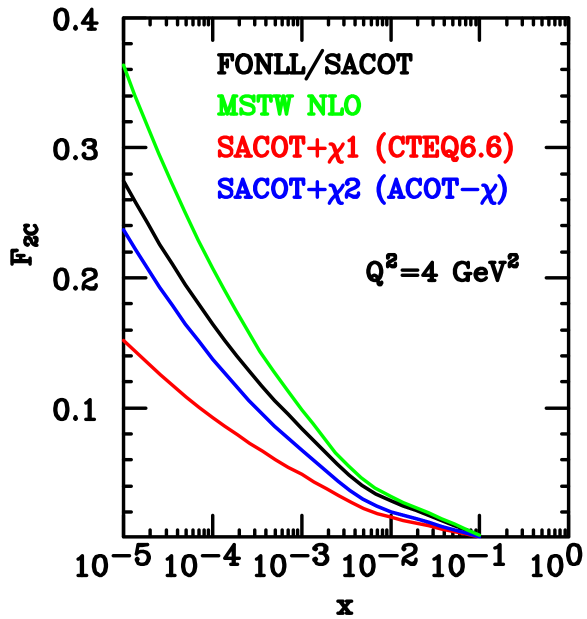
THE IMPACT ON PHENOMENOLOGY

- 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

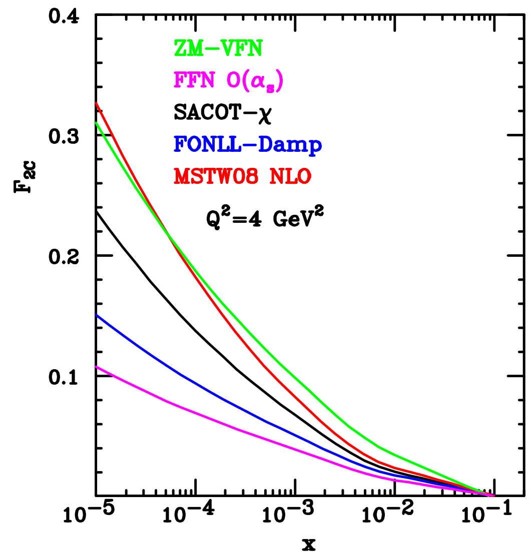


(Rojo et al., 2010)

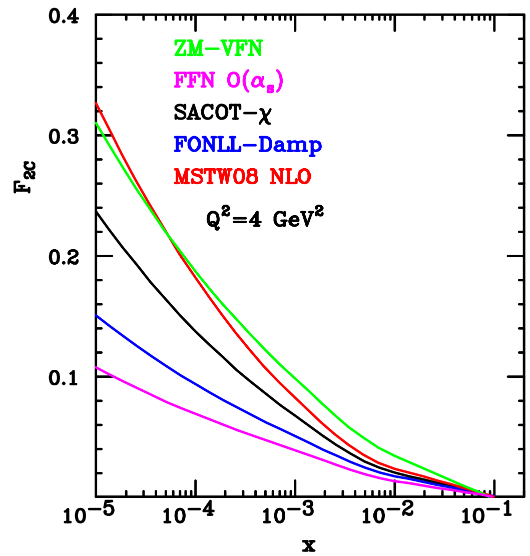
- TR, FONLL AND ACOT FOR DIS BENCHMARKED AT NLO AND NNLO
- $O(\alpha_s)$ FONLL, ACOT COINCIDE EXACTLY, TR' DIFFERS BY SUBLEADING $O(\alpha_s^2(m_c))$ Q^2–INDEP. TERM
- DIFFERENCES BETWEEN DAMPING PRESCRIPTIONS SIZABLE

THE PROBLEM OF DAMPING TERMS:
PHENOMENOLOGY

- IMPACT OF SUBLEADING TERMS SIZABLE CLOSE TO THRESHOLD
- 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)

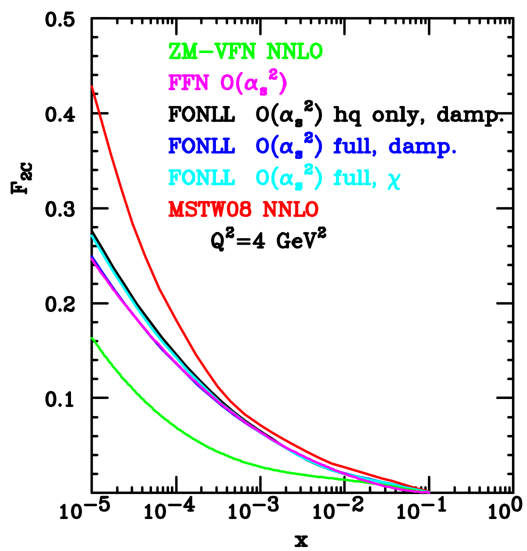


THE PROBLEM OF DAMPING TERMS:
PHENOMENOLOGY



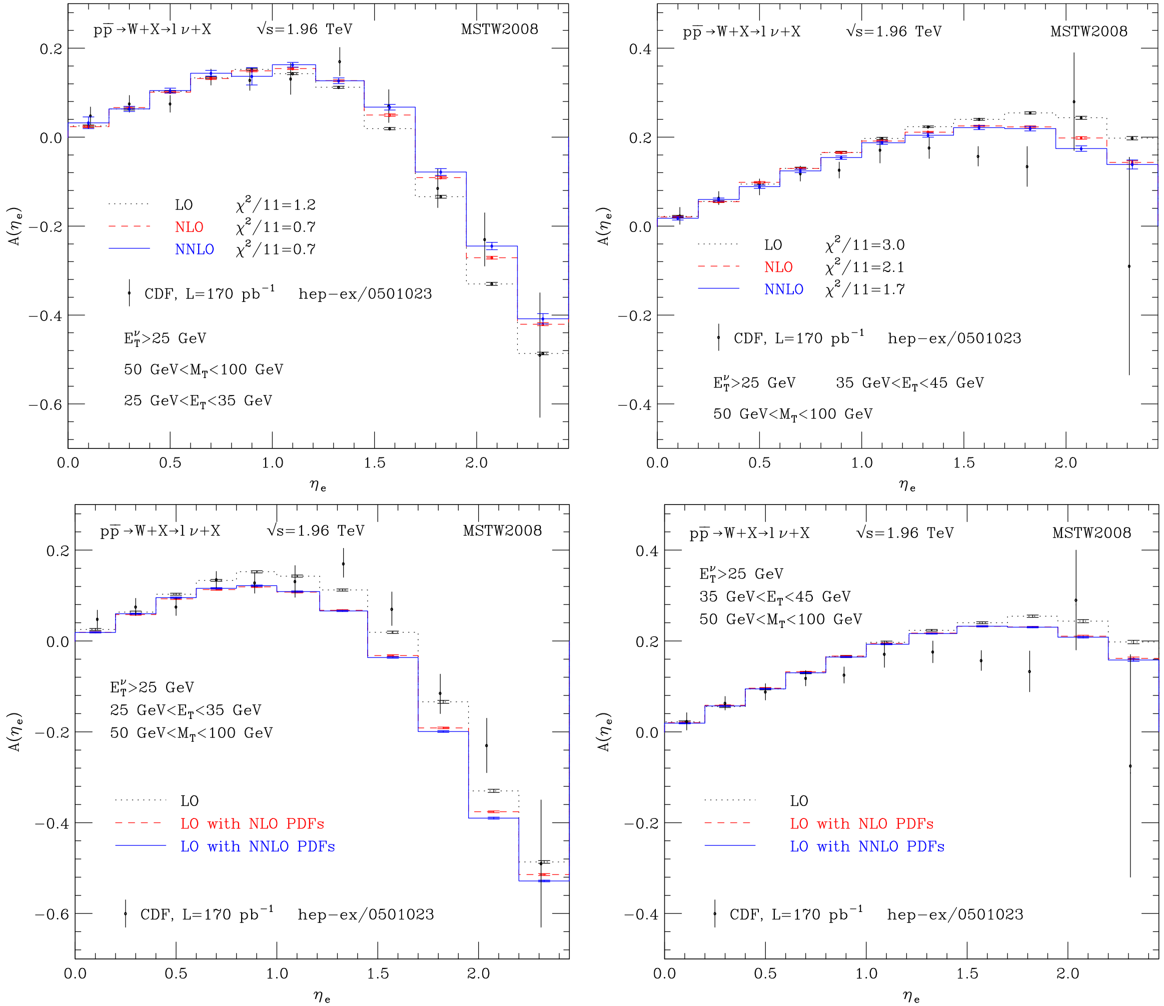
- IMPACT OF SUBLEADING TERMS SIZABLE CLOSE TO THRESHOLD
- DIFFERENCE BETWEEN DIFFERENT PRESCRIPTIONS (ACOT-χ–SCALING, FONLL-DAMPING, MSTW-MATCHING) AS LARGE AS DIFFERENCE BETWEEN FFN (NO DGLAP RESUMMATION FOR CHARM) AND ZMVFN (NO CHARM MASS)

THE SOLUTION:
GO UP ONE ORDER



- IF EVERYTHING AT ONE EXTRA ORDER IN α_s, DIFFERENCES MINOR
- IN FONLL, CAN COMBINE $O(\alpha_s^2)$ TREATMENT OF HQ WITH STANDARD NLO $O(\alpha_s)$ TREATMENT OF LIGHT QUARKS $\Rightarrow$ EXCELLENT APPROX TO FULL RESULT (s.f., Laenen, Nason, Rojo 2010)
- RECENT PROGRESS: $O(\alpha^3)$ MASSLESS LIMIT OF HQ PRODUCTION COEFF. FCTNS. COMPUTED (Bierenbaum, Blümlein, Klein, 2009)

NNLO CORRECTIONS: THE W CHARGE ASYMMETRY

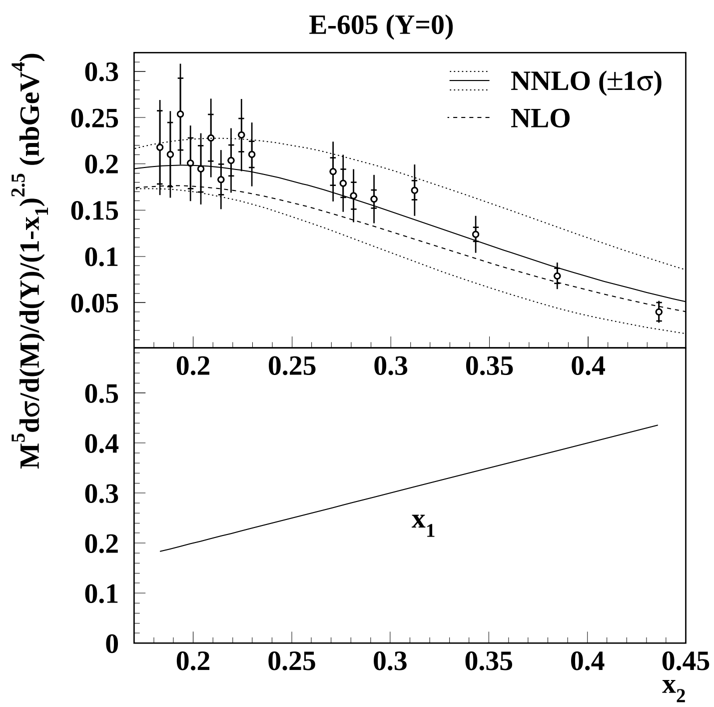


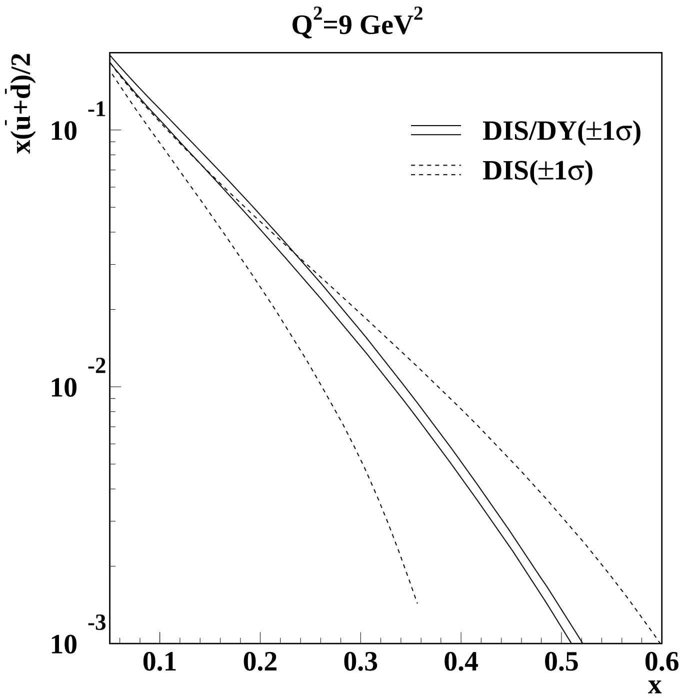
Catani, Ferrera, Grazzini, 2010

- NNLO CORRECTIONS VISIBLY IMPROVE AGREEMENT WITH DATA
- EFFECT ON MATRIX ELEMENT COMPARABLE TO EFFECT ON PDFs, BUT IN DIFFERENT REGIONS

NNLO NEEDED FOR STANDARD CANDLES!

NNLO PARTON DISTRIBUTIONS?





Alekhin, Melnikov, Petriello, 2006

- CURRENT GLOBAL PDF FITS ARE NLO
- MSTW08 NNLO TREATS DIS AT NNLO, JETS AT NLO, DRELL-YAN AT LO+K–FACTORS
- HERAPDF+ALEKHIN-SERIES FITS GENUINELY NNLO, BUT SMALLER DATASET
- BUT IMPACT NOT NEGLIGIBLE...