PARTON DISTRIBUTIONS:
THE IMPACT OF LHC DATA

STEFANO FORTE
UNIVERSITÀ DI MILANO & INFN

UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI FISICA

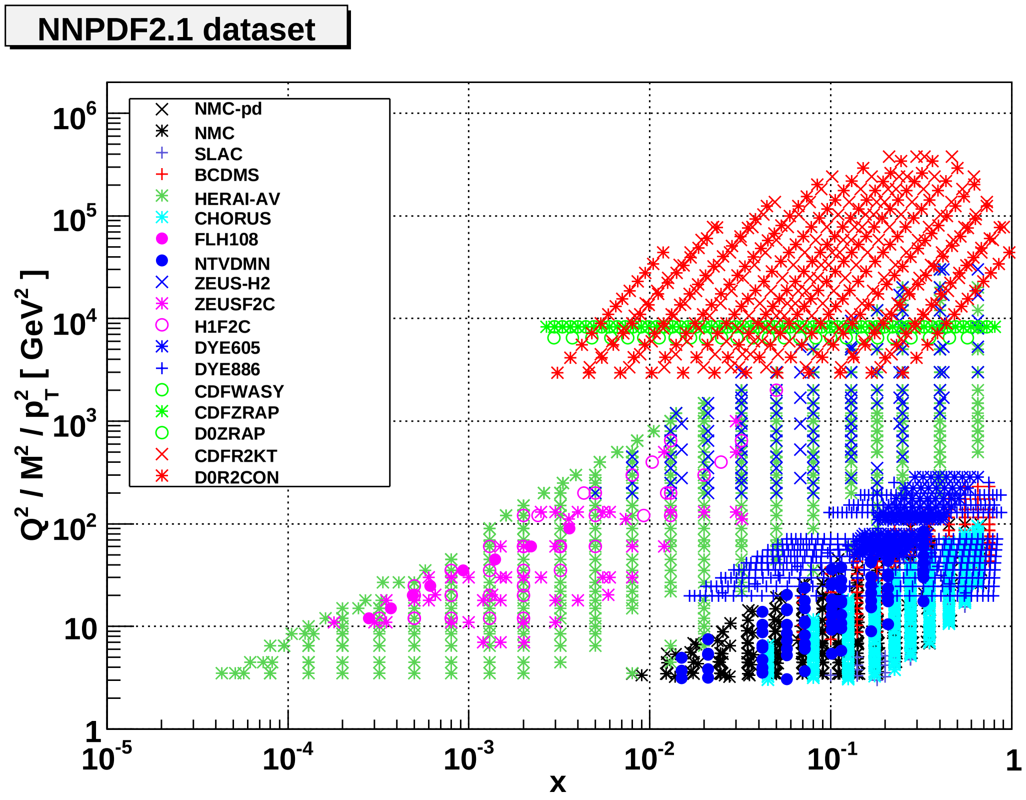
INFN Istituto Nazionale di Fisica Nucleare

SM@LHC

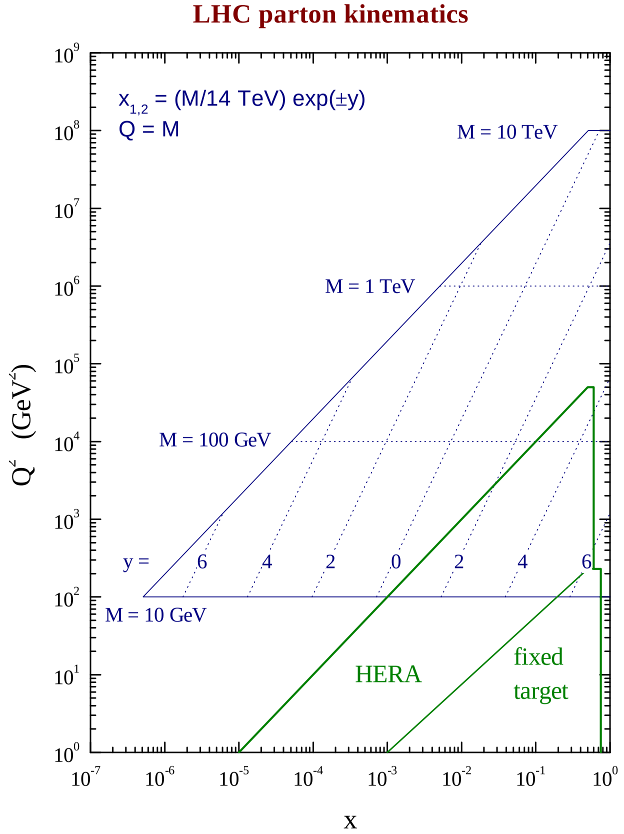
DURHAM, APRIL 12, 2011

PDFs TODAY

A GLOBAL DATASET



LHC kinematics



- CURRENT GLOBAL PDF FITS INCLUDE
  - DIS: NEUTRAL AND CHARGED CURRENT, CHARGED LEPTON AND NEUTRINO BEAMS, INCLUSIVE AND CHARM-TAGGED
  - DRELL-YAN: FIXED TARGET AND COLLIDER, NEUTRAL γ^* AND Z AND CHARGED CURRENT W PRODUCTION
  - INCLUSIVE JETS
- WHERE ARE WE NOW?
- WHERE ARE LHC DATA TAKING US?

SUMMARY

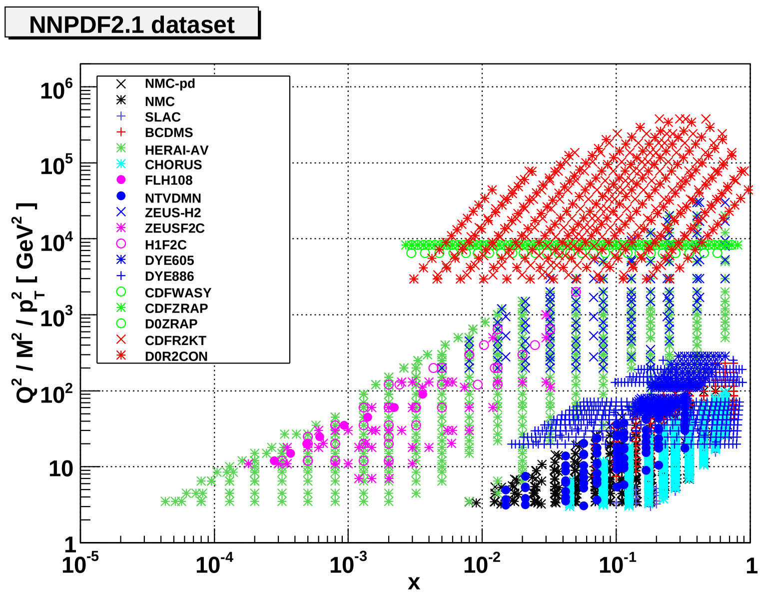
- STANDARD CANDLES
  - WHERE DO WE STAND?
  - DO WE UNDERSTAND WHAT WE ARE DOING?
- THE GLUON
  - JETS: FROM THE TEVATRON TO THE LHC
  - PROMPT PHOTONS
  - TOP: NOT YET A STANDARD CANDLE
- FLAVOR SEPARATION: W AND Z PRODUCTION
  - THE IMPACT OF TEVATRON DRELL-YAN DATA
  - INCLUSION OF LHC W ASYMMETRIES
  - THE FORWARD-BACKWARD ASYMMETRY
- A ROAD MAP FOR THE FUTURE?

STANDARD CANDLES

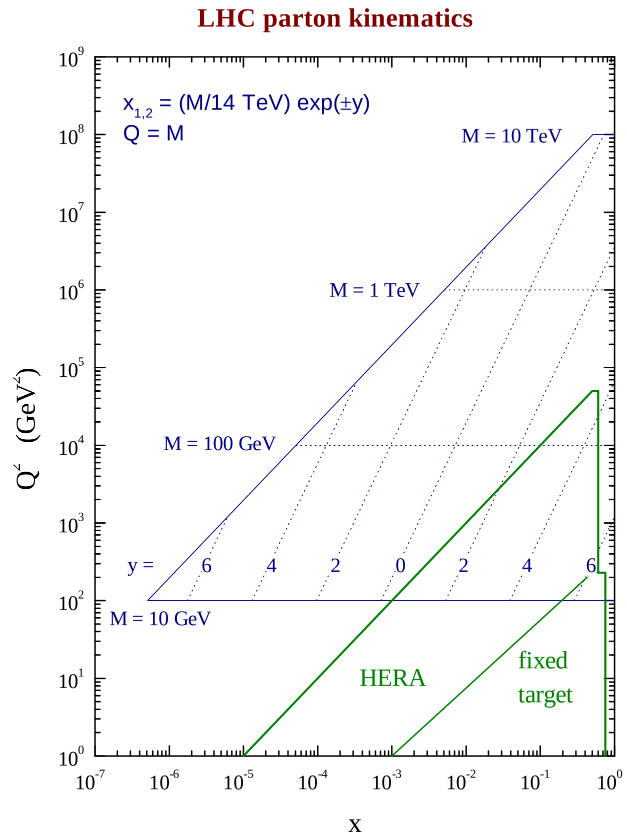
CURRENT PDF SETS

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

data for a global fit (NNPDF2.1)



LHC kinematics



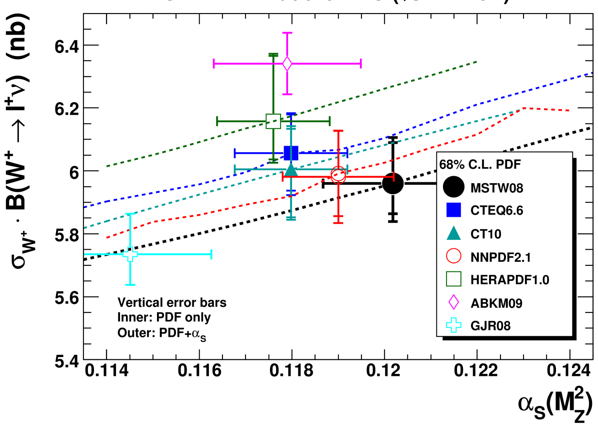
- CT10: GLOBAL, NLO, SEVERAL α_s VALUES, VFN
- MSTW08: GLOBAL, NLO & NNLO, SEVERAL α_s VALUES, VFN
- NNPDF2.1: GLOBAL, NLO, SEVERAL α_s VALUES, VFN
- ALEKHIN ABKM: DIS+SOME DY, NLO & NNLO, SINGLE α_s VALUE, FFN
- HERAPDF1.5: HERA DIS+JETS, NLO & NNLO, SEVERAL α_s VALS, VFN
- SETS BASED ON MODEL ASSUMPTIONS (GRV/GJR, STATISTICAL PDFS,...)

FEATURES, TRADEOFFS AND CHOICES

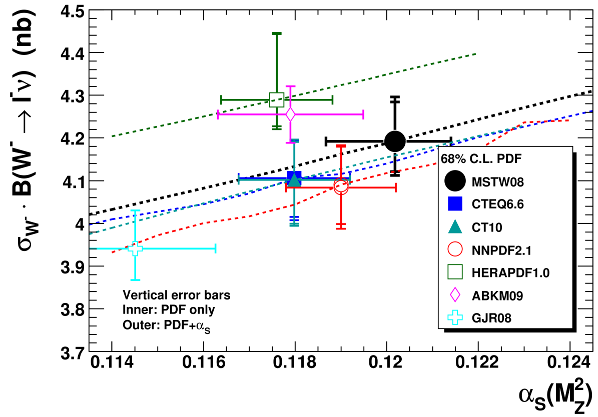
- DATASET: CTEQ, MSTW, NNPDF FIXED TARGET AND COLLIDER, eP AND $\bar{P}P$ DATA; ABKM, GJR GLOBAL DIS+ FIXET-TARGET DY; HERAPDF: HERA ONLY
- STATISTICAL TREATMENT: CTEQ, MSTW HESSIAN WITH DYNAMICAL TOLERANCE; HERAPDF, STANDARD HESSIAN+PARM. ERROR ANALYSIS; GJR, HESSIAN WITH FIXED TOLERANCE; ABKM STANDARD 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;
- 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
- PERTURBATIVE ORDER: CTEQ: NLO, BUT DY LO WITH K-FACTORS; MSTW: NNLO, BUT DY LO WITH K-FACTORS; NNPDF FULL NLO, NNLO DIS; ABKM, GJR: FULL NNLO
- HEAVY QUARKS: CTEQ: GM-VFN (SACOT-χ SCHEME); MSTW: GM-VFN (ACOT+TR' SCHEME); NNPDF: GM-VFN (FONLL SCHEME); ABKM: FFN ($N_f = 3, 4$ MATCHED WITH BMSN SCHEME); GJR: FFN ($N_f = 3$)

WHERE DO WE STAND?
LHC STANDARD CANDLES

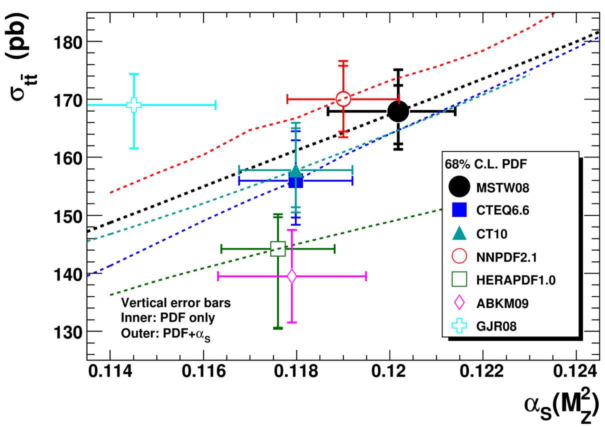
W^+



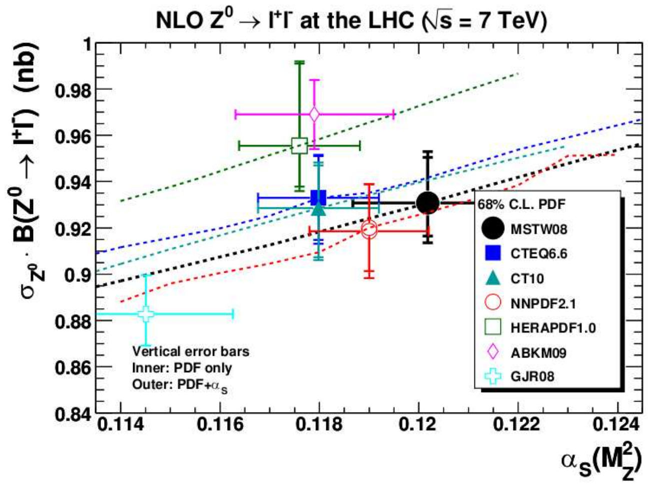
W^-



TOP



Z

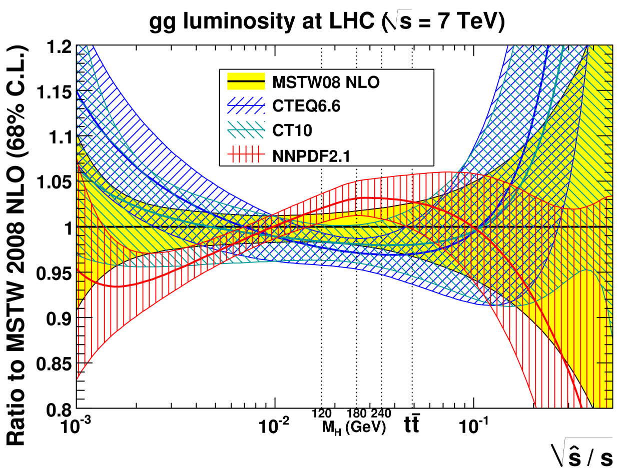


(G. Watt, 2011)

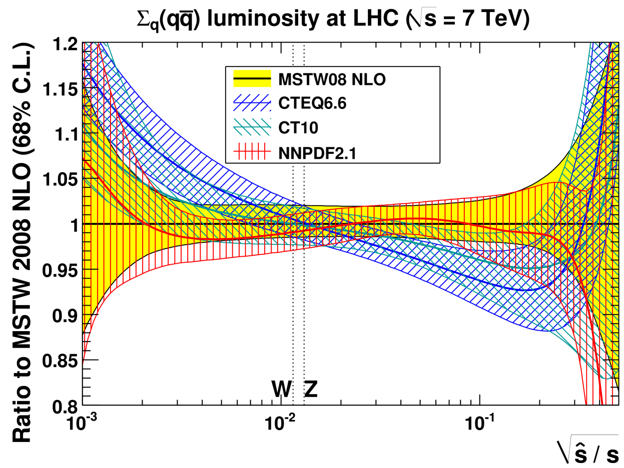
- GLOBAL FITS IN GOOD MUTUAL AGREEMENT
- CHOICE OF α_s VALUE IMPORTANT

WHERE DO WE STAND?
PARTON LUMINOSITIES

GLUON-GLUON

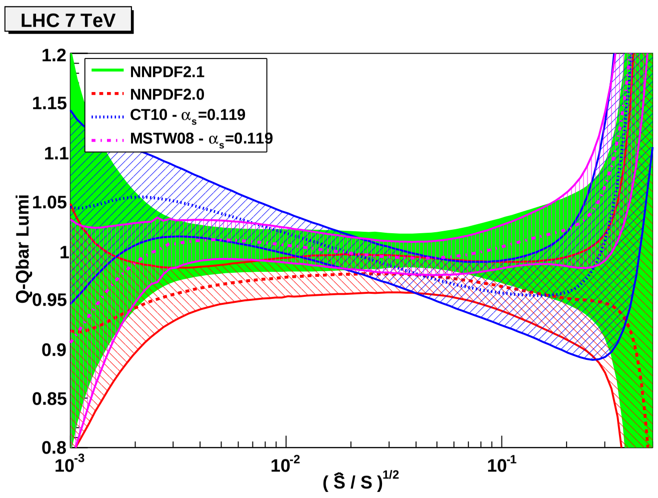


QUARK-QUARK

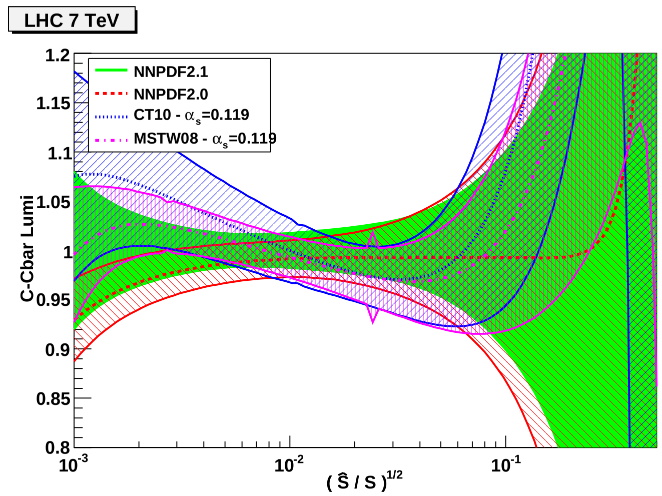


(G. Watt, 2011)

QUARK-GLUON



CHARM-CHARM



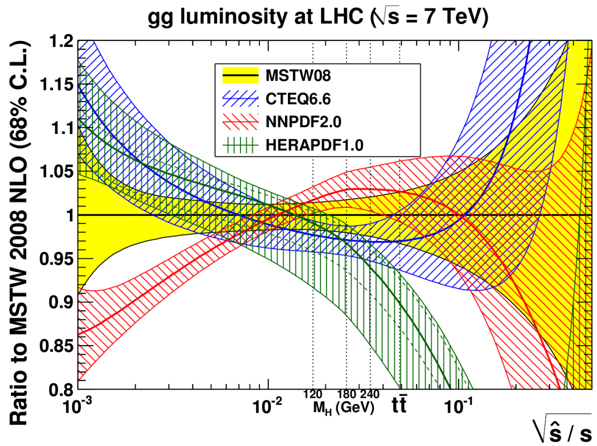
(NNPDF, 2011)

HOW DID WE GET THERE?

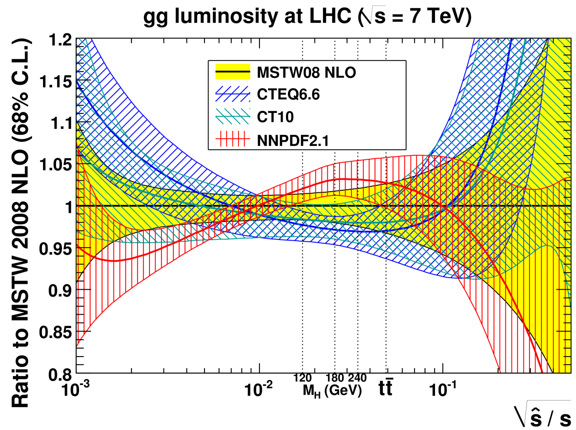
PARTON LUMINOSITIES

GLUON-GLUON

2010

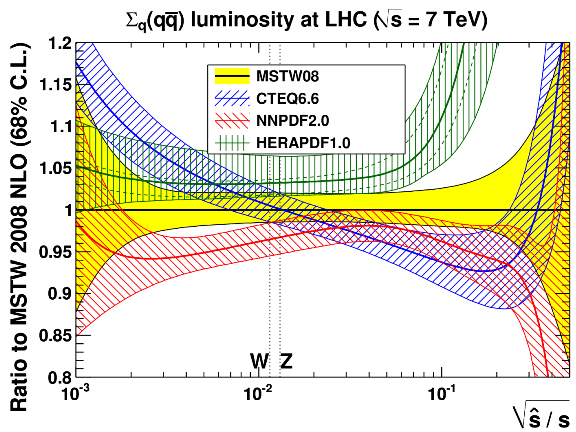


2011

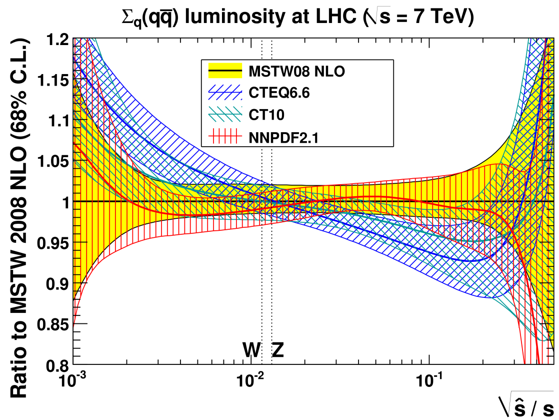


QUARK-QUARK

2010



2011



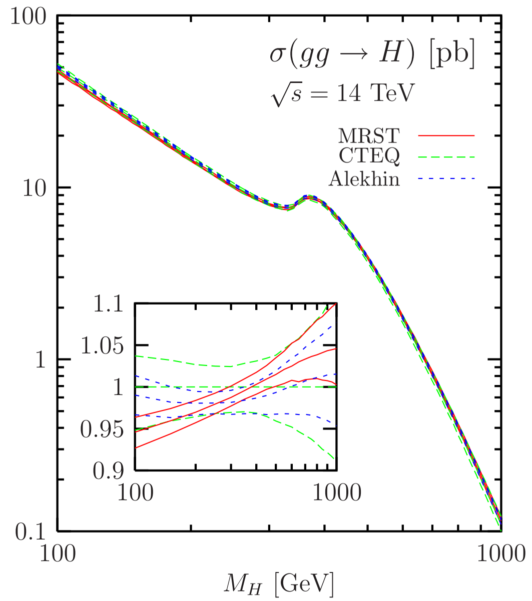
(G. Watt, 2011)

- NNPDF2.0 $\rightarrow$ NNPDF2.1: INCLUSION OF HW MASSES
- CTEQ6.6 $\rightarrow$ CT10: IMPROVED STATISTICS, PARAMETRIZATION, DATA
- CONSIDERABLY IMPROVED AGREEMENT!

THE HIGGS CROSS SECTION

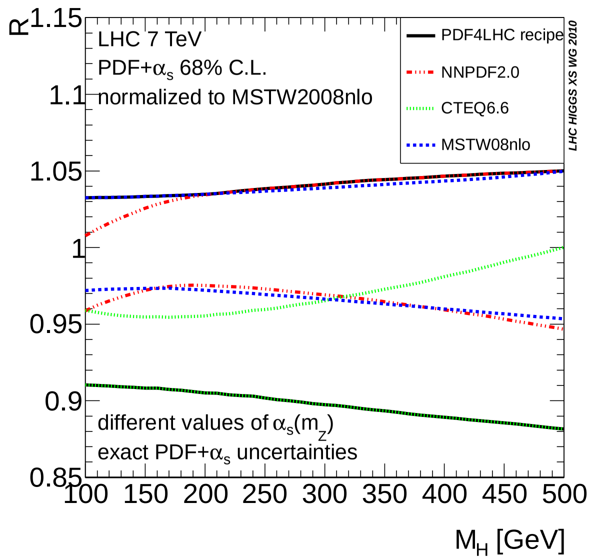
WHERE WE WERE

FIRST PDFs WITH ERRORS



(Djouadi, Ferrag, 2004)

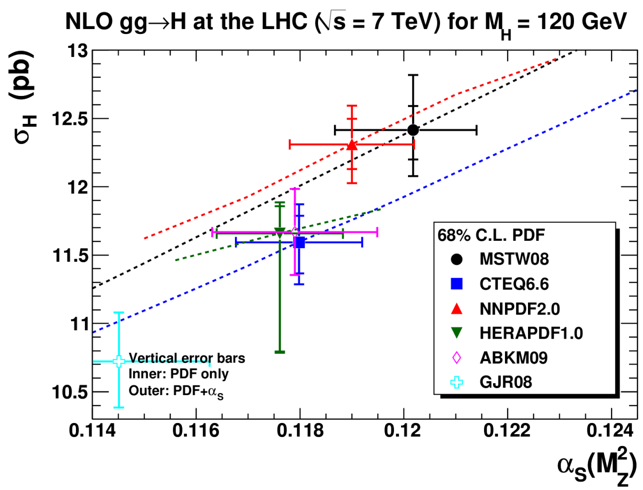
THE PDF4LHC RECIPE



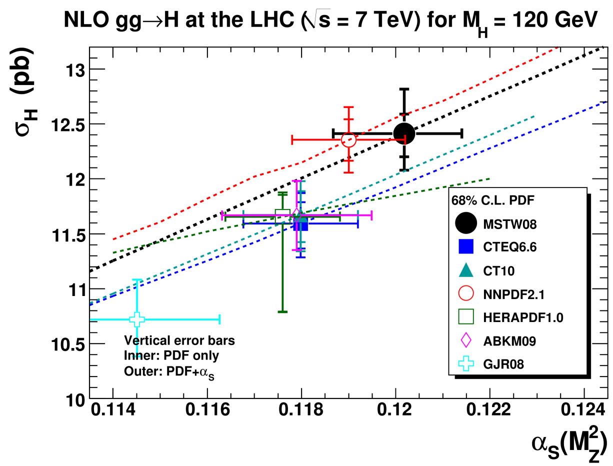
(Higgs WG, 2011)

WHERE WE ARE GOING

$m_H = 120$ GeV: 2010



$m_H = 120$ GeV: 2011



(G. Watt)

DO PDF UNCERTAINTIES HAVE A STATISTICAL MEANING?

SOMETIMES STATED THAT "PDF UNCERTAINTIES ARE THEORETICAL UNCERTAINTIES" (THUS DEVOID OF STATISTICAL MEANING) **IS IT TRUE?**

DO PDF UNCERTAINTIES HAVE A STATISTICAL MEANING?

SOMETIMES STATED THAT "PDF UNCERTAINTIES ARE THEORETICAL UNCERTAINTIES" (THUS DEVOID OF STATISTICAL MEANING) IS IT TRUE?

A WAY OF TESTING:

NEW DATA $\Rightarrow$ BAYES' THEOREM

$$\langle \mathcal{O} \rangle_{\rm new} = \int \mathcal{O}[f]\, \mathcal{P}_{\rm new}(f)\, Df, = \mathcal{N}_\chi \int \mathcal{O}[f]\, \mathcal{P}(\chi^2|f)\mathcal{P}_{\rm old}(f)\, Df,$$

IN A MONTE CARLO APPROACH...

$$\langle \mathcal{O} \rangle_{\rm new} = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \mathcal{N}_\chi \mathcal{P}(\chi^2|f_k)\mathcal{O}[f_k] = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} w_k\, \mathcal{O}[f_k], \quad w_k = \mathcal{N}(\chi^2_k)^{n/2-1} e^{-\frac{1}{2}\chi^2_k}$$

$\Rightarrow$ EFFECT OF NEW DATA IS ACCOUNTED FOR BY REWEIGHTING MONTE CARLO AVERAGES

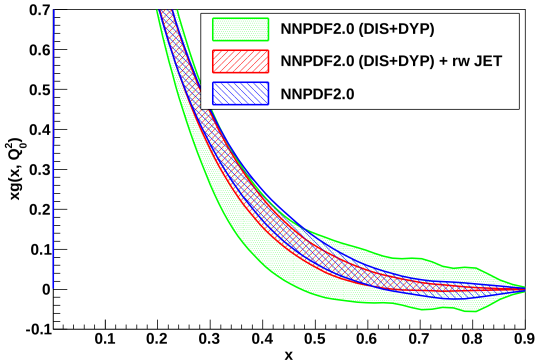
(CAN ALSO BE IMPLEMENTED IN HESSIAN FRAMEWORK (de Lorenzi, McNulty, 2010-2011))

- DETERMINE PDFS INCLUDING SOME DATA BY BAYES' THEOREM (REWEIGHTING)
- DETERMINE PDFS BY ENLARGING THE DATASET TO THE NEW DATA (REFITTING)
- COMPARE RESULTS $\Rightarrow$ IF THEY AGREE, PDFS DO HAVE A STATISTICAL INTERPRETATIONS ALRIGHT!

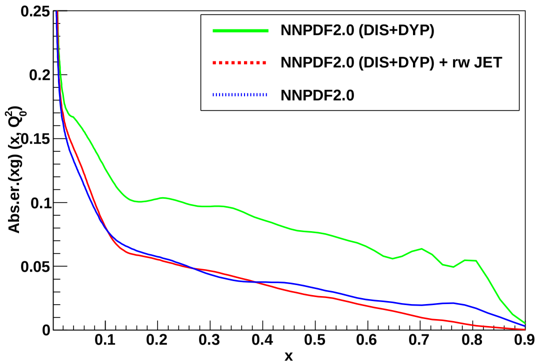
RUNNING THE TEST

INCLUSION OF JET DATA: REWEIGHTING VS. REFITTING
NNPDF2.0DIS+DY VS. NNPDF2.0FULL

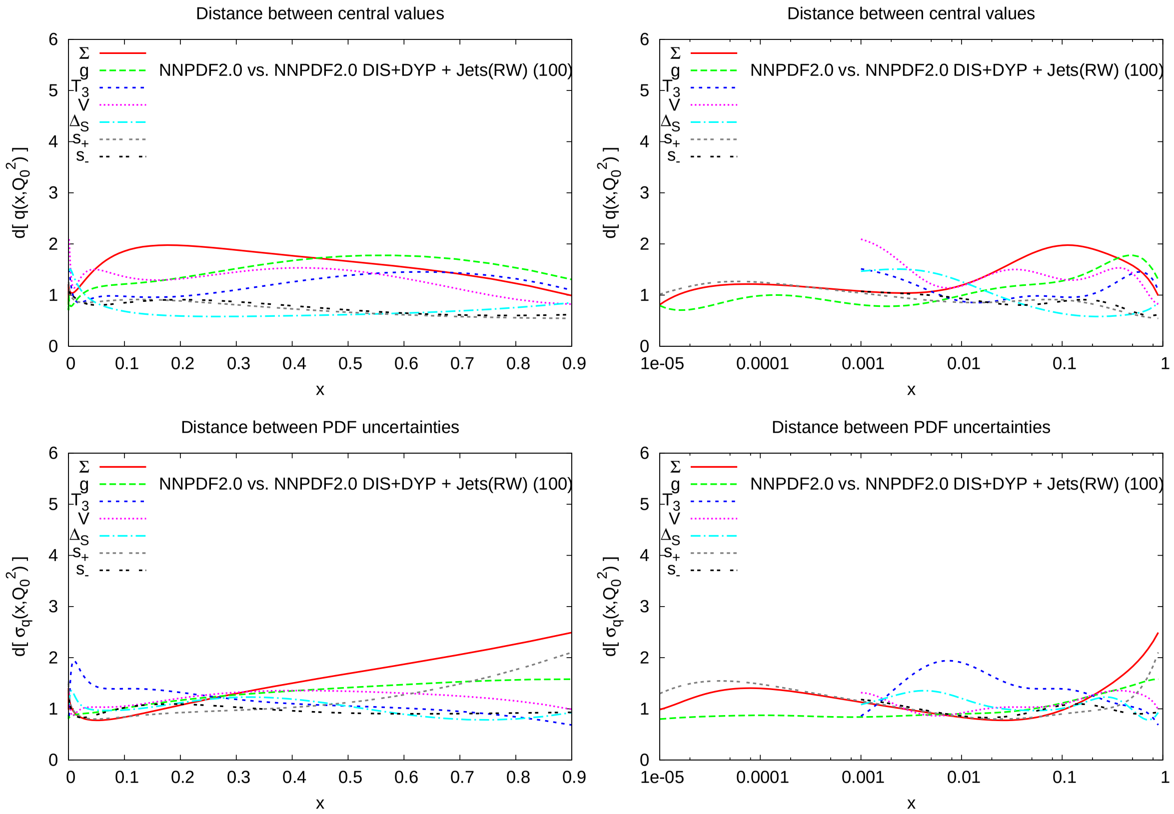
GLUON



GLUON UNCERTAINTY



DISTANCES



$d \sim 1 \Rightarrow$ STATISTICAL EQUIVALENCE

IT LOOKS LIKE WE DO KNOW WHAT WE ARE DOING....

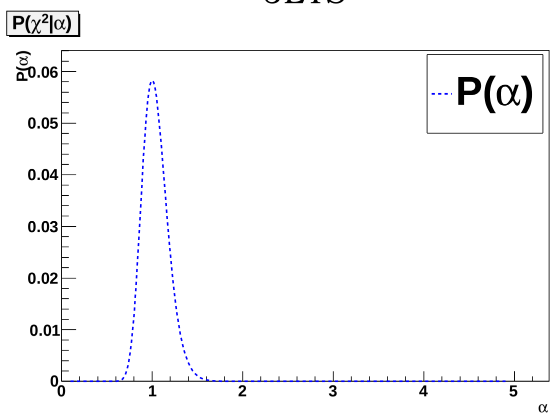
A NEW TOOL: TOLERANCE ESTIMATION AND COMPATIBILITY

- INCONSISTENT DATA $\Leftrightarrow$ UNDERESTIMATED UNCERTAINTIES

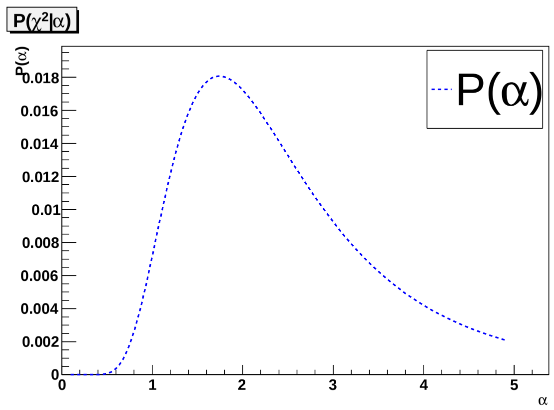
- RESCALE ALL UNCERTAINTIES IN A GIVEN EXPERIMENT BY SOME FACTOR α: $\chi^2_\alpha = \chi^2/\alpha$ (TOLERANCE)

- DETERMINE PROBABILITY DISTRIBUTION OF α VALUES BY BAYES' THEOREM $\Rightarrow$ REWEIGHTING: $\mathcal{P}(\alpha) = \frac{\mathcal{N}}{\alpha} \sum_{k=1}^{N} w_k w_k(\alpha)$.

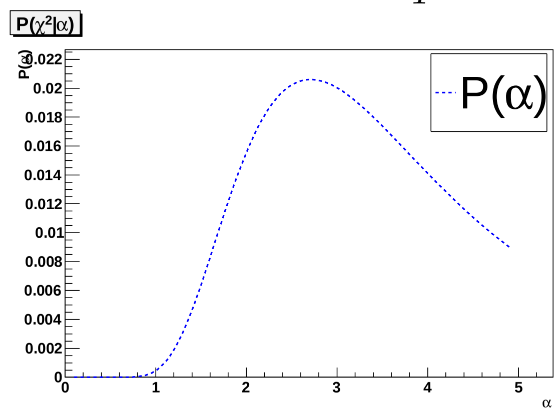
JETS



D0 e INCLUSIVE



D0 e HIGH E_T



- JETS: $\Rightarrow$ CONSISTENT DATA

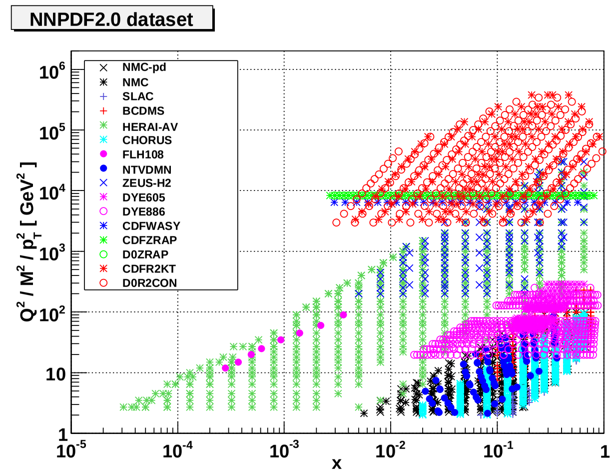
- $W^\pm$ CHARGE ASYMMETRIES, D0 INCLUSIVE e DATA $\Rightarrow$ UNCERTAINTIES UNDERESTIMATED BY $\sim 30\%$ (PROB. PEAKS AT $\alpha \sim 1.7$)

- $W^\pm$ CHARGE ASYMMETRIES, D0 e DATA WITH $E_T > 35$ GEV $\Rightarrow$ INCONSISTENT DATA

THE GLUON

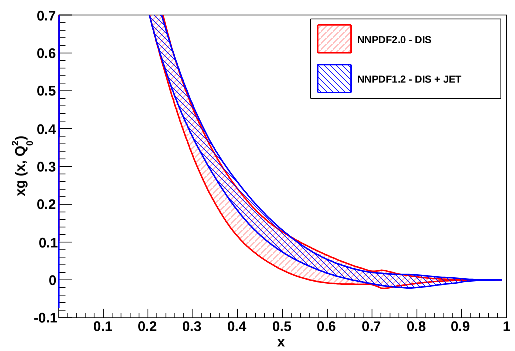
THE IMPACT OF JET DATA (NNPDF2.0)

	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

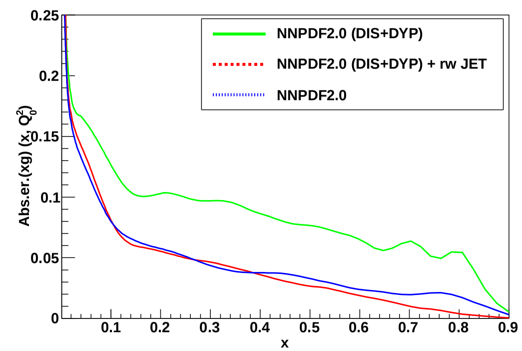


- HIGH E_T JET DATA WELL REPRODUCED EVEN WHEN NOT FITTED $\Rightarrow$ LARGE x GLUON WELL DETERMINED BY SCALING VIOLATIONS!
- SIGNIFICANT IMPROVEMENT GLUON ACCURACY FOR $0.1 \lesssim x \lesssim 0.7$

GLUON



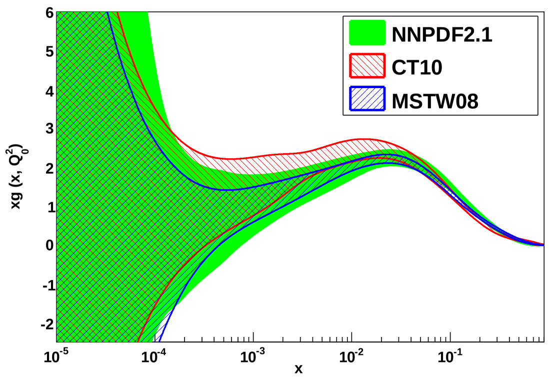
GLUON UNCERTAINTY



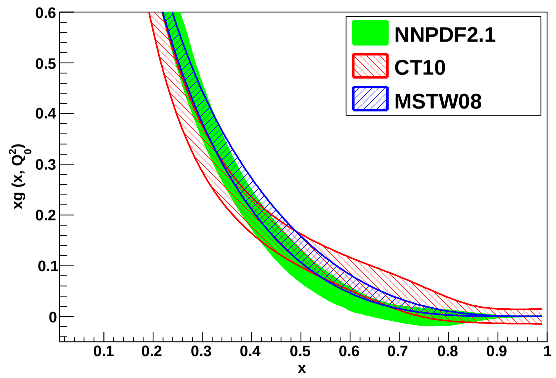
JETS AT THE LHC

- FOR SAME p_T & RAPIDITY, TEVATRON $\to$ LHC, $x \to x/7$
- LARGEST IMPACT EXPECTED IN MEDIUM–x REGION $0.05 \lesssim x \lesssim 0.2$
- CURRENTLY SOME DISAGREEMENT IN THIS REGION

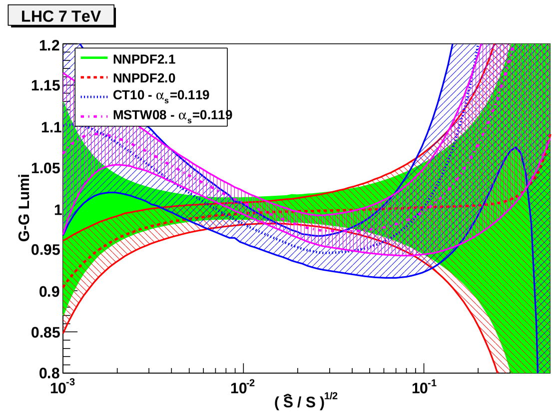
GLUON SMALL x



GLUON LARGE x



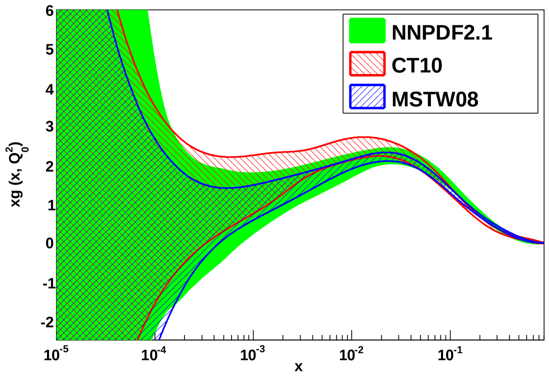
GLUON-GLUON LUMI



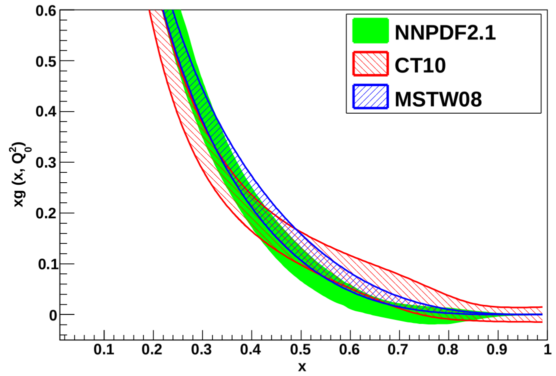
JETS AT THE LHC

- FOR SAME p_T & RAPIDITY, TEVATRON $\rightarrow$ LHC, $x \rightarrow x/7$
- LARGEST IMPACT EXPECTED IN MEDIUM–x REGION $0.05 \lesssim x \lesssim 0.2$
- CURRENTLY SOME DISAGREEMENT IN THIS REGION

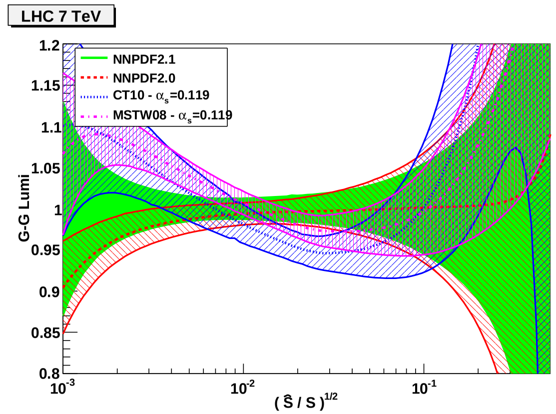
GLUON SMALL x



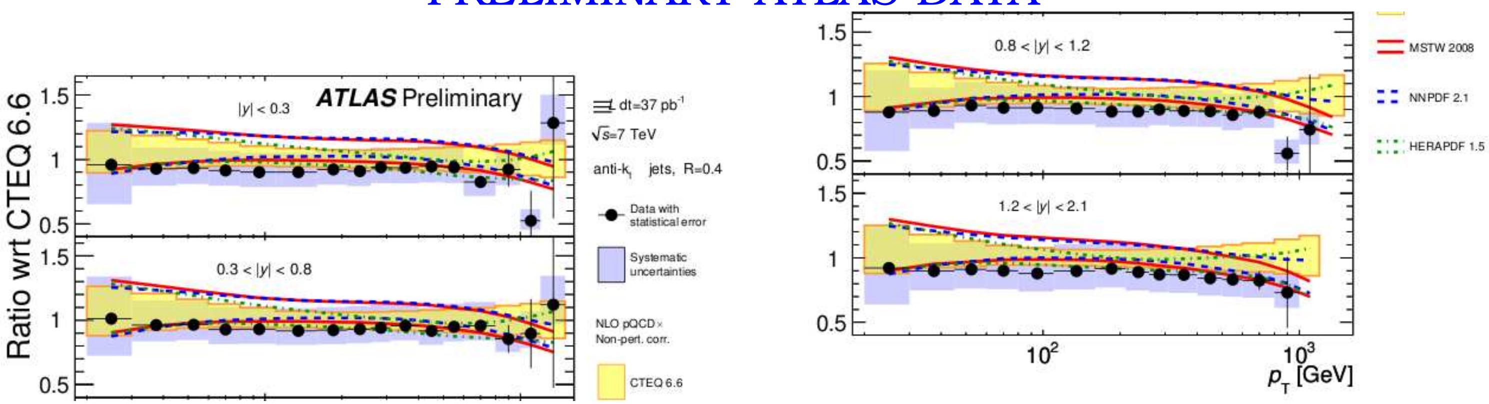
GLUON LARGE x



GLUON-GLUON LUMI

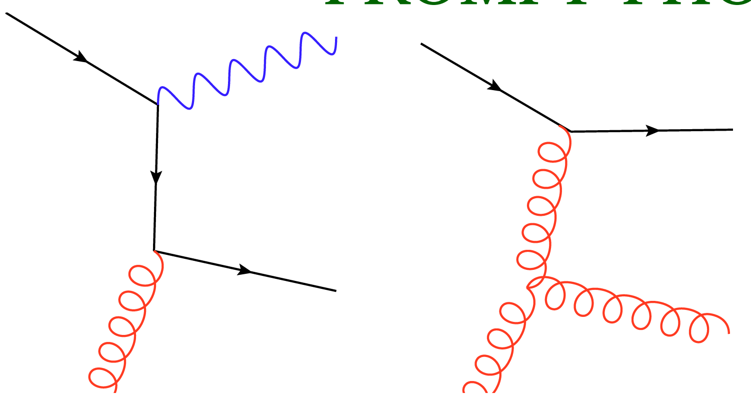


PRELIMINARY ATLAS DATA

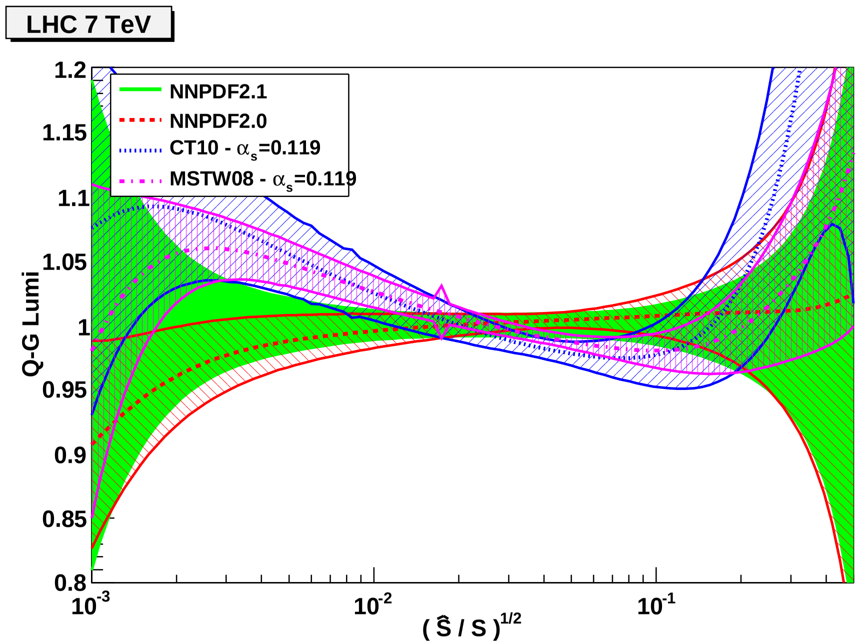


SYSTEMATICS LARGE, BUT GETTING THERE...

PROMPT PHOTON PRODUCTION

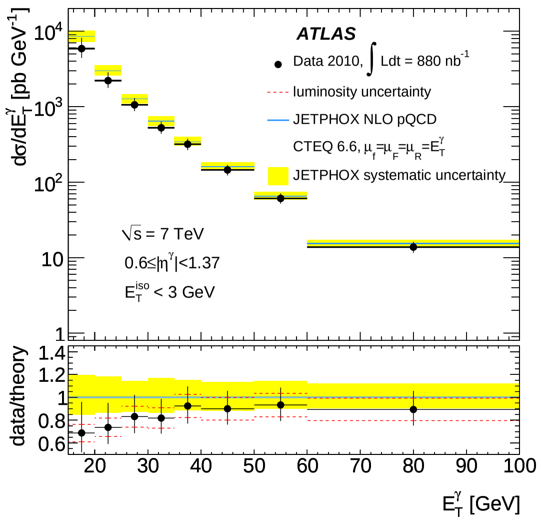


QUARK-GLUON LUMI

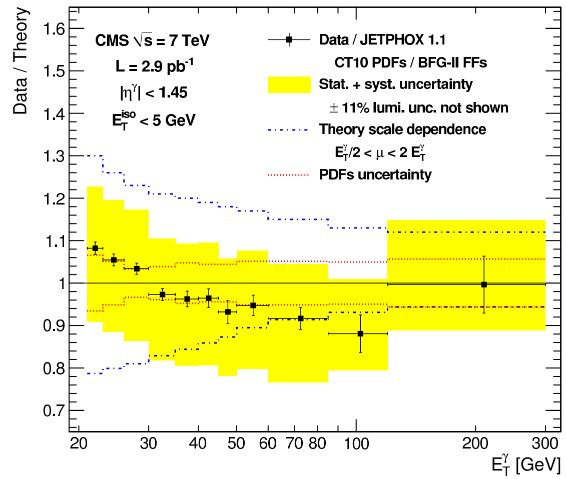


- DIRECT PROBE OF THE qg LUMINOSITY
- $0.01 \lesssim x_1 x_2 = \frac{4p_T}{s} \lesssim 0.2$
- MEASURED BY CMS+ATLAS, UN-CLEAR WHETHER DATA AGREE WITH EACH OTHER & THEORY

ATLAS



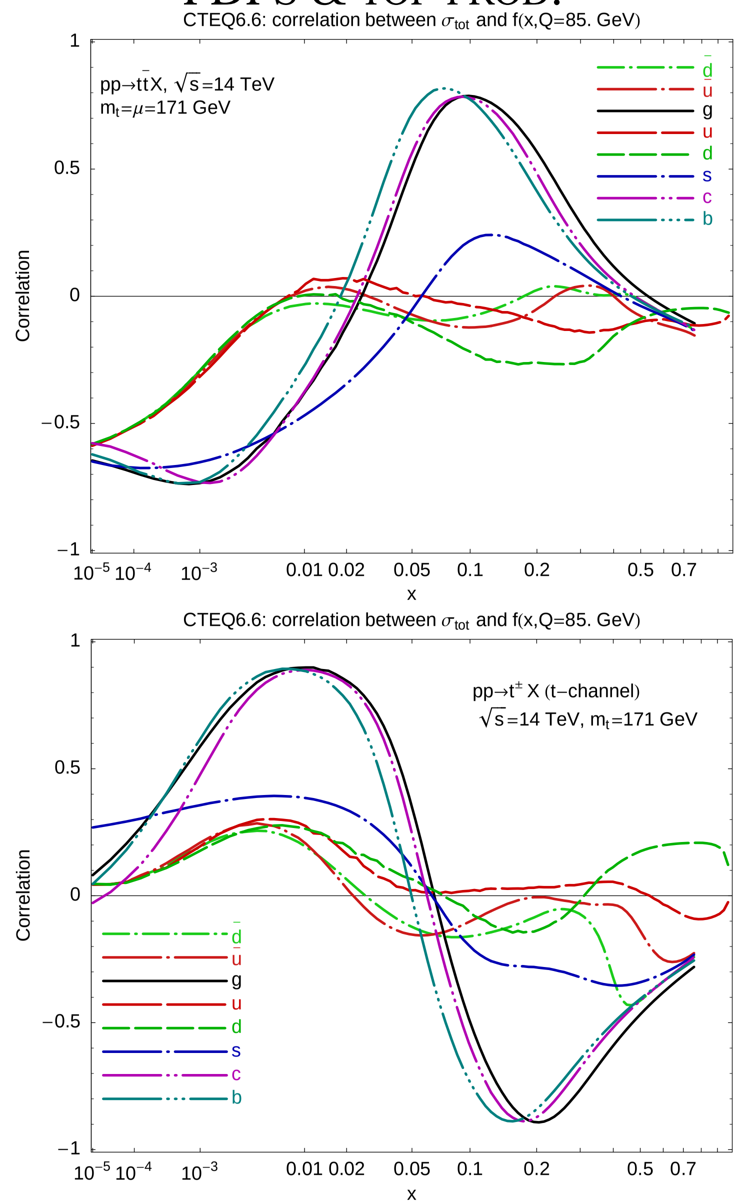
CMS



TOP PRODUCTION

- $t\bar{t}$ STRONGLY CORRELATED TO ALL PDFs
- SINGLE TOP STRONGLY CORRELATED TO GLUON & HEAVY FLAVORS
- SENSITIVE PROBE OF GLUON

CORRELATION BETWEEN PDFs & TOP PROD.

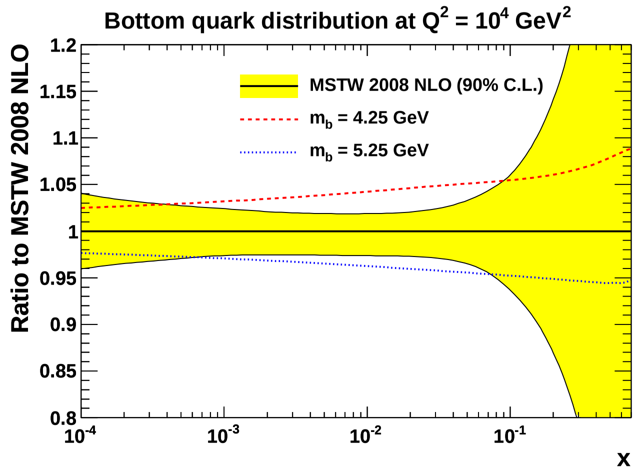


(CT10, 2010)

TOP PRODUCTION

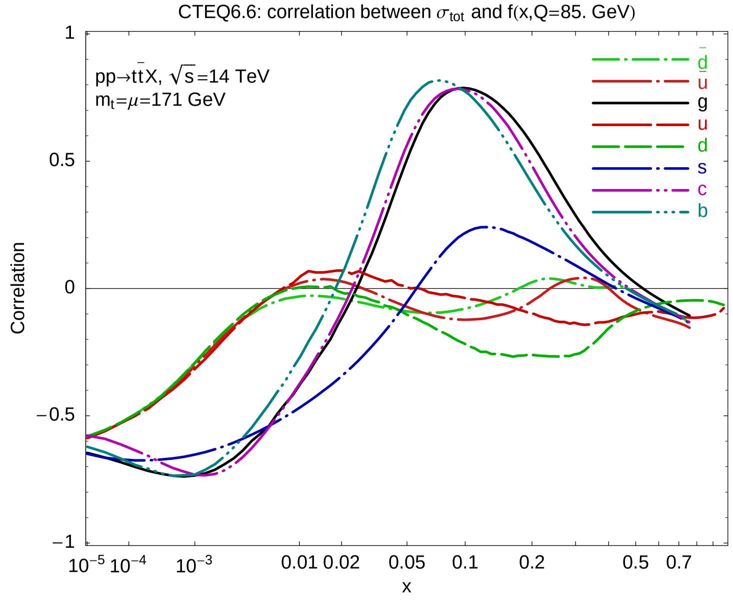
- $t\bar{t}$ STRONGLY CORRELATED TO ALL PDFs
- SINGLE TOP STRONGLY CORRELATED TO GLUON & HEAVY FLAVORS
- SENSITIVE PROBE OF GLUON
- HIGH SENSITIVITY TO b MASS & PDF

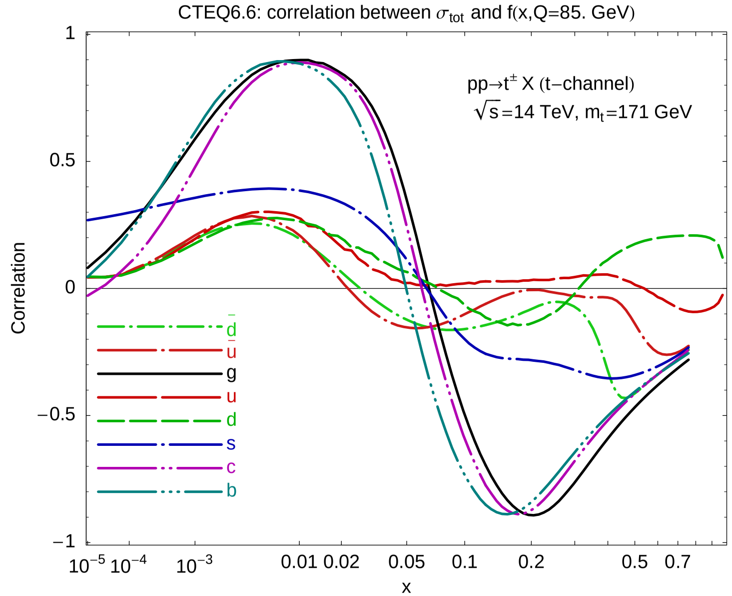
B VS M_b



(MSTW, 2010)

CORRELATION BETWEEN PDFs & TOP PROD.



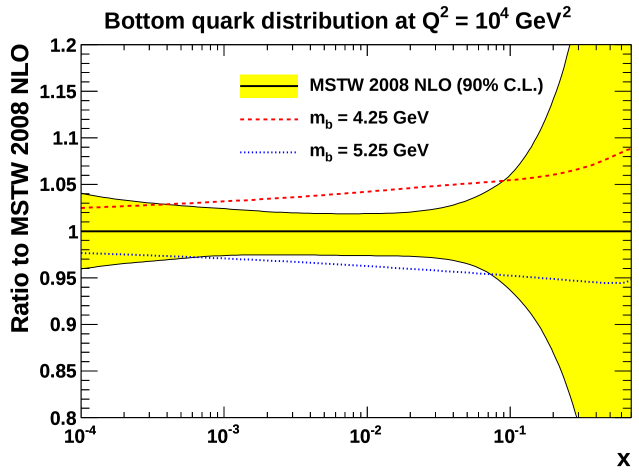


(CT10, 2010)

TOP PRODUCTION

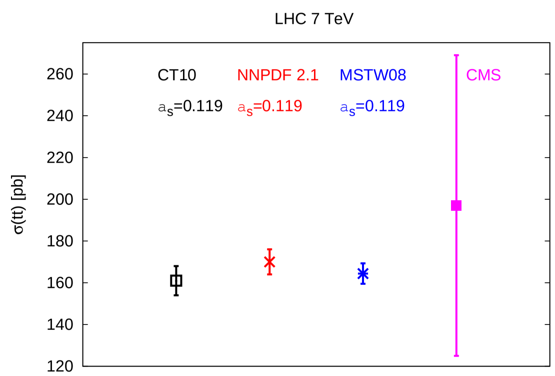
- $t\bar{t}$ STRONGLY CORRELATED TO ALL PDFs
- SINGLE TOP STRONGLY CORRELATED TO GLUON & HEAVY FLAVORS
- SENSITIVE PROBE OF GLUON
- HIGH SENSITIVITY TO b MASS & PDF
- NOT YET COMPETITIVE BUT....

B VS M_b

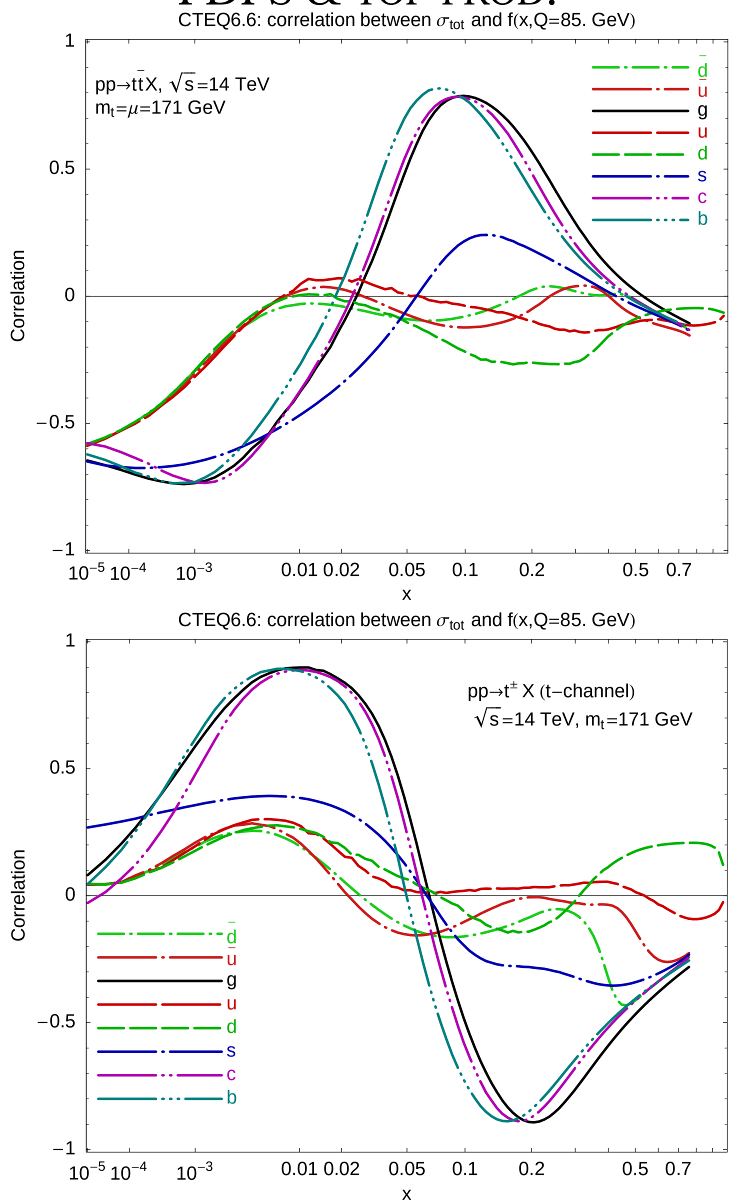


(MSTW, 2010)

TOP XSECT



CORRELATION BETWEEN PDFs & TOP PROD.



(CT10, 2010)

FLAVOR SEPARATION:

DRELL-YAN, W & Z

DIS VS. DRELL-YAN

DIS

NC	${F_1}^{\gamma} = \sum_i e_i^2 \, (q_i + \bar{q}_i)$
NC	${F_1}^{Z,\,\mathrm{int.}} = \sum_i B_i(Q^2/M_Z^2)\,(q_i + \bar{q}_i)$
NC	${F_3}^{Z,\,\mathrm{int.}} = \sum_i D_i(Q^2/M_Z^2)\,(q_i + \bar{q}_i)$
CC	$F_1^{W^+} = \bar{u} + d + s + \bar{c}$
CC	$-F_3^{W^+}/2 = \bar{u} - d - s + \bar{c},$

DRELL-YAN

γ	$\frac{d\sigma}{dM_X^2\,dy}(M_X^2, y) = \frac{4\pi\alpha^2}{9M_X^2 s} \sum_i e_i^2 L^{ii}(x_1^0, x_2^0)$
W	$\frac{d\sigma}{dy} = \frac{\pi G_F M_V^2 \sqrt{2}}{3s} \sum_{i,j} \lvert V_{ij}^{\mathrm{CKM}} \rvert L^{ij}(x_1^0, x_2^0)$
Z	$\frac{d\sigma}{dy} = \frac{\pi G_F M_V^2 \sqrt{2}}{3s} \sum_i \left(V_i^2 + A_i^2\right) L^{ii}(x_1^0, x_2^0)$

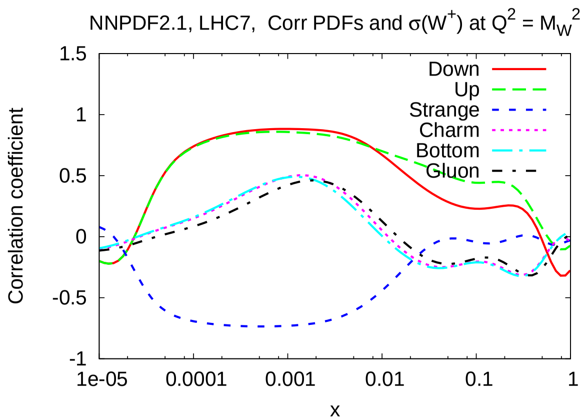
$$L^{ij}\,(x_1, x_2) \equiv q_i(x_1, M_X^2)\bar{q}_j(x_2, M_X^2) + q_i(x_2, M_X^2)\bar{q}_j(x_1, M_X^2)$$

- FIXED TARGET DIS $\Rightarrow$ SINGLET-TRIPLET SEPARATION $F_2^p - F_2^d = \frac{1}{3}[(u^p + \bar{u}^p) - (d^p + \bar{d}^p]$
- HERA THREE INDEPENDENT COMBINATIONS OF PDFS (NC+CC WITH $e^{\pm}$), ONE MORE FROM Q^2 OF Z CONTRIBUTION $\Rightarrow$ u, $\bar{u}$, $d + s$, $\bar{d} + \bar{s}$
  - $F_2^c \Rightarrow$ CHARM
- NEUTRINO DIS $\Rightarrow$ FOUR COMBINATIONS OF PDFS (NC+CC WITH ν, $\bar{\nu}$)
  - "DIMUON" $F_2^{\nu,\,c}$, $F_2^{\bar{\nu},\,c}$ $\Rightarrow$ STRANGE AND ANTISTRANGE
- FIXED TARGET DRELL-YAN (P AND NUCLEAR TARGETS) $\Rightarrow$ ANTIUP/ANTIDOWN SEPARATION
- W AND Z PRODUCTION $\Rightarrow$ UP/DOWN FLAVOR ASYMMETRIES; STRANGENESS

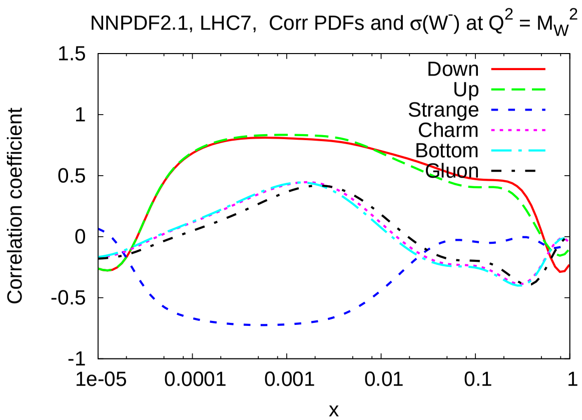
CORRELATIONS: INCLUSIVE DRELL-YAN VS. PDFS

LHC

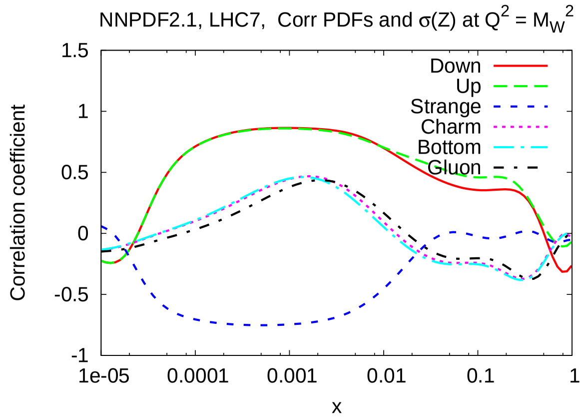
W^+



W^-

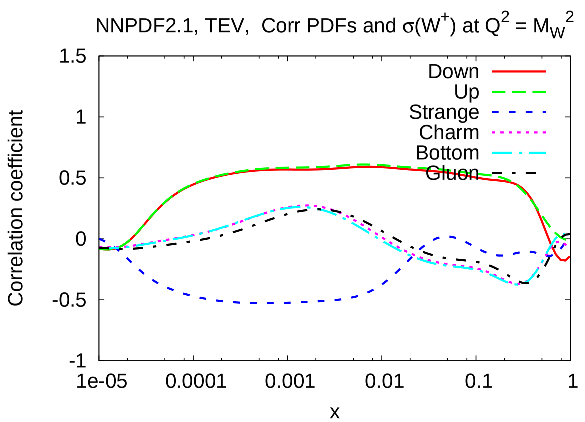


Z

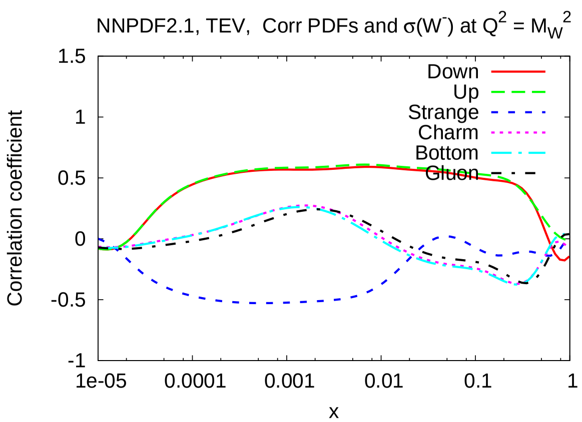


TEVATRON

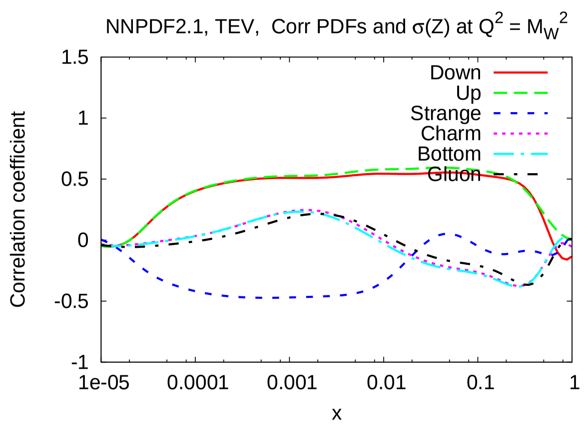
W^+



W^-



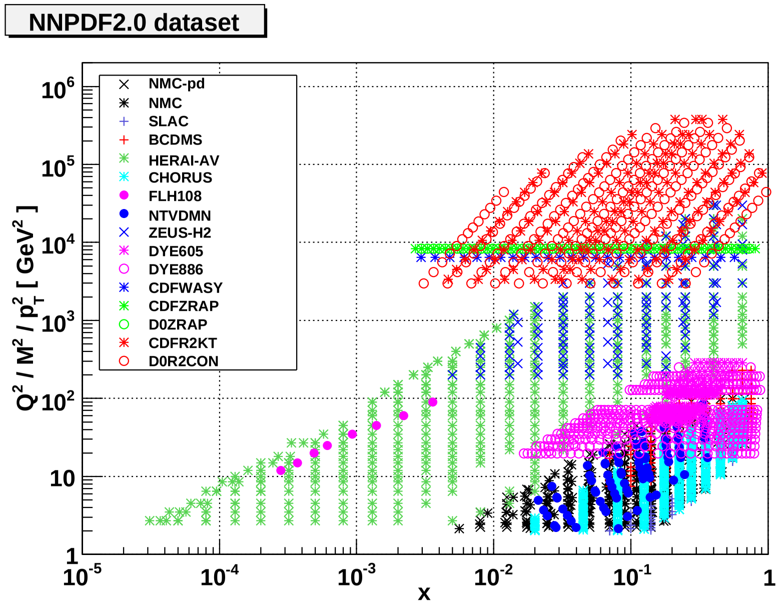
Z



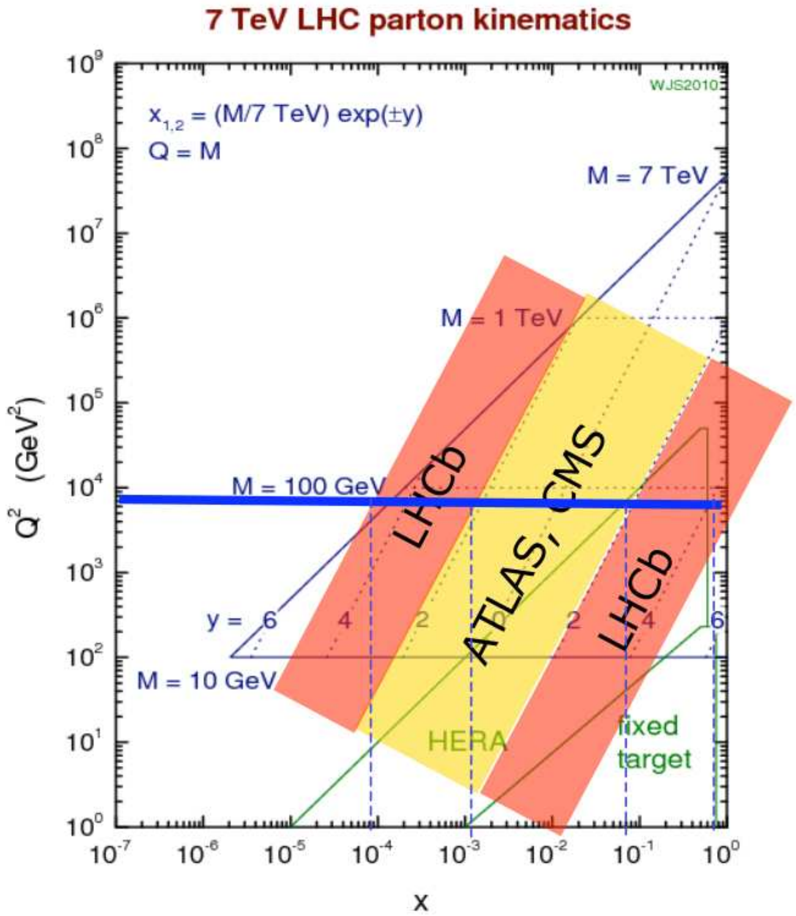
- LHC VS. TEVATRON $\Rightarrow$ CORRELNS. STRONGER FOR $x \lesssim 0.01$, WEAKER @ LARGE x
- LIGHT FLAVORS $\Rightarrow$ LARGE CORRELATION IN WIDE RANGE $3 \cdot 10^{-5} \lesssim x \lesssim 0.3$
- GLUONS & HEAVY FLAVORS $\Rightarrow$ WEAKER CORRELN., MOSTLY AT MEDIUM-SMALL x

RAPIDITY DISTRIBUTIONS: KINEMATIC REGIONS

CURRENT FITS



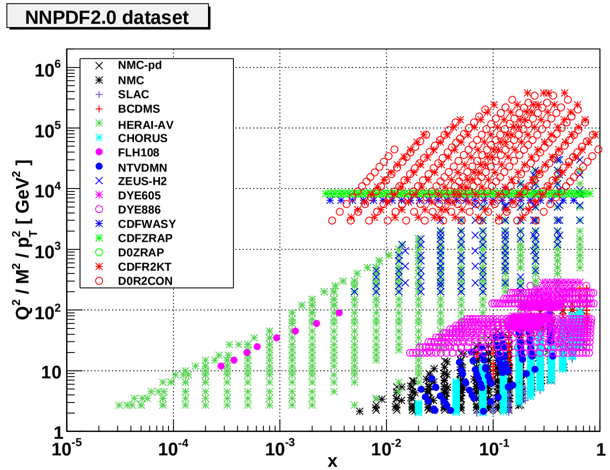
LHC



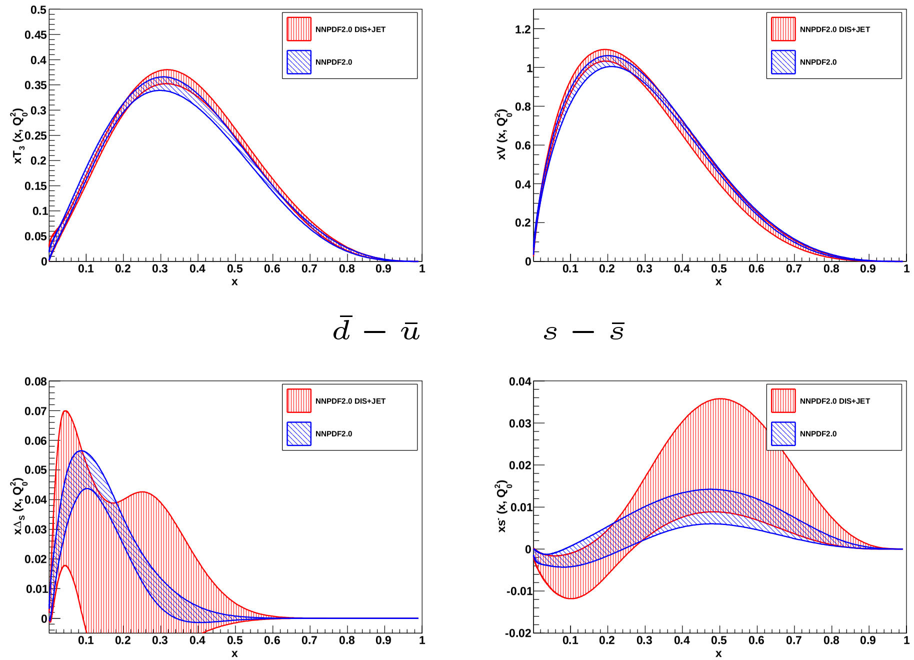
- ATLAS, CMS $\Rightarrow$ BROAD COVERAGE IN CENTRAL RAPIDITY REGION
- LHCb: OUTER RAPIDITY REGIONS $\Rightarrow$ SMALL AND LARGE x PDFs

TEVATRON DRELL-YAN+W-PROD. DATA(NNPDF2.0)

	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.[illegible]**



ISOTRIPLET VALENCE



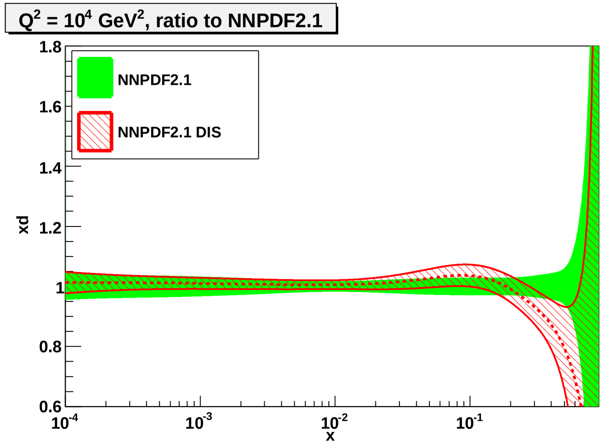
$\bar{d} - \bar{u}$ $s - \bar{s}$

- VERY SUBSTANTIAL IMPROVEMENT IN FIT QUALITY WHEN DATA INCLUDED $\Rightarrow$ 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 $(u + \bar{u} - (d + \bar{d}))$

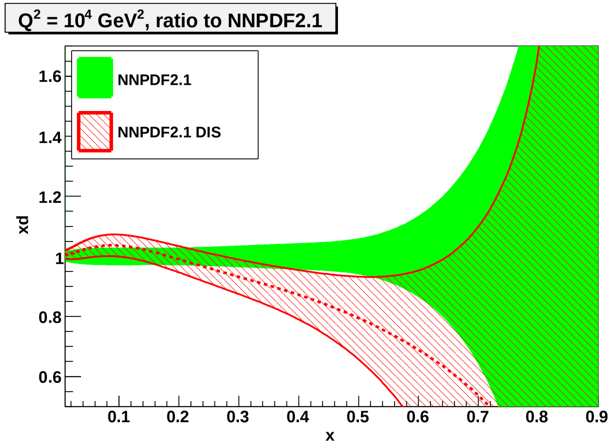
Tevatron Drell-Yan+W-prod. data

quarks: down

SMALL x



LARGE x

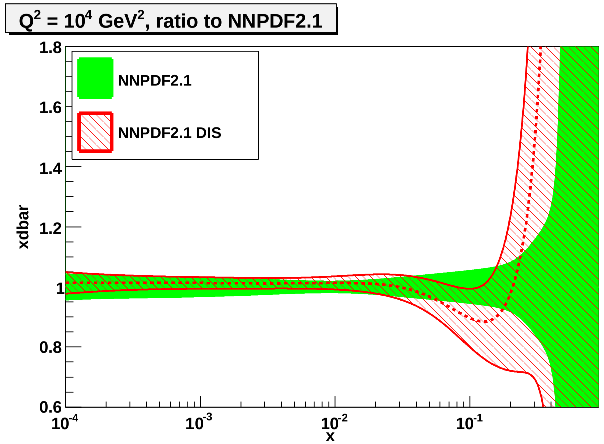


- WELL DETERMINED BY DIS AT SMALL X x
- SOME IMPROVEMENTS IN ACCURACY AT LARGE x
- TENSION BETWEEN DIS AND DY AT LARGE x

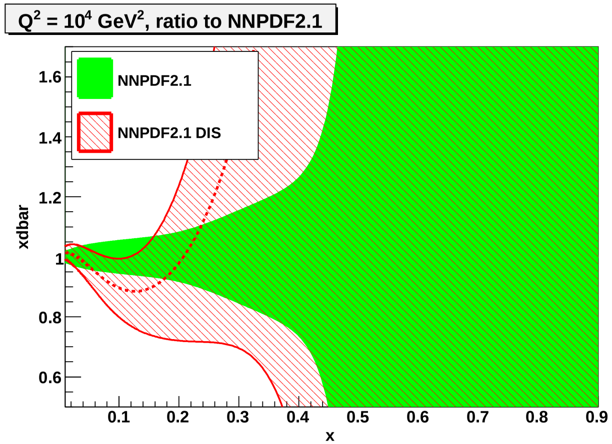
antiquarks: antidown

- AT SMALL x ALL FLAVORS ARE ALIKE $\longrightarrow$ WELL DETERMINED BY DIS DATA
- SOME TENSION BETWEEN DIS AND DY AT MEDIUM-SMALL x
- DRAMATIC IMPROVEMENT IN ACCURACY AT LARGE x

SMALL x



LARGE x



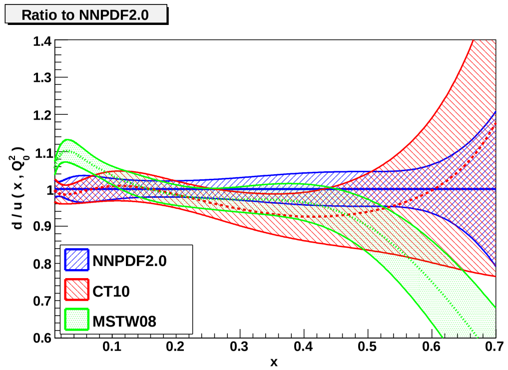
TENSION?

- SHAPES OF THE d QUARK FAVORED BY DIS AND DY DATA DO NOT AGREE
- IS THERE A PROBLEM WITH THE u/d RATIO?

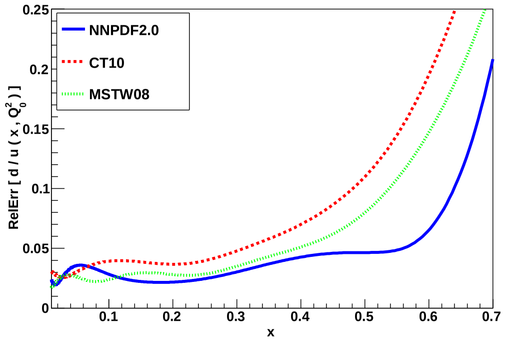
THE UP/DOWN RATIO

COMPARISON BETWEEN PDF SETS

u/d AT $Q^2 = 2$ GEV^2



RELATIVE UNCERTAINTY

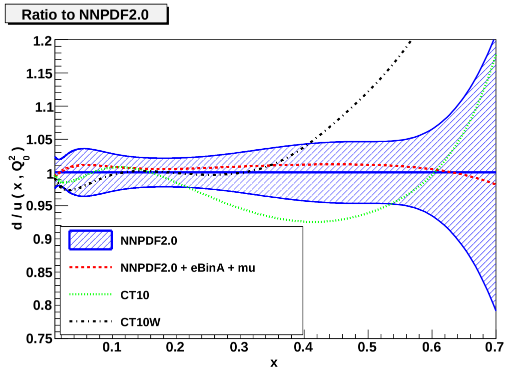


- u/d RATIO IN DIS DOMINATED BY BCDMS AND NMC DATA KNOWN TO HAVE CONSISTENCY PROBLEMS
- PERHAPS NUCLEAR CORRECTIONS & HIGHER TWISTS PLAY A ROLE IN NMC & BCDMS
- DIFFERENT PDF SETS BASED ON THE SAME DATA (BUT WITH DIFFERENT FIT QUALITY) IN PARTIAL DISAGREEMENT

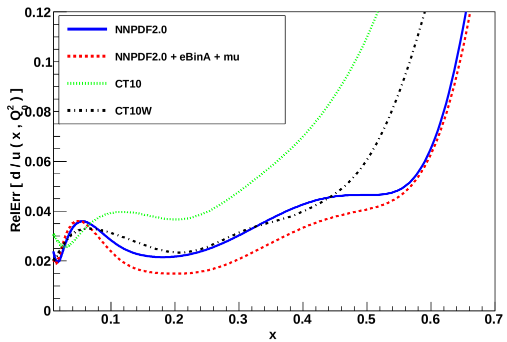
THE RUN II W ASYMMETRY DATA

- PRECISE D0 RUN II LEPTON-LEVEL DATA INCLUDED VIA REWEIGHTING
- NNPDF: SIZABLE IMPROVEMENT IN FIT QUALITY
- CTEQ: NO GLOBALLY CONSISTENT FIT FOUND: CT10W WITH W DATA, CT10 WITHOUT

u/d AT $Q^2 = 2$ GEV^2



RELATIVE UNCERTAINTY



THE W ASYMMETRY AT THE TEVATRON

- PRECISE W ASYMMETRY DATA DETERMINE THE u/d RATIO ACCURATELY
- EVENTUALLY, IT WOULD BE BETTER TO DO WITHOUT NMC/BCDMS DATA (NO LOW Q^2 DATA, NO NUCLEAR TARGETS)

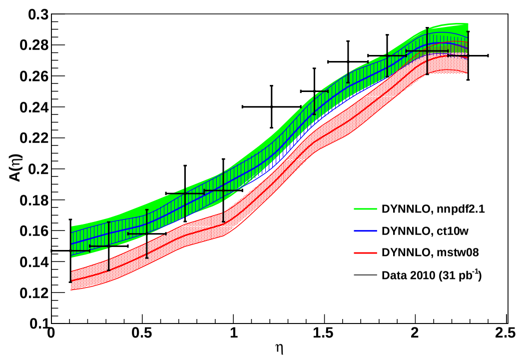
The W lepton asymmetry at the LHC

The $p\bar{p} \to W^{\pm} \to l^{\pm}\nu$ asymmetry: $A_\ell(y_\ell) = \frac{d\sigma(W^+)/dy_\ell - d\sigma(W^-)/dy_\ell}{d\sigma(W^+)/dy_\ell + d\sigma(W^-)/dy_\ell}$

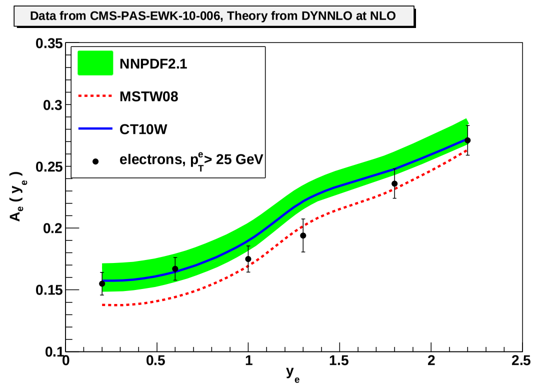
- Accurate ATLAS (arXiv:1103.2929) and CMS (arXiv:1103.3470) data published earlier this month
- LHCb preliminary data (uncorrected for EW, not discussed here)

$\chi^2/N_{\rm dat}$ (el)	ATLAS	CMS $p_T^l \geq 25$ GeV	CMS $p_T^l \geq 30$ GeV
NNPDF2.1	0.68	1.8	1.6
MSTW08	3.24	1.8	2.3
CT10W	0.81	1.2	1.3

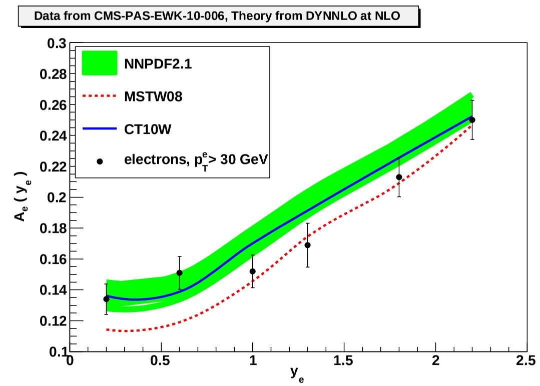
ATLAS



CMS $p_t > 25$ GeV



CMS $p_t > 30$ GeV

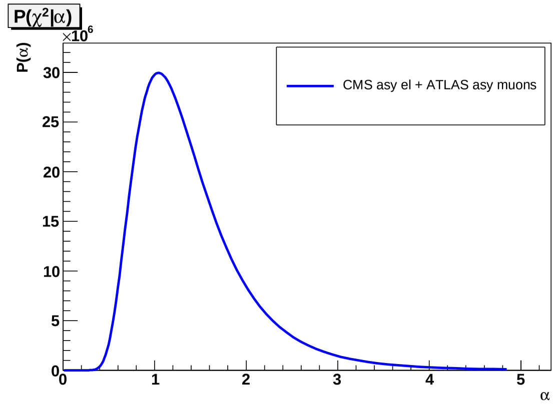


including the LHC W asymmetry in PDF fits

- LHC data can be included in the PDF determination using the reweighting method

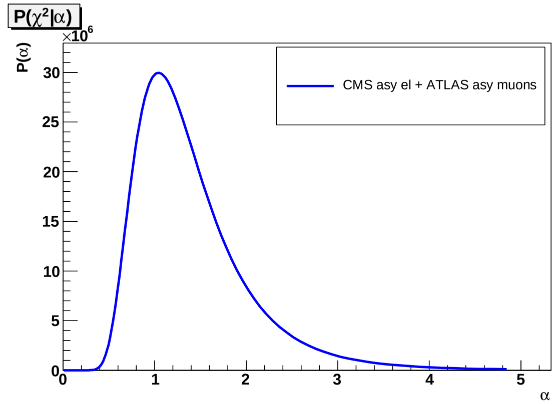
INCLUDING THE LHC W ASYMMETRY IN PDF FITS

- LHC DATA CAN BE INCLUDED IN THE PDF DETERMINATION USING THE REWEIGHTING METHOD
- CAN CHECK FOR CONSISTENCY

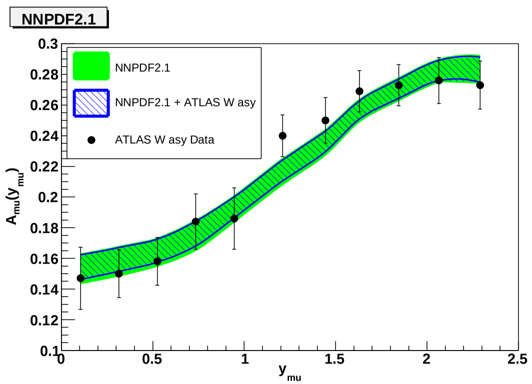


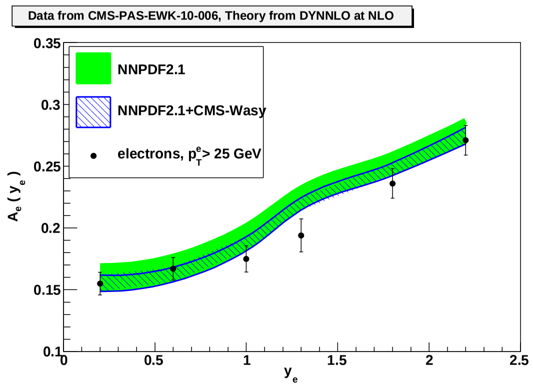
INCLUDING THE LHC W ASYMMETRY IN PDF FITS

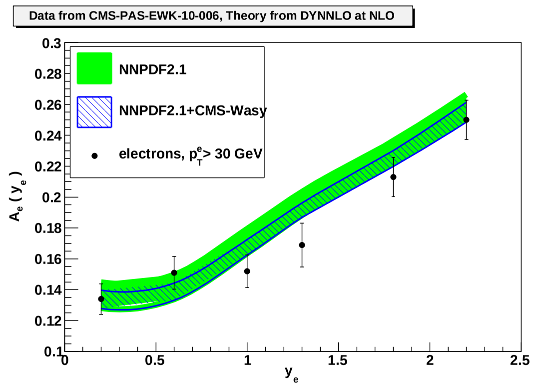
- LHC DATA CAN BE INCLUDED IN THE PDF DETERMINATION USING THE REWEIGHTING METHOD
- CAN CHECK FOR CONSISTENCY
- RECOMPUTE OBSERVABLES AFTER INCLUSION & VERIFIY FIT QUALITY



$\chi^2/N_{\rm dat}$ (EL)	ATLAS	CMS $p_T^l \geq 25$ GEV	CMS $p_T^l \geq 30$ GEV	ALL
NNPDF2.1	0.68	1.8	1.6	0.95
NNPDF2.1+EACH DATASET	0.63	1.03	1.16	-
NNPDF2.1+ALL DATA	-	-	-	0.92





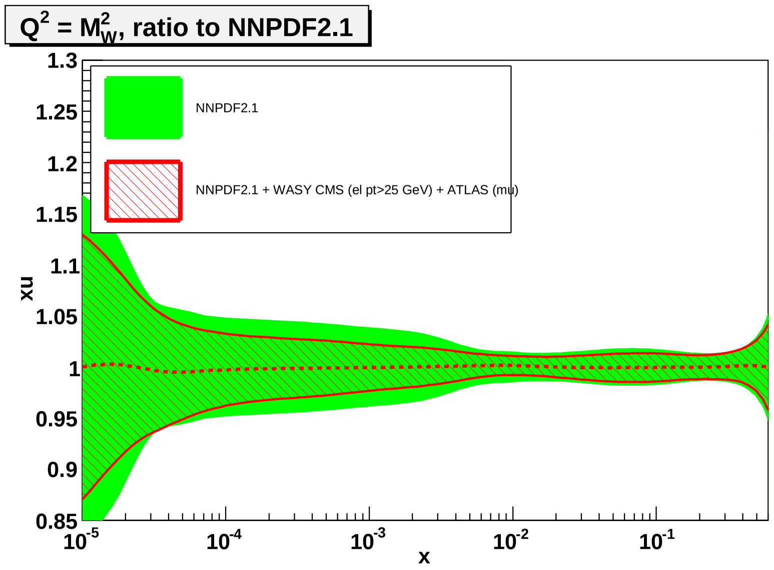


LHC W ASYMMETRY $\Rightarrow$ PDFS
REWEIGHTED PDFS

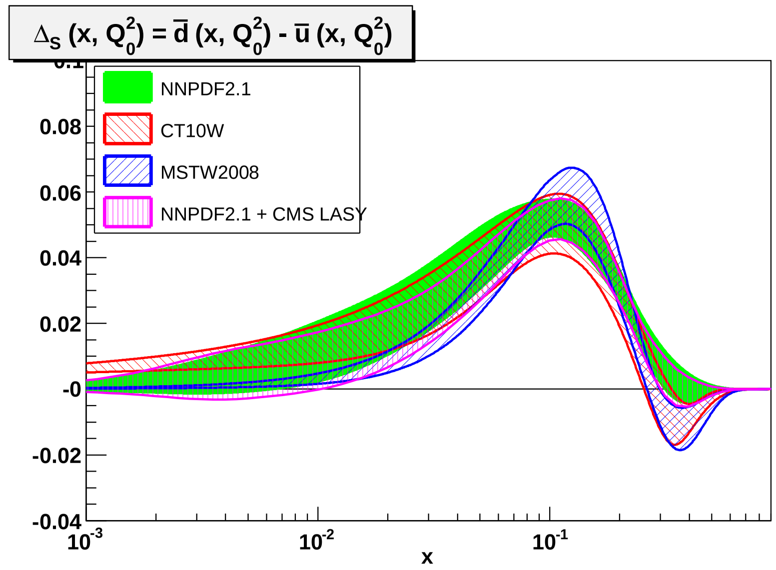
WHAT IS THE EFFECT OF THE LHC DATA ON PDFS?

- ALL LIGHT QUARK UNCERTAINTIES REDUCED
- GOOD CONSISTENCY $\Rightarrow$
  - CENTRAL VALUES DON'T CHANGE MUCH
  - UNCERTAINTIES SHRINK
- DIFFERENCE W.T. TO OTHER PDFS: MOSTLY $\bar{u} - \bar{d}$
  - NNPDF CENTRAL VALUE UNCHANGED WITH SMALLER UNCERTAINTY
  - MSTW (CTEQ) SHAPE SOMEWHAT DIFFERENT $\Rightarrow$ DISFAVORED?

THE UP DISTRIBUTION



$\bar{u} - \bar{d}$

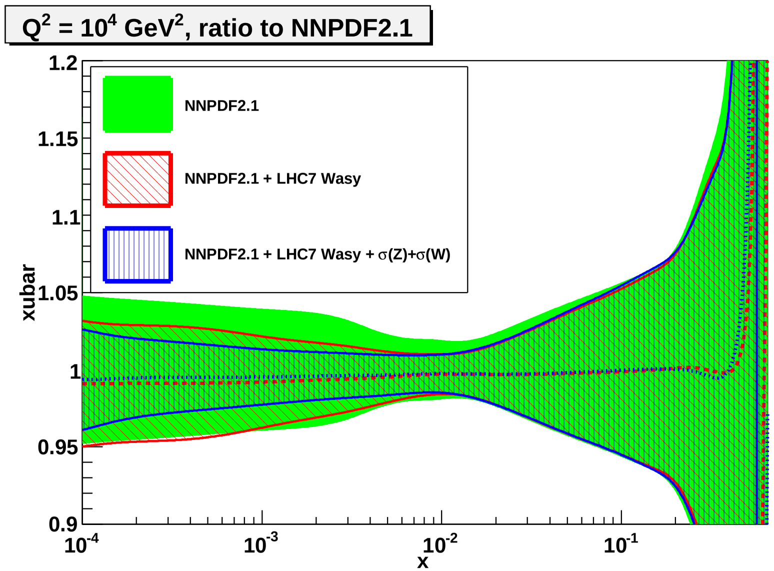


FUTURE LHC DATA

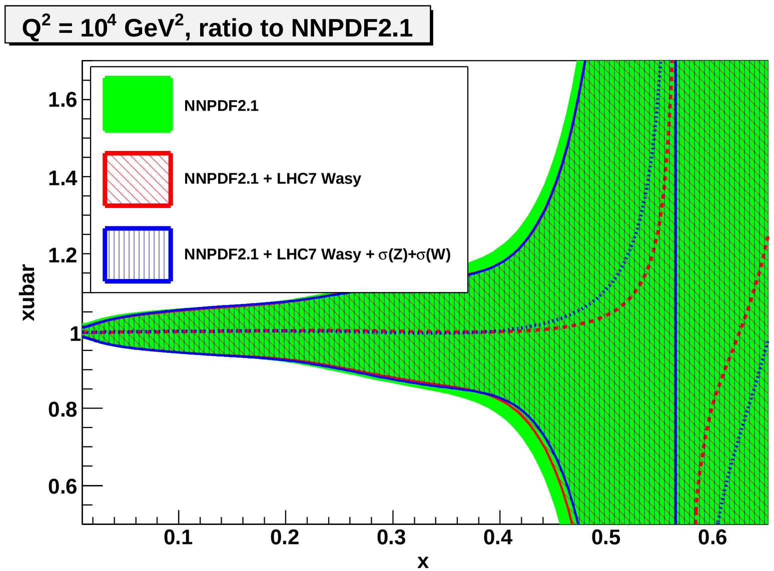
TEST THE IMPACT OF LHC 7TEV W AND Z DATA USING REWEIGHTING IN THE FOLLOWING “TOY” SCENARIOS:

- W ASYMMETRY MEASURED WITH $\sim 5\%$ ACCURACY AT ATLAS (KINEMATICS COURTESY OF A. GLAZOV) $\Rightarrow$ SIGNIFICANT IMPROVEMENT FOR ALL FLAVORS AT SMALL x
- W ASYMMETRY + TOTAL W AND Z CROSS-SECTIONS MEASURED TO 2% ACCURACY $\Rightarrow$
  - FURTHER IMPROVEMENT AT SMALL x
  - NOTE EFFECT SEEN ON ALL FLAVORS AND ANTIFLAVORS

ANTIUP SMALL x



ANTIUP LARGE x



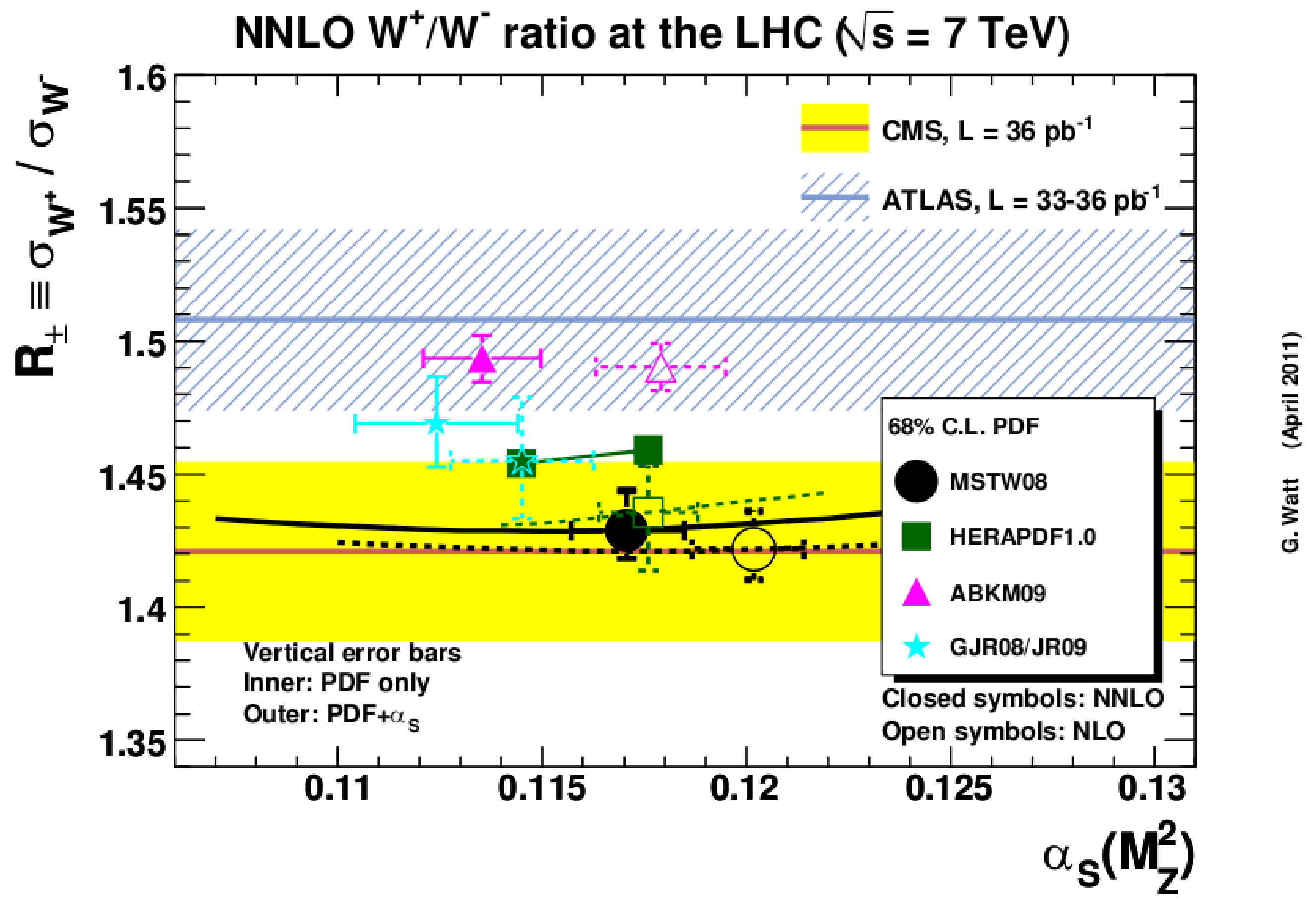
PERCENTAGE ACCURACY ON W, Z CROSS SECTIONS?

- NNLO PROBABLY NECESSARY
- REQUIRES INDEPENDENT ACCURATE MEASUREMENT OF LUMINOSITY
- ACCURATE ACCEPTANCE CALCULATIONS
- EVEN FOR W^+/W^- RATIO NOT EASY

  $\Rightarrow$

PERCENTAGE ACCURACY ON W, Z CROSS SECTIONS?

- NNLO PROBABLY NECESSARY
- REQUIRES INDEPENDENT ACCURATE MEASUREMENT OF LUMINOSITY
- ACCURATE ACCEPTANCE CALCULATIONS
- EVEN FOR W^+/W^- RATIO NOT EASY
  $\Rightarrow$CURRENT CMS-ATLAS DISAGREEMENT (?)



ACCEPTANCE: PROBLEMS?

- CMS [arXiv:1012.2466, CMS PAS EWK-10-005] use POWHEG (with CTEQ6.6/CT10, MSTW08NLO, NNPDF2.0), check with FEWZ.

$$\Rightarrow \frac{\sigma^{\rm tot}_{W^+}}{\sigma^{\rm tot}_{W^-}} = \begin{cases} 1.418 \pm 0.008(\text{stat.}) \pm 0.022(\text{syst.}) \pm 0.029(\text{th.}) & (e) \\ 1.423 \pm 0.008(\text{stat.}) \pm 0.019(\text{syst.}) \pm 0.031(\text{th.}) & (\mu) \\ 1.421 \pm 0.006(\text{stat.}) \pm 0.014(\text{syst.}) \pm 0.030(\text{th.}) & (e+\mu) \end{cases}$$

- ATLAS [arXiv:1010.2130, ATLAS-CONF-2011-041] use PYTHIA (with MRSTLO*) $\Rightarrow \sigma^{\rm tot}_{W^+}/\sigma^{\rm tot}_{W^-} = 1.508 \pm 0.034\ (e+\mu)$.

MC+PDF	A_{W^+}/A_{W^-}		$\sigma^{\rm tot}_{W^+}/\sigma^{\rm tot}_{W^-}$	
	e	μ	e	μ
PYTHIA+MRSTLO*	1.02	1.02	1.50	1.51
PYTHIA+CTEQ6.6	1.05	1.05	1.46	1.47
PYTHIA+HERAPDF1.0	1.03	1.04	1.48	1.49
MC@NLO+HERAPDF1.0	1.05	1.05	1.46	1.47
MC@NLO+CTEQ6.6	1.06	1.06	1.45	1.46

$$\frac{\sigma^{\rm tot}_{W^+}}{\sigma^{\rm tot}_{W^-}} = \frac{\sigma^{\rm fid}_{W^+}}{\sigma^{\rm fid}_{W^-}} \cdot \frac{A_{W^-}}{A_{W^+}}$$

Alternative acceptance calculations move ATLAS $\sigma^{\rm tot}_{W^+}/\sigma^{\rm tot}_{W^-}$ much closer to CMS.

- Precision requires better acceptance calculations and/or data to theory comparisons at level of fiducial cross section.

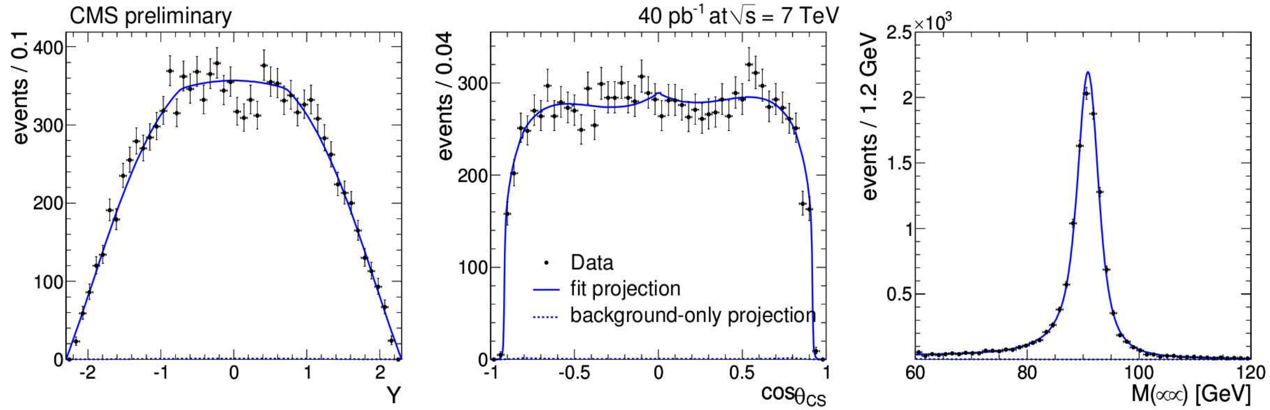
(G.Watt, 2011)

$\sin^2\theta_W$ FROM A_{FB} AT THE LHC?

- FORWARD/BACKWARD ASYMMETRY OF LEPTONS PRODUCED BY Z DECAY $\Rightarrow$ ELECTROWEAK MIXING ANGLE
- CURRENTLY $\sin^2\theta_W = 0.23153 \pm 0.00016$ FROM GLOBAL FIT BUT TWO MOST PRECISE MEASUREMENTS $\sin^2\theta_W = 0.23098 \pm 0.00026$ (SLD) & $\sin^2\theta_W = 0.23221 \pm 0.00029$ (B) DISAGREE BY THREE SIGMA
- $\rightarrow$ AIM AT 0.0001 LEVEL ACCURACY!
- IS IT POSSIBLE AT HADRON COLLIDER WHERE PDF UNCERTAINTIES PLAY A ROLE?

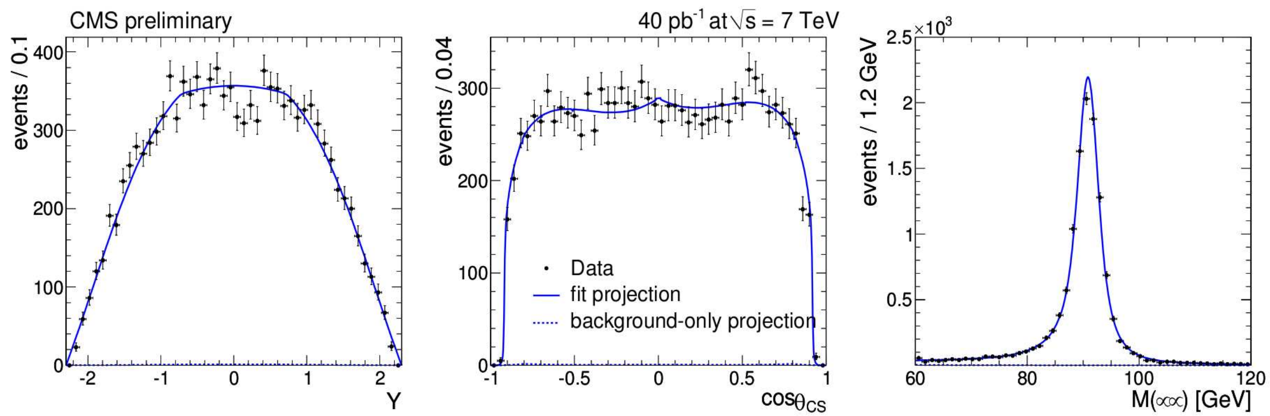
$\sin^2\theta_W$ FROM A_{FB} AT THE LHC?

- FORWARD/BACKWARD ASYMMETRY OF LEPTONS PRODUCED BY Z DECAY $\Rightarrow$ ELECTROWEAK MIXING ANGLE
- CURRENTLY $\sin^2\theta_W = 0.23153 \pm 0.00016$ FROM GLOBAL FIT BUT TWO MOST PRECISE MEASUREMENTS $\sin^2\theta_W = 0.23098 \pm 0.00026$ (SLD) & $\sin^2\theta_W = 0.23221 \pm 0.00029$ (B) DISAGREE BY THREE SIGMA
- $\rightarrow$ AIM AT 0.0001 LEVEL ACCURACY!
- IS IT POSSIBLE AT HADRON COLLIDER WHERE PDF UNCERTAINTIES PLAY A ROLE?
- MEASUREMENT PERFORMED BY CMS WITH 40 pb^{-1} $\Rightarrow$ $\sin^2\theta_W = 0.2287 \pm 0.0077$ (STAT.) ± 0.0036 (SYSR.)



$\sin^2\theta_W$ FROM A_{FB} AT THE LHC?

- FORWARD/BACKWARD ASYMMETRY OF LEPTONS PRODUCED BY Z DECAY $\Rightarrow$ ELECTROWEAK MIXING ANGLE
- CURRENTLY $\sin^2\theta_W = 0.23153 \pm 0.00016$ FROM GLOBAL FIT BUT TWO MOST PRECISE MEASUREMENTS $\sin^2\theta_W = 0.23098 \pm 0.00026$ (SLD) & $\sin^2\theta_W = 0.23221 \pm 0.00029$ (B) DISAGREE BY THREE SIGMA
- $\rightarrow$ AIM AT 0.0001 LEVEL ACCURACY!
- IS IT POSSIBLE AT HADRON COLLIDER WHERE PDF UNCERTAINTIES PLAY A ROLE?
- MEASUREMENT PERFORMED BY CMS WITH 40 pb^{-1} $\Rightarrow$ $\sin^2\theta_W = 0.2287 \pm 0.0077$ (STAT.) ± 0.0036 (SYSR.)
- GREATER SENSITIVITY AT LARGE RAPIDITY $\Rightarrow$ LHCB PROPOSES MEASUREMENT FOR $2 \lesssim y \lesssim 4$ @ Z PEAK
- 0.0001 STAT. ACCURACY CAN BE ATTAINED WITH 50 fb^{-1}, BUT LARGE & SMALL x PDFS INVOLVED
- PDF+ THEORY (RESUMMATION) UNCERTAINTIES HARD!



OUTLOOK

WISHLIST ⇒ ROADMAP

- LHC CAN PROVIDE US PRECISION INFORMATION ON PDFs
- TOWARDS A "COLLIDER ONLY" HERA+LHC PDF FIT
  (TEVATRON DATA MIGHT BE SUPERFLUOUS)
  - MEDIUM & LARGE x GLUON
    - PROMPT PHOTONS AVAILABLE
    - (PRECISION) JETS IN PROGRESS
  - LIGHT FLAVORS AT MEDIUM @ SMALL x, FLAVOR SEPARATION @ SMALL x
    - LOW–MASS DRELL-YAN PRELIM.
    - Z RAPIDITY DISTRIBUTIONS PRELIM.
    - W ASYMMETRIES AVAILABLE
  - STRANGENES & HEAVY FLAVORS
    - STRANGENESS $\Rightarrow W + c$ FUTURE?
    - CHARM $\Rightarrow Z + c$, $\gamma + c$ FUTURE?
    - BOTTOM $Z + b$ IN PROGRESS

WISHLIST ⇒ ROADMAP

- LHC CAN PROVIDE US PRECISION INFORMATION ON PDFS
- TOWARDS A "COLLIDER ONLY" HERA+LHC PDF FIT
  (TEVATRON DATA MIGHT BE SUPERFLUOUS)
  - MEDIUM & LARGE x GLUON
    * PROMPT PHOTONS AVAILABLE
    * (PRECISION) JETS IN PROGRESS
  - LIGHT FLAVORS AT MEDIUM @ SMALL x, FLAVOR SEPARATION @ SMALL x
    * LOW–MASS DRELL-YAN PRELIM.
    * Z RAPIDITY DISTRIBUTIONS PRELIM.
    * W ASYMMETRIES AVAILABLE
  - STRANGENES & HEAVY FLAVORS
    * STRANGENESS $\Rightarrow W + c$ FUTURE?
    * CHARM $\Rightarrow Z + c$, $\gamma + c$ FUTURE?
    * BOTTOM $Z + b$ IN PROGRESS
- PRECISION "LEP" PHYSICS POSSIBLE @ LHC!
  - NEW PHYSICS FROM EW SECTOR
  - NEW QCD EFFECTS

EXTRAS

NMC DATA: A PROBLEM?

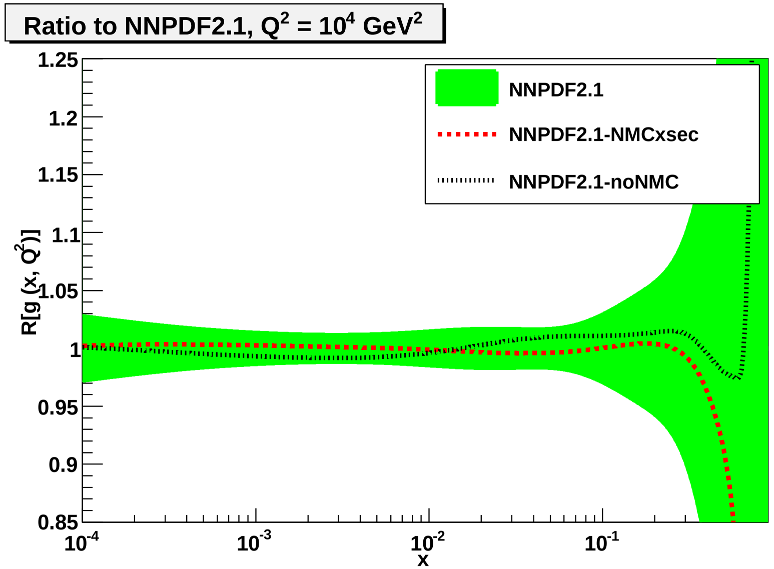
- NMC PROTON AND DEUTERON DATA FOR DIS F_2 STRUCTURE FUNCTION
- F_2STRUCTURE FUNCTION DATA USED IN MSTW, CTEQ, NNPDF FITS, WOULD THE USE OF CROSS SECTION DATA MAKE A DIFFERENCE? (Alekhin, Blümlein, Moch, 2011)
  $\widetilde{\sigma}(x, y, Q^2) = F_2(x, Q^2)\left(2 - 2y + y^2/[1 + R(x, Q^2)]\right)$

NMC DATA: A PROBLEM?

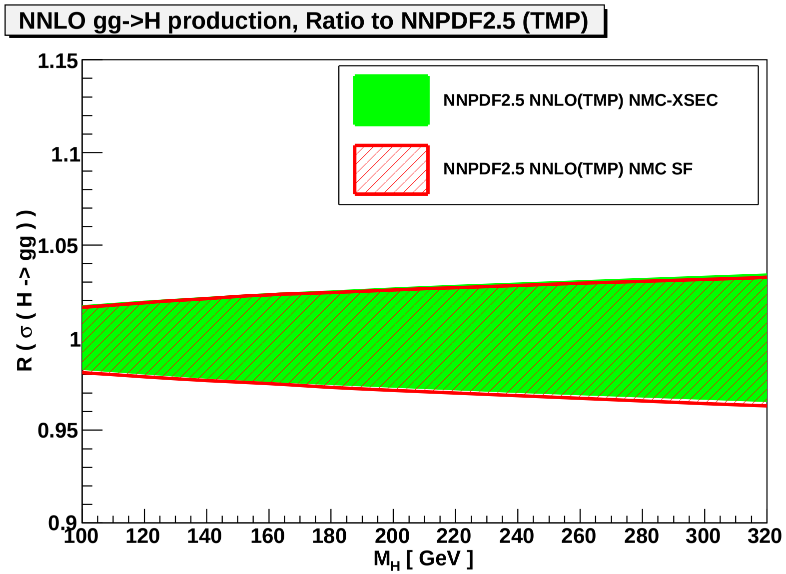
- NMC PROTON AND DEUTERON DATA FOR DIS F_2 STRUCTURE FUNCTION
- F_2STRUCTURE FUNCTION DATA USED IN MSTW, CTEQ, NNPDF FITS, WOULD THE USE OF CROSS SECTION DATA MAKE A DIFFERENCE? (Alekhin, Blümlein, Moch, 2011)
  $\widetilde{\sigma}(x, y, Q^2) = F_2(x, Q^2)\left(2 - 2y + y^2/[1 + R(x, Q^2)]\right)$

NNPDF FIT REPEATED WITH NMC XSECT OR NO NMC

EFFECT ON GLUON

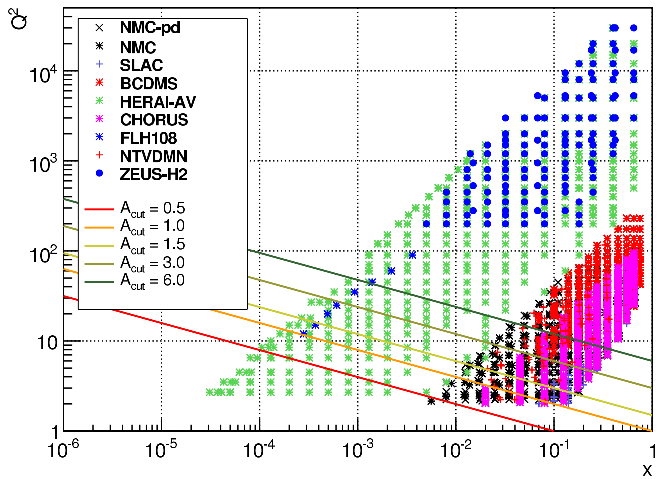


EFFECT ON HIGGS XSECT



NO SIGNIFICANT CHANGE OBSERVED

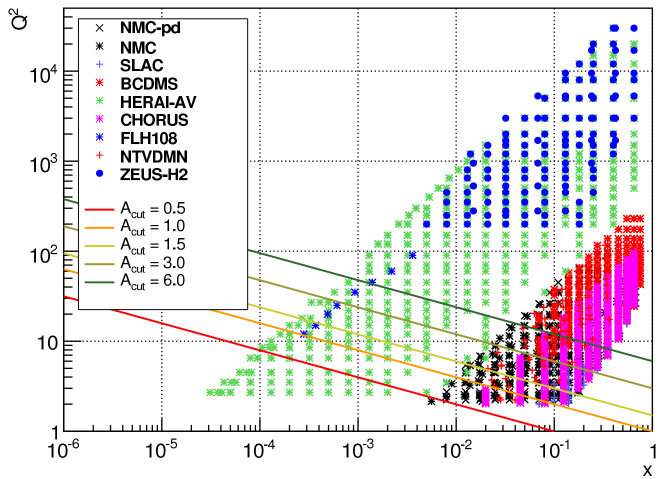
BEYOND DGLAP?
DISCOVERING A NEW QCD EFFECT IN HERA DATA



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

- CUT OUT DATA IN THE "DANGEROUS" (SMALL x) 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

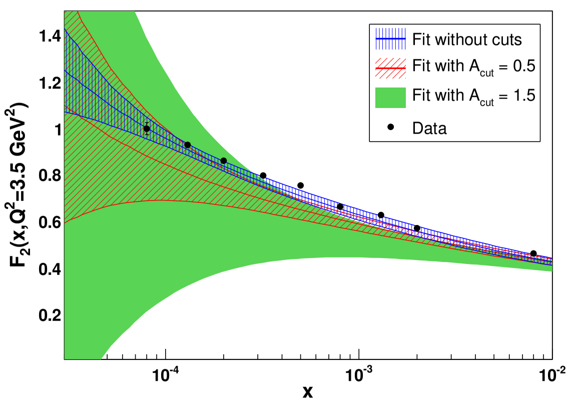


IDEA: (Gélis, 2008, $\Rightarrow$ Caola, s.f. ,Rojo 2010)

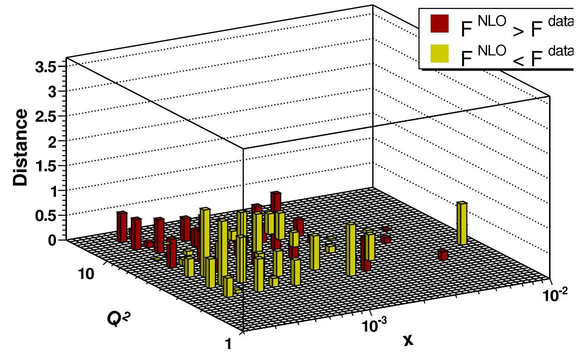
- CUT OUT DATA IN THE "DANGEROUS" (SMALL x) REGION
- DETERMINE PDFs IN THE "SAFE" (LARGE x AND Q^2) REGION
- EVOLVE BACKWARDS AND COMPARE TO DATA

OLD HERA DATA

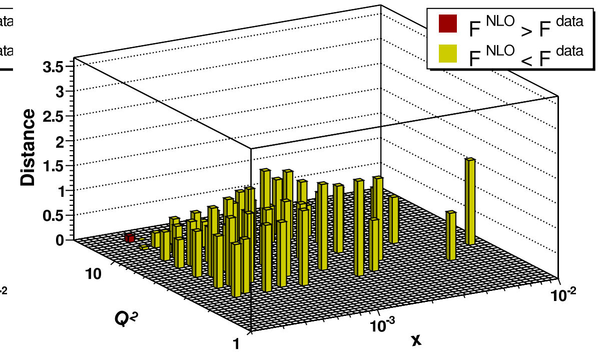
BACKWARD EV. VS DATA



DAT/TH DIST: NO CUT



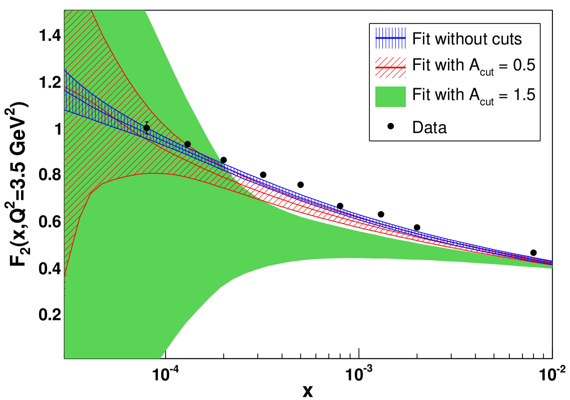
DAT/TH DIST: CUT



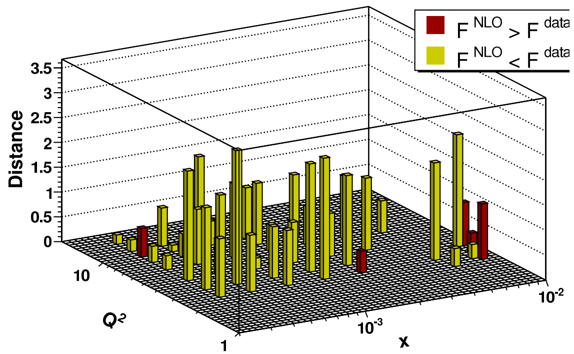
- BACKWARD EVOLVED FIT LIES SYSTEMATICALLY BELOW DATA
- DATA AT LOW x AND Q^2 SHOW LESS EVOLUTION THAN PREDICTED BY NLO DGLAP
- IF LOW x AND Q^2 DATA INCLUDED, THE FIT COMPENSATES READJUSTING PDFs

NEW COMBINED HERA DATA

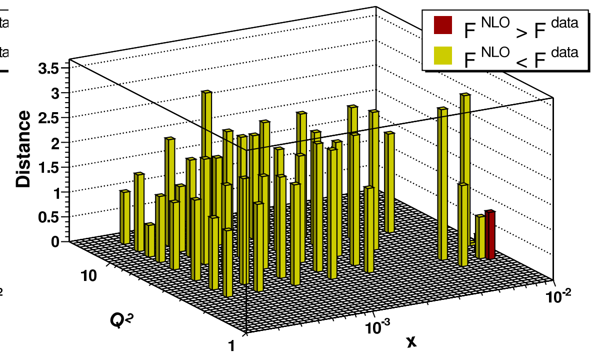
BACKWARD EV. VS DATA



DAT/TH DIST: NO CUT



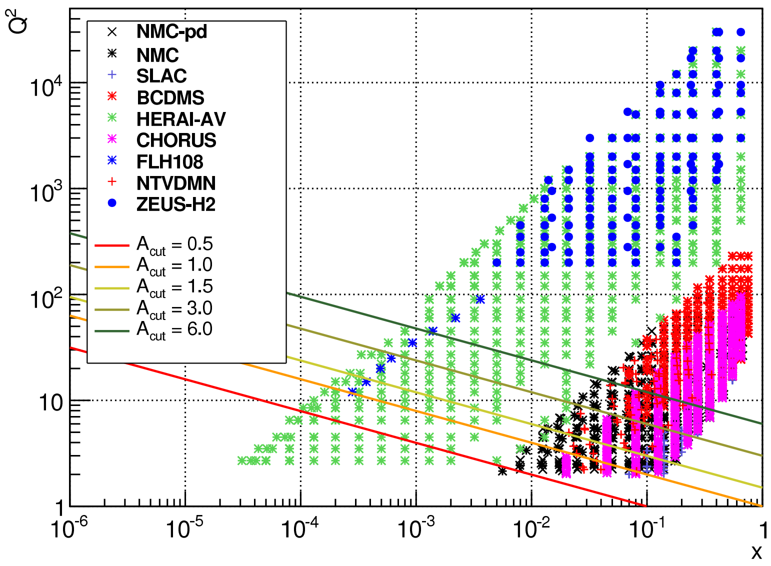
DAT/TH DIST: CUT

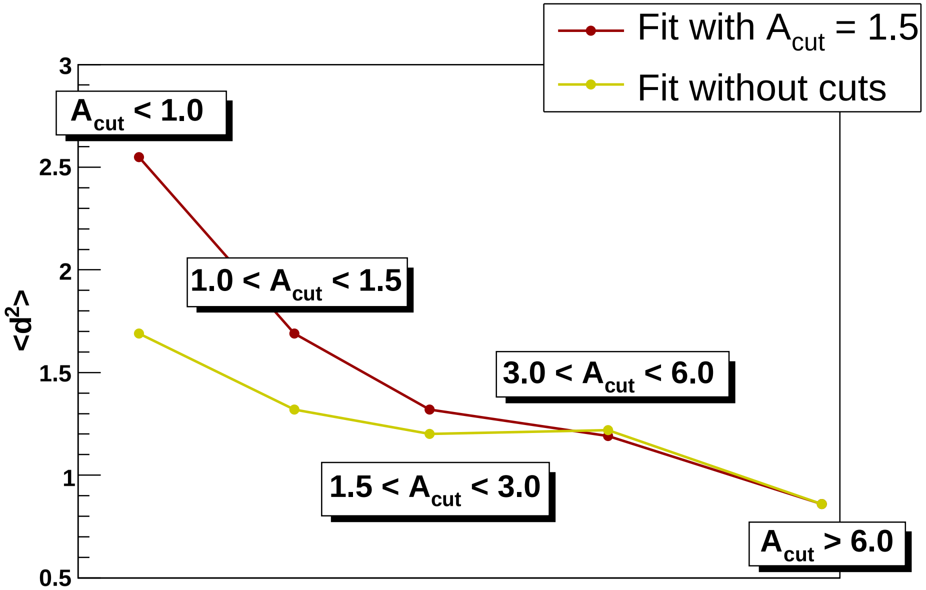


- DATA AT LOW x AND Q^2 SHOW LESS EVOLUTION THAN PREDICTED BY NLO DGLAP
- BACKWARD EVOLVED FIT LIES SYSTEMATICALLY BELOW DATA
- WITH MORE PRECISE DATA, THE FIT NO LONGER MANAGES TO COMPENSATE BY READJUSTING THE PDFS: EVEN FULL FIT LIES BELOW DATA

DETERIORATION IN FIT QUALITY:

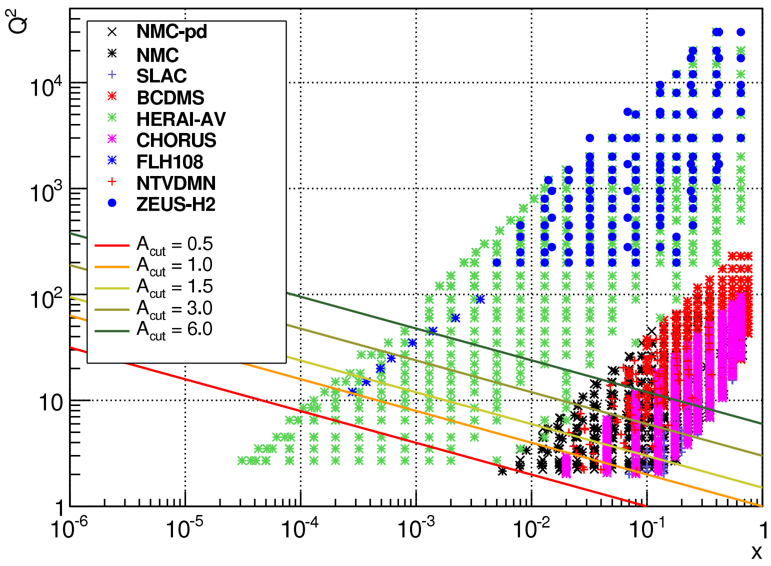
$\langle d^2 \rangle$ VS x SLICES



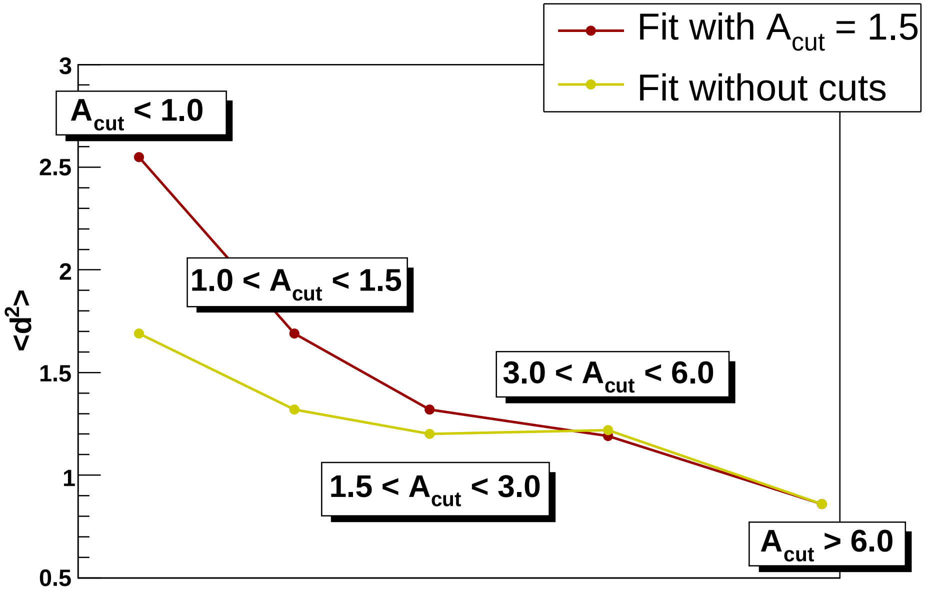


- QUALITY OF UNCUT FIT DETERIORATES IN LOW x REGIONS
- QUALITY OF CUT FIT INCREASINGLY POOR AS x DECREASES
- DISTANCE RISES DESPITE HUGE INCREASE IN UNCERTAINTY

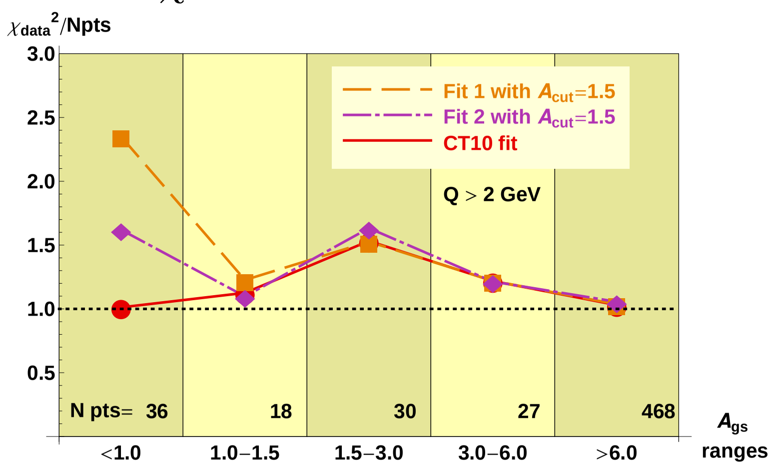
DETERIORATION IN FIT QUALITY:



$\langle d^2 \rangle$ VS x SLICES



χ^2 VS x SLICES FOR CT10



- QUALITY OF UNCUT FIT DETERIORATES IN LOW x REGIONS
- QUALITY OF CUT FIT INCREASINGLY POOR AS x DECREASES
- DISTANCE RISES DESPITE HUGE INCREASE IN UNCERTAINTY
- IN HESSIAN FIT (CTEQ) RESULTS DEPEND ON PARAMETRIZATION $\Rightarrow$ EVIDENCE INCONCLUSIVE

THE STANDARD (HESSIAN) SOLUTION

(MSTW, CTEQ):

FUNCTIONAL PARTON FITTING

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

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

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

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

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

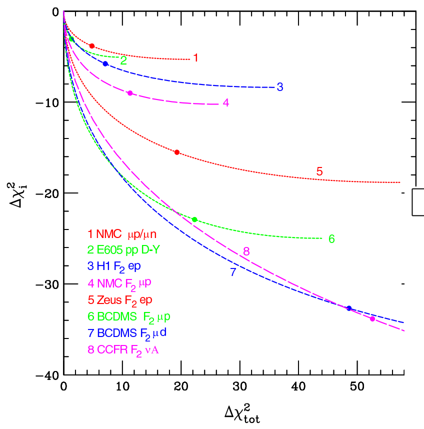
    a_6, a_7 ONLY USED FOR GLUON
  - BASIS FUNCTIONS: $u_v \equiv u-\bar{u}$, $d_v \equiv d-\bar{d}$, $\bar{u}\pm\bar{d}$ (MSTW) OR $\bar{u}$, $\bar{d}$ (CTEQ), $s^{\pm} \equiv s\pm\bar{s}$ (CTEQ $\bar{s}+s$ ONLY), g.
- EVOLVE TO DESIRED SCALE & COMPUTE PHYSICAL OBSERVABLES
- DETERMINE BEST-FIT VALUES OF PARAMETERS
- DETERMINE ERROR BY PROPAGATION OF ERROR ON PARMS. ('HESSIAN METHOD') PARM. SCANS ALSO POSSIBLE ('LAGR. MULTIPLIER METHOD')

THE TOLERANCE PROBLEM

MSTW/CTEQ: ONE–σ IS DEFINED UP TO A "TOLERANCE"

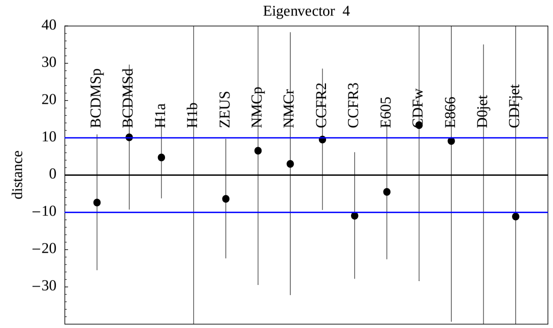
- STANDARD $\Delta\chi^2 = 1$ BANDS TOO NARROW $\Rightarrow$ SPREAD OF χ^2 FOR INDIVIDUAL EXPERIMENTS MUCH TOO LARGE!
- TOLERANCE $\Rightarrow$ ENVELOPE OF UNCERTAINTIES OF EXPERIMENTS
- DYNAMICAL $\Rightarrow$ SEPARATELY DETERMINED FOR EACH HESSIAN EIGENVECTOR

MINIMUM χ_i^2 VS GLOBAL χ^2

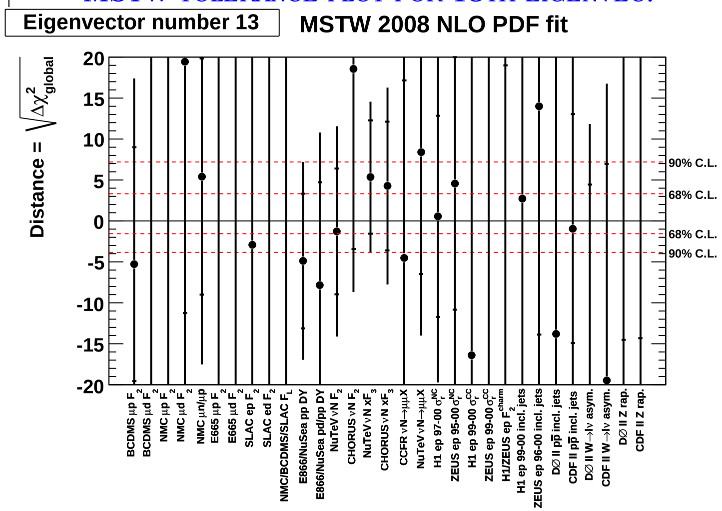


Collins, Pumplin 2001

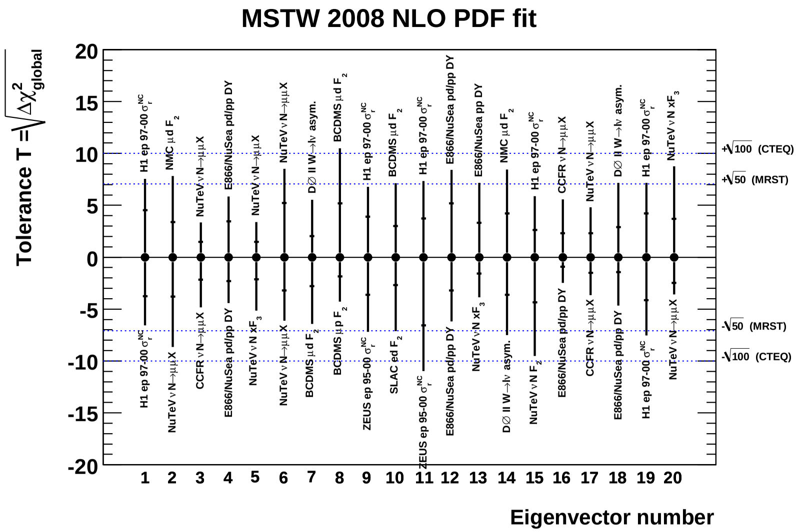
CTEQ TOLERANCE PLOT FOR 4TH EIGENVEC.



MSTW TOLERANCE PLOT FOR 13TH EIGENVEC.



GLOBAL MSTW TOLERANCE

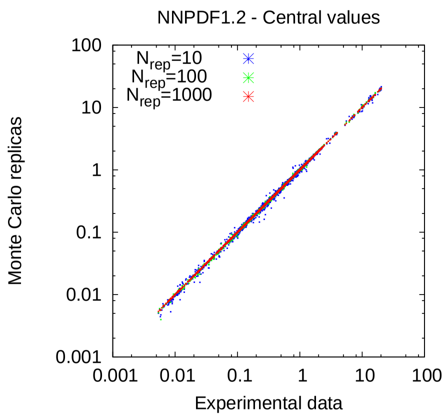


THE NNPDF APPROACH:
THE NEURAL MONTE CARLO

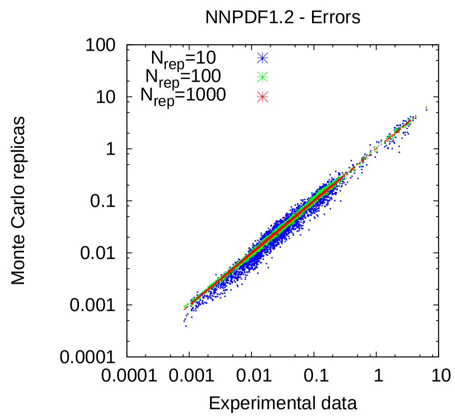
BASIC IDEA: MONTE CARLO SAMPLING OF THE PROBABILITY MEASURE IN THE (FUNCTION) SPACE OF PDFs

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

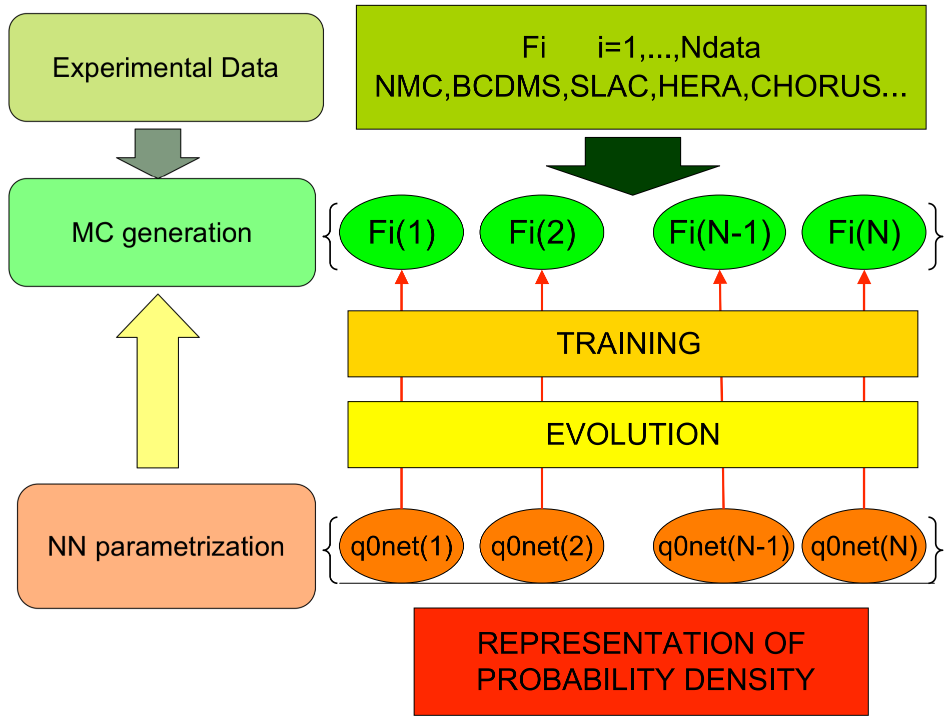
replica averages vs. central values



replica standard dev. vs. uncertainties

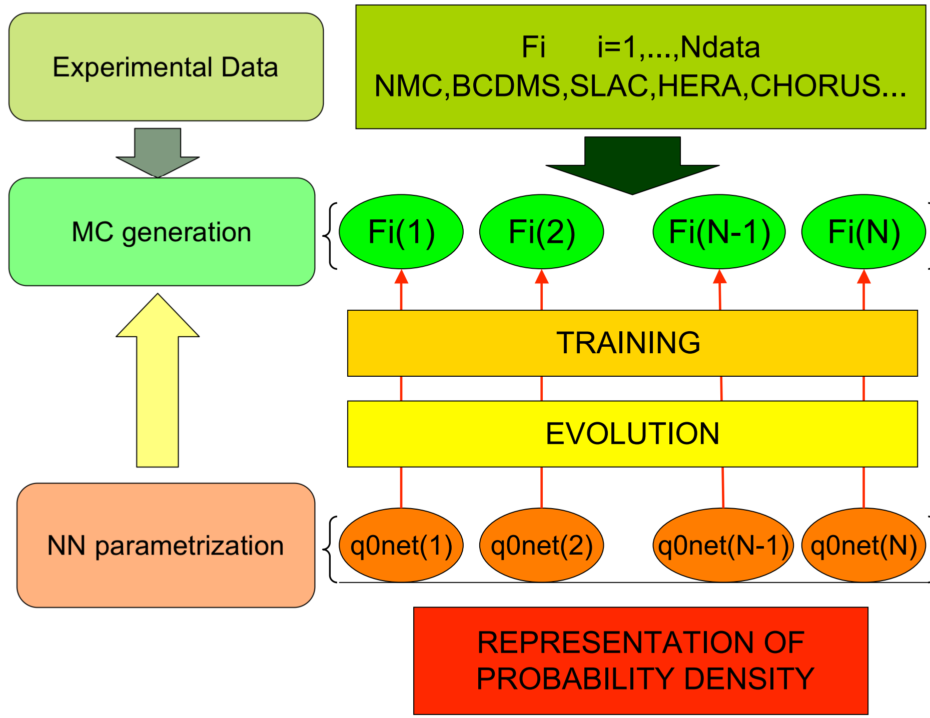


10 REPLICAS ENOUGH FOR CENTRAL VALS, 100 FOR UNCERTAINTIES, 1000 FOR CORRELNS



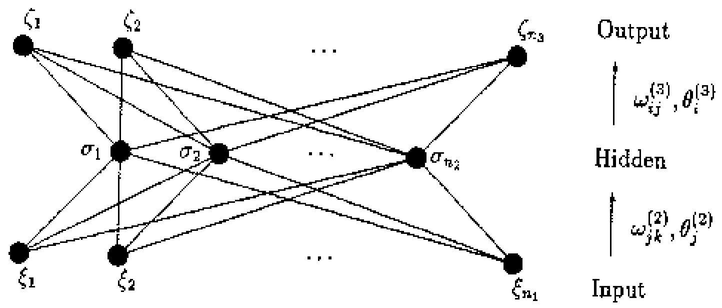
DATA MONTE CARLO $\Rightarrow$ PDF MONTE CARLO
NEURAL NETWORK PARM+ CROSS-VALIDATION METHOD

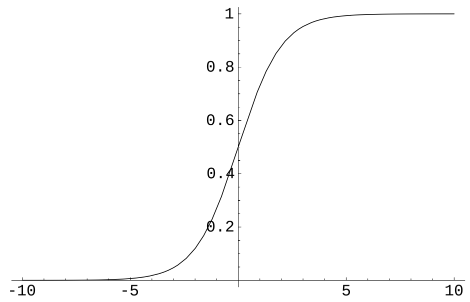
- EACH PDF$\leftrightarrow$ NEURAL NETWORK PARAMETRIZED BY 37 PARAMETERS
- NNPDF1.X, NNPDF2.X: $37 \times 7 = 259$ PARMS (RECALL MSTW, CTEQ $\rightarrow$ 20 FREE PARAMETERS) "INFINITE" NUMBER OF PARAMETERS$\Rightarrow$ CAN REPRESENT ANY FUNCTION
- COMPLEX SHAPES (LARGE NO.OF PARAMETERS) REQUIRE LONGER 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")



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

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

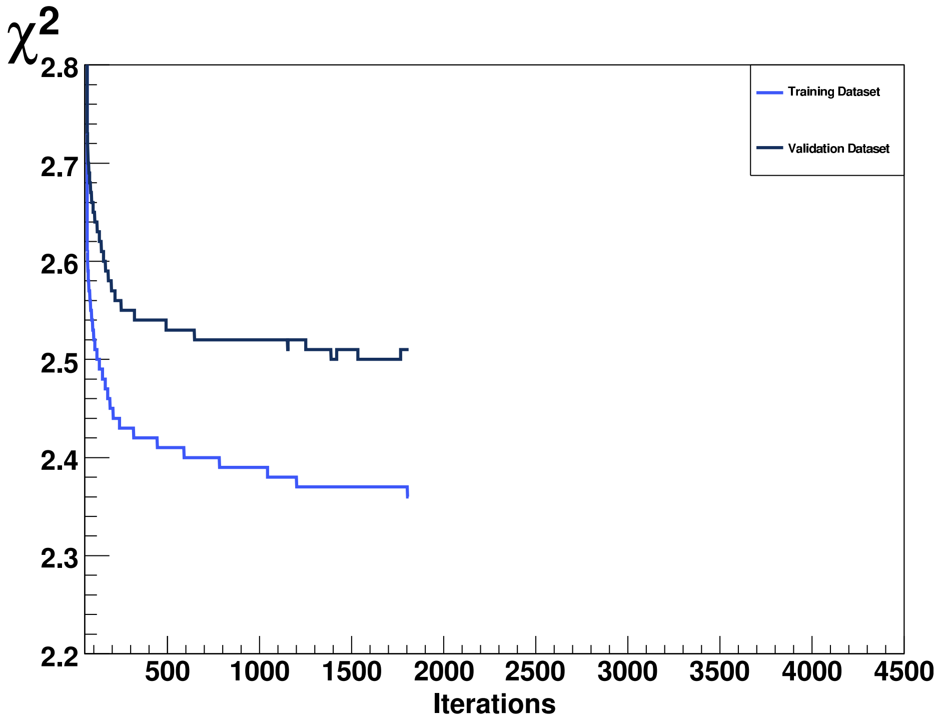
- ANY FUNCTION CAN BE REPRESENTED BY A SUFFICIENTLY BIG NEURAL NETWORK $\Rightarrow$ CAN CHECK INDEPENDECE OF SIZE
- MINIMZATION PERFORMED BY GENETIC ALGORITHM
- INFORMATION CONTAINED ("COARSENESS") IN NN INCREASES DURING MINIMIZATION

DATA MONTE CARLO ⇒ PDF MONTE CARLO
CROSS-VALIDATION

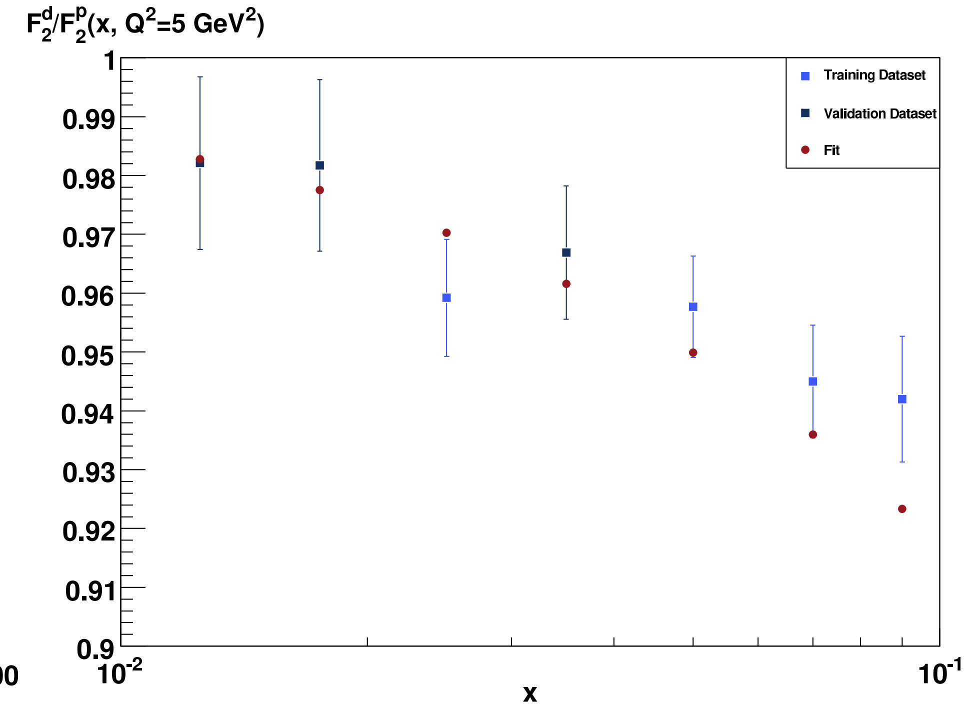
- REPLICAS ARE FITTED TO A DATA SUBSET
- A DIFFERENT SUBSET OF DATA USE FOR EACH REPLICA
- OPTIMAL FIT WHEN FIT TO VALIDATION (CONTROL) DATA STOPS IMPROVING
-

OPTIMAL FITTING

χ^2



FIT TO DATA

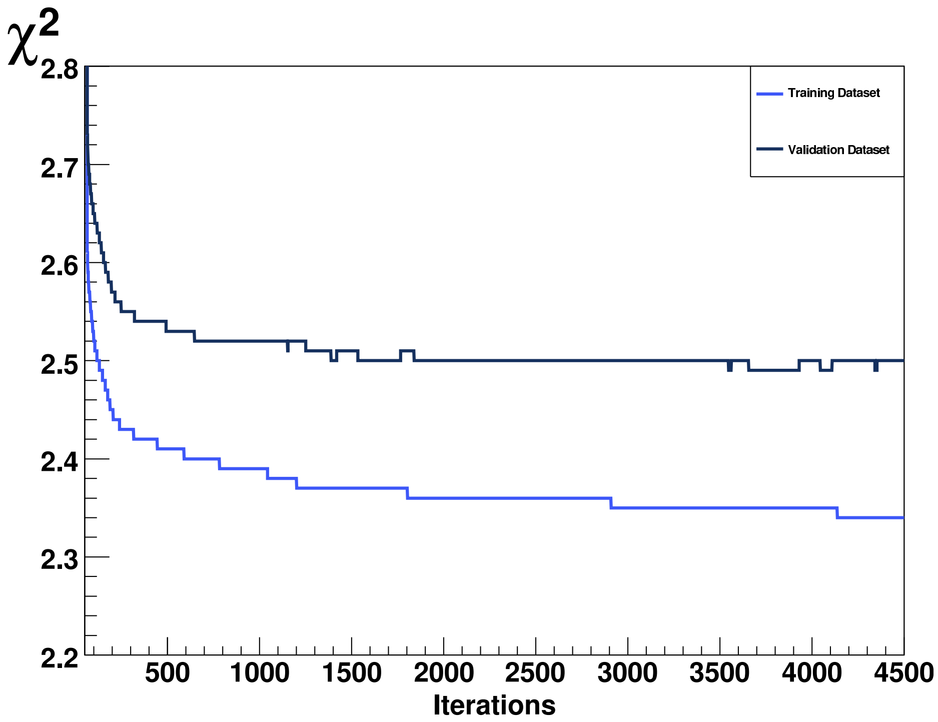


DATA MONTE CARLO ⇒ PDF MONTE CARLO
CROSS-VALIDATION

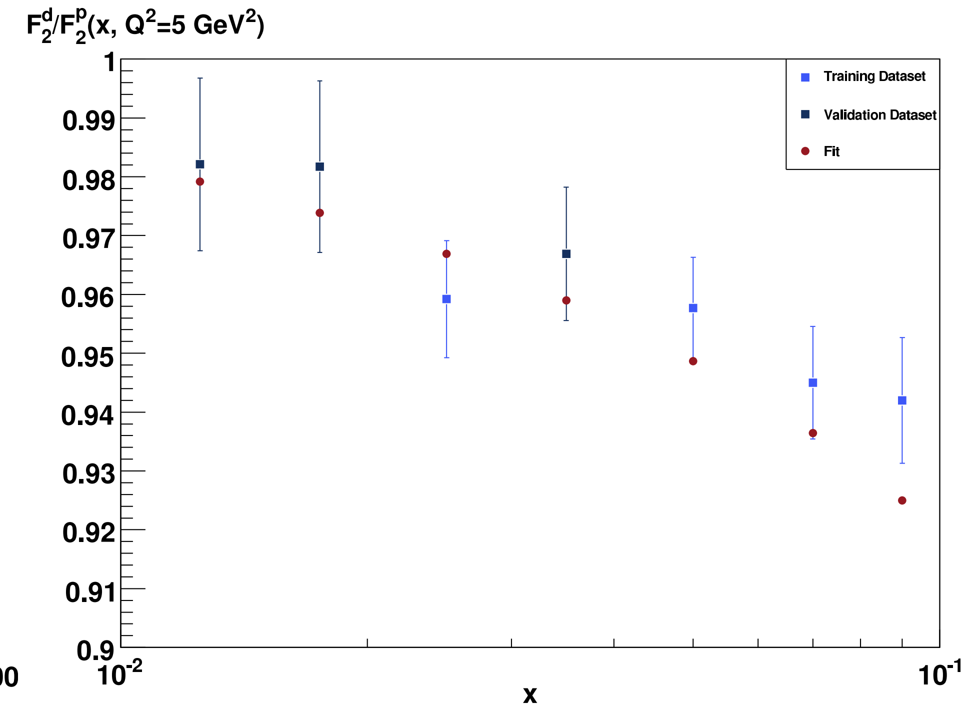
- REPLICAS ARE FITTED TO A DATA SUBSET
- A DIFFERENT SUBSET OF DATA USE FOR EACH REPLICA
- OPTIMAL FIT WHEN FIT TO VALIDATION (CONTROL) DATA STOPS IMPROVING
- THE BEST FIT IS NOT AT THE MINIMUM OF THE χ^2

OVERFITTING

χ^2



FIT TO DATA

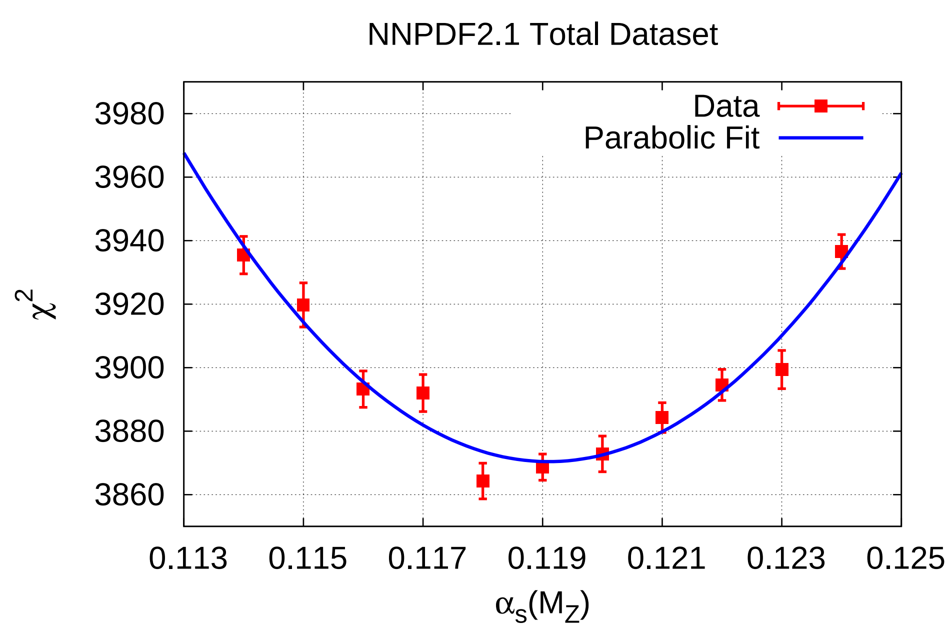


THE RESULT: GLOBAL FIT

NLO NNPDF2.1 GLOBAL DETERMINATION (ONLY STAT. ERROR KNOWN)

$N_{\rm rep} = 500$ PER VALUE OF α_s

$$\alpha_s(M_z) = 0.1191 \pm 0.0006(\text{stat.}) \pm 0.0001(\text{proc.})$$



χ^2/d.o.f. $= 1.6$ for the parabolic fit

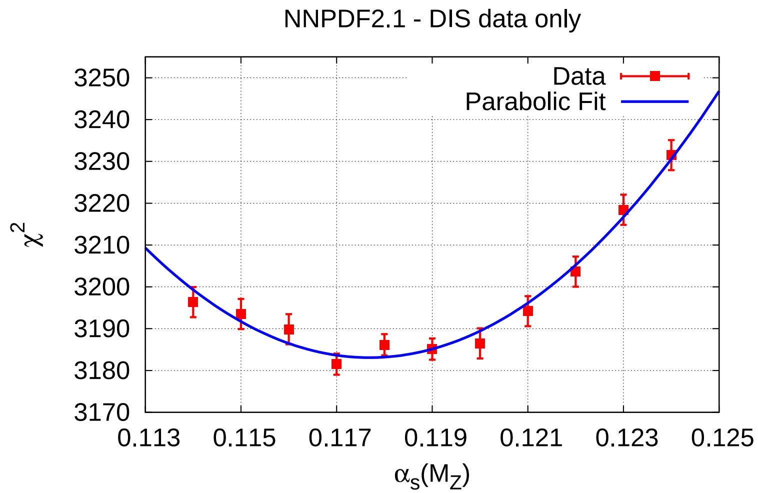
THE RESULT: DIS ONLY

DO DIS DATA PREFER A SMALLER VALUE?

NLO NNPDF2.1 DEEP-INELASTIC DATA

$N_{\rm rep} = 500$ PER VALUE OF α_s

$$\alpha_s(M_z) = 0.1177 \pm 0.0009(\text{stat.}) \pm 0.0002(\text{proc.})$$

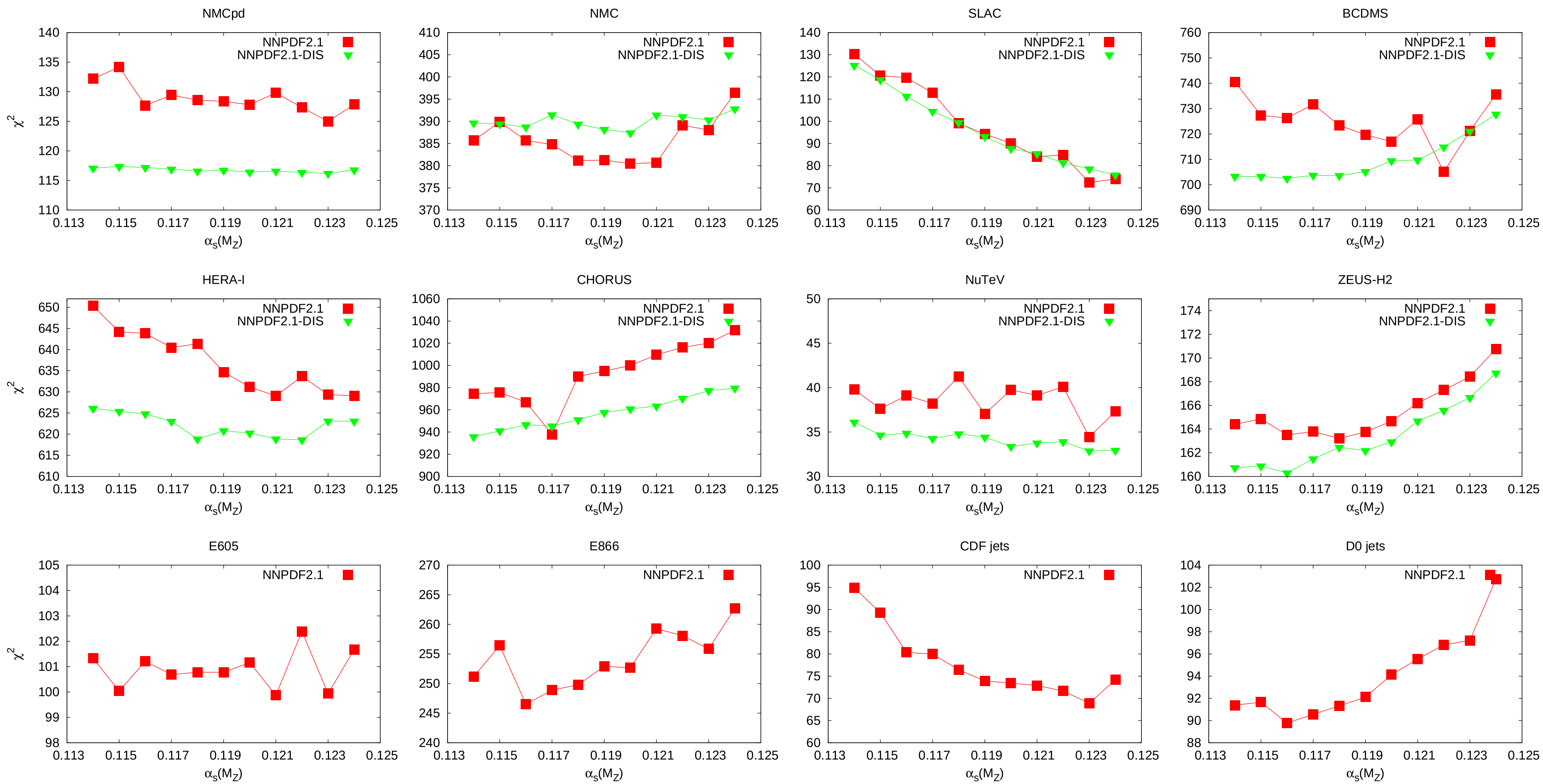


$$\chi^2/\text{d.o.f.} = 0.8 \text{ for the parabolic fit}$$

- YES
- BUT NOT MUCH SMALLER & WITH LARGER UNCERTAINTY (COMPATIBLE WITHIN UNCERTAINTIES, AS IT OUGHT TO)

INDIVIDUAL EXPERIMENTS: WHAT'S GOING ON?

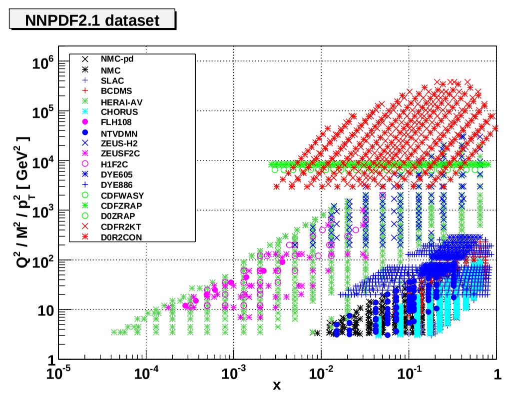
χ^2 PROFILE FOR INDIVIDUAL EXPERIMENTS IN GLOBAL & DIS FITS



- BCDMS SIMILAR TO HERA, NMC PRETTY FLAT
- DIS EXPERIMENTS (BCDMS+HERA) IN DIS FIT HAVE A "RUNAWAY DIRECTION" AT SMALL α_s,
  ABSENT IN GLOBAL FIT
- JET EXPERIMENTS EXCLUDE SMALL α_s

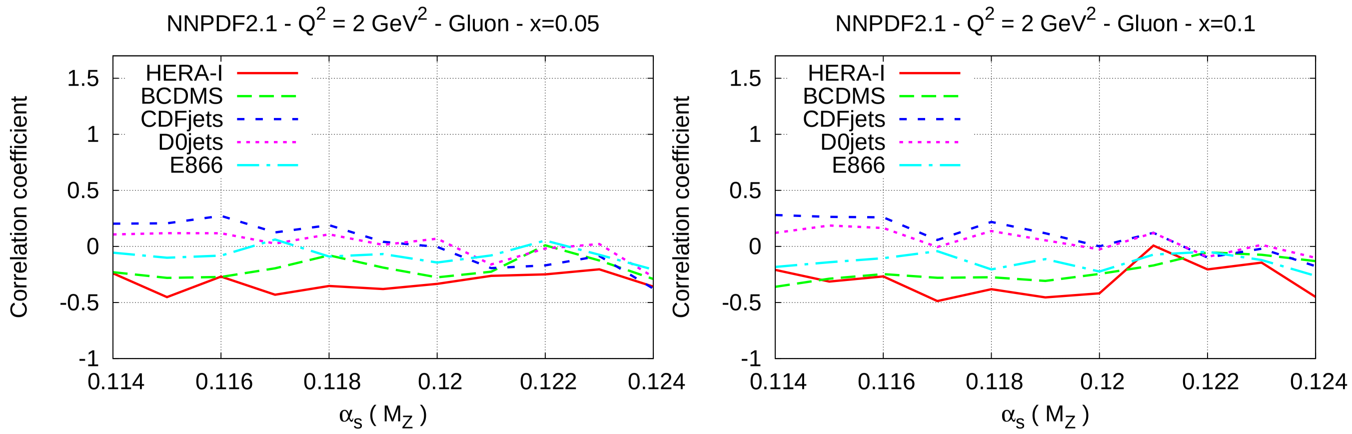
IS THERE A PROBLEM WITH DIS?

CONJECTURE



- IN DIS, GLUON DETERMINED BY SCALING VIOLATIONS $\Rightarrow$ CAN COMPENSATE SMALLER α_s WITH LARGER GLUON (OR CONVERSELY)
  $\Rightarrow$ RUNAWAY DIRECTIONS POSSIBLE IN DIS FIT
- JET DATA ARE AT MUCH LARGER SCALE, IF α_s TOO SMALL OR LARGE HIGH-SCALE GLUON WILL COME OUT WRONG
  $\Rightarrow$ RUNAWAY DIRECTION QUENCHED IN GLOBAL FIT

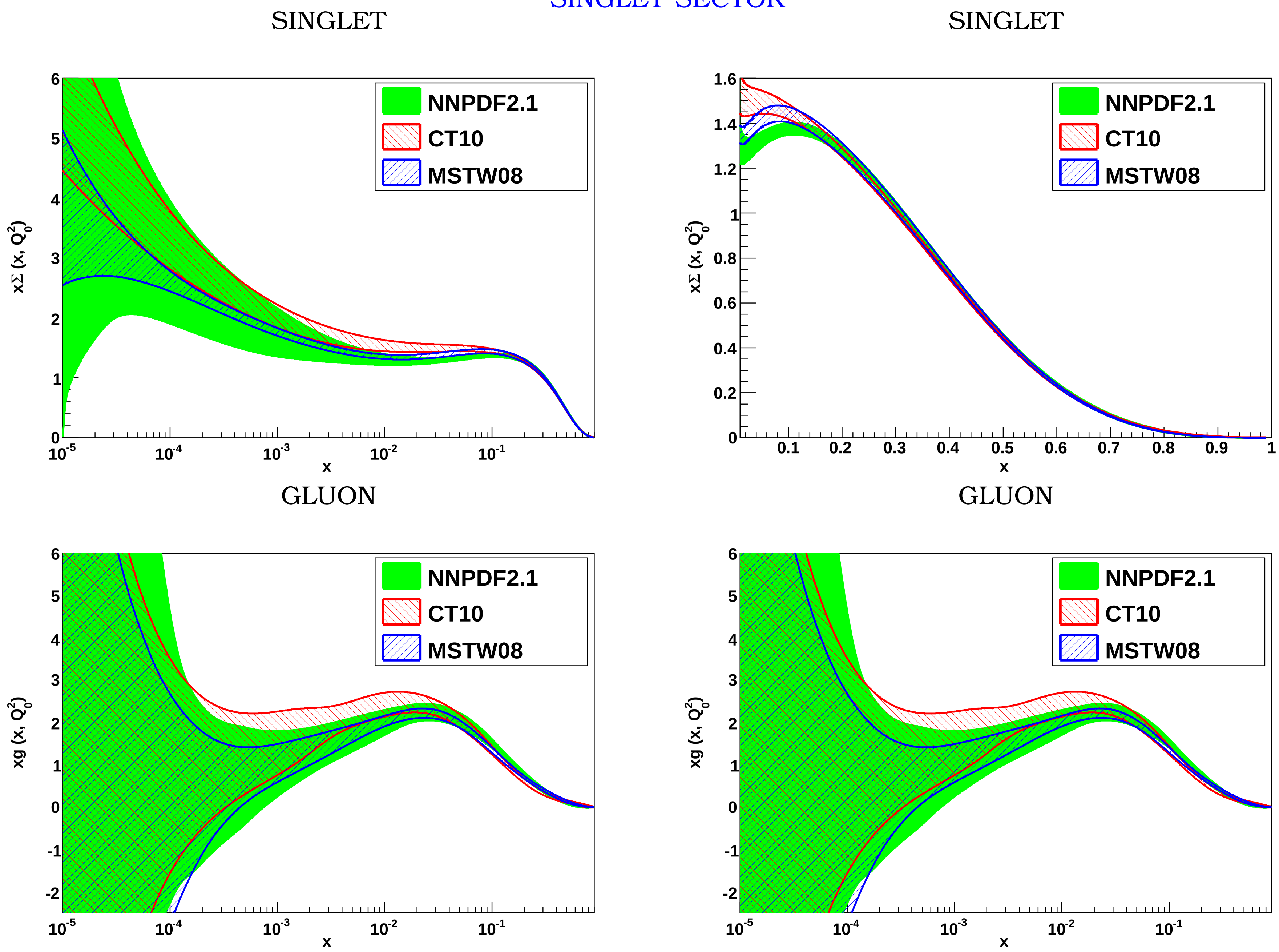
THE EVIDENCE



- COMPUTE THE CORRELATION BETWEEN χ^2 & PDFs AT THE GLOBAL BEST FIT
  LARGE (POSITIVE OR NEGATIVE) CORRN. $\Leftrightarrow$ RUNAWAY DIRECTION
- OPPOSITE SIGN CORRELATIONS $\Leftrightarrow$ DATA PULLING IN OPPOSITE DIRECTIONS
- AT LOW $\alpha_s = 0.114$ DIS FIT HAS A RUNAWAY DIRECTION STABILIZED BY JETS
- AT HIGH $\alpha_s = 0.122$ FIT IS GENERALLY STABLE,
  THOUGH STILL SOME PULL IN HERA LOW x

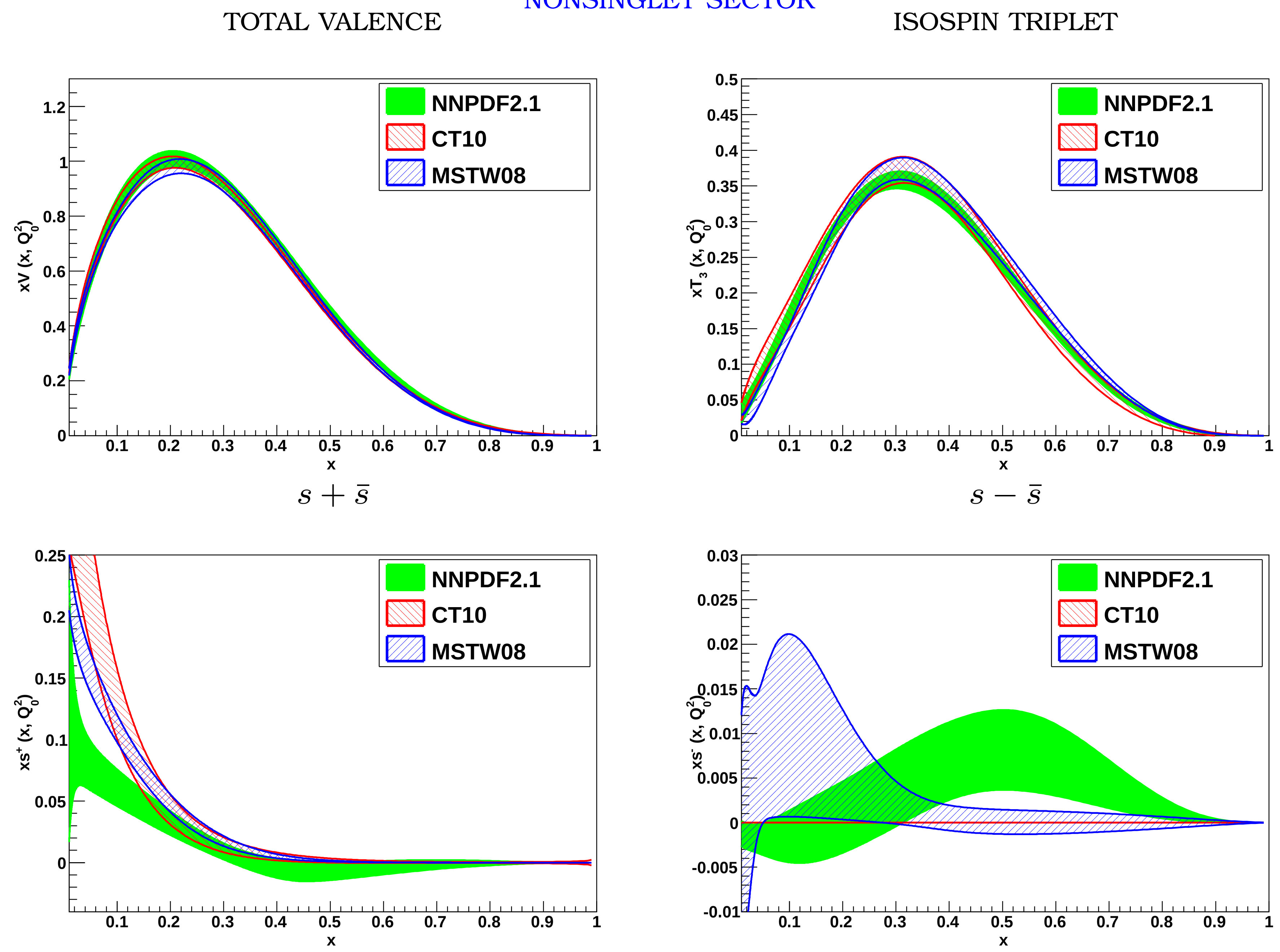
NNPDF2.1, MSTW08, CT10 COMPARISON

SINGLET SECTOR



NNPDF2.1, MSTW08, CT10 COMPARISON

NONSINGLET SECTOR



Monte Carlo data generation

- BCDMS+ NMC proton & deuteron F_2 data (full correlated systematics available), taken at 4 beam energies

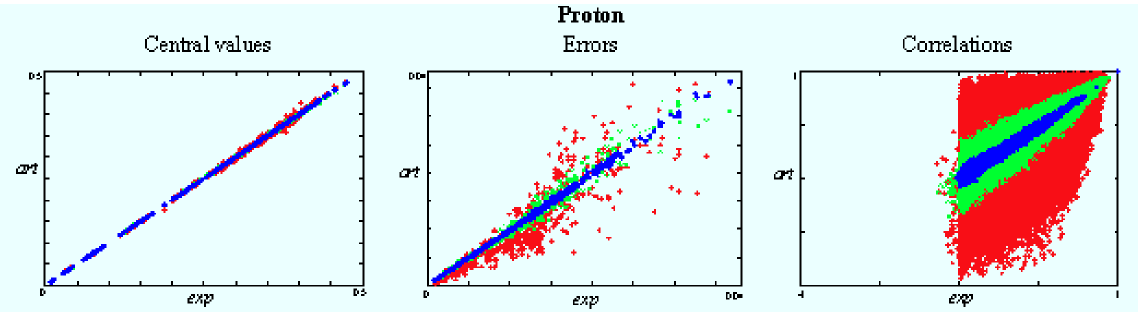
- on top of stat. errors, 4 systematics + 1 normalization (NMC) or 6 systematics + 1 absolute & 2 relative normalizations (BCDMS), with various forms of correlation (full, or for each target, or for each beam energy)

Generate data according to a multigaussian distribution

$$F_i^{(art)\,(k)} =$$

$$(1 + r_5^{(k)}\,\sigma_N)\sqrt{1 + r_{i,6}^{(k)}\,\sigma_{N_t}}\sqrt{1 + r_{i,7}^{(k)}\,\sigma_{N_b}}\left[F_i^{(exp)} + \frac{r_{i,1}^{(k)}\,f_b + r_{i,2}^{(k)}\,f_{i,s} + r_{i,3}^{(k)}\,f_{i,r}}{100}F_i^{(exp)} + r_{i,s}^{(k)}\,\sigma_s^i\right]$$

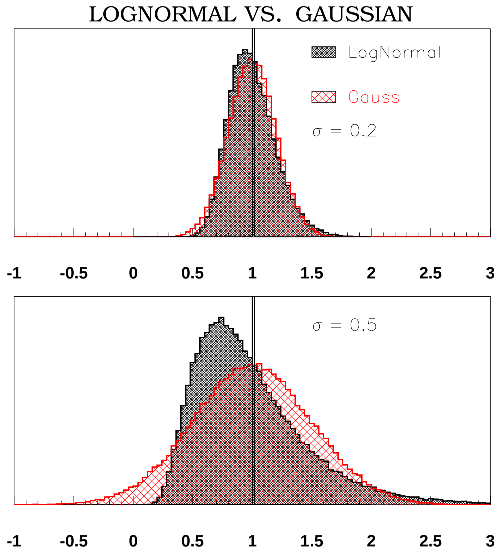
r univariate gaussian random nos., one $r_{i,s}$ for each data, but single $r_{i,j}$ for all correlated data



scatter plot art. vs. exp. for 10 (red) 100 (green) and 1000 (blue) replicas

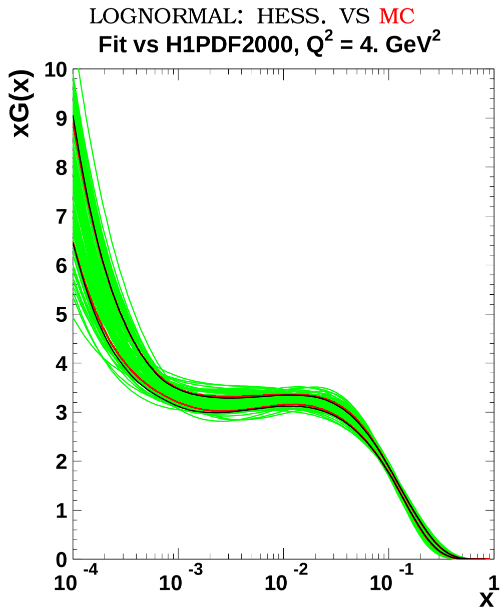
need 1000 replicas to reproduce correlations to percent accuracy

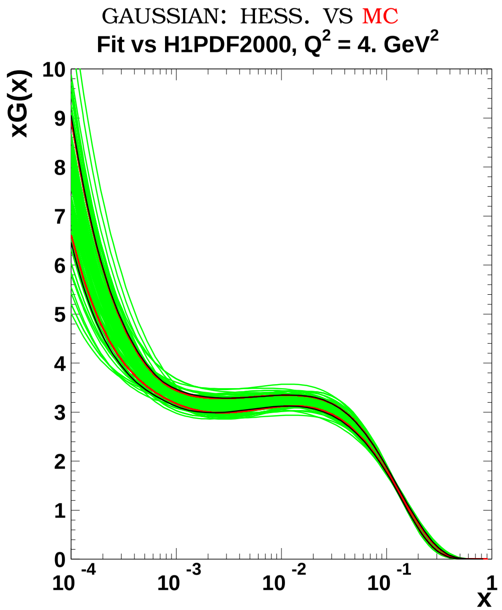
PARAMETRIZATION UNCERTAINTIES?
NONGAUSSIAN BEHAVIOUR?



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

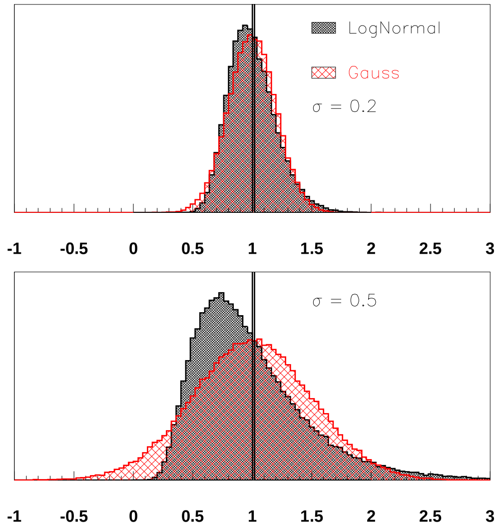




- NO DIFFERENCE BETWEEN LOGNORMAL, GAUSSIAN, MC, HESSIAN

PARAMETRIZATION UNCERTAINTIES?
NONGAUSSIAN BEHAVIOUR?

LOGNORMAL VS. GAUSSIAN

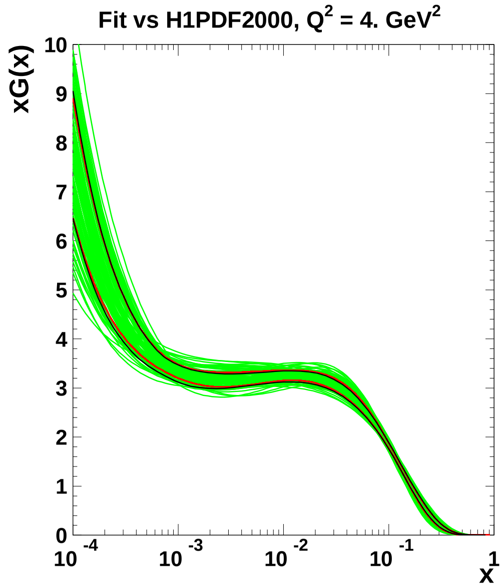


THE HERALHC BENCHMARK

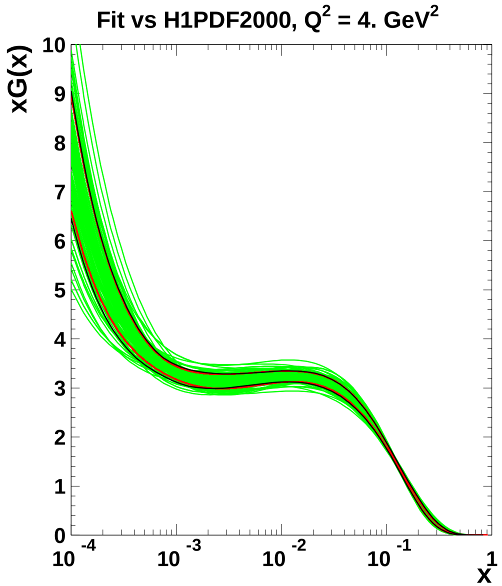
(Feltesse, Glazov, Radescu + NNPDF 2008)

- 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

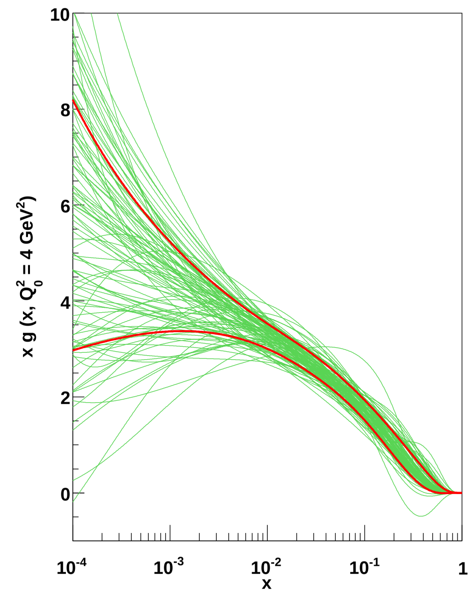
LOGNORMAL: HESS. VS MC



GAUSSIAN: HESS. VS MC



NNPDF

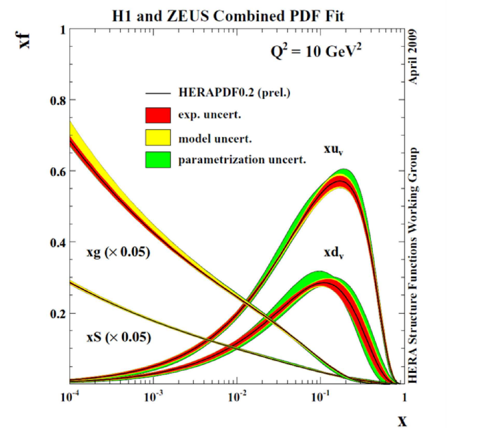


- NO DIFFERENCE BETWEEN LOGNORMAL, GAUSSIAN, MC, HESSIAN
- SIZABLE DIFFERENCE WR TO FLEXIBLE NNPDF PARAMETRIZATION

PARAMETRIZATION UNCERTAINTIES?

EXPLORING THE SPACE OF PARAMETERS: HESSIAN APPROACH

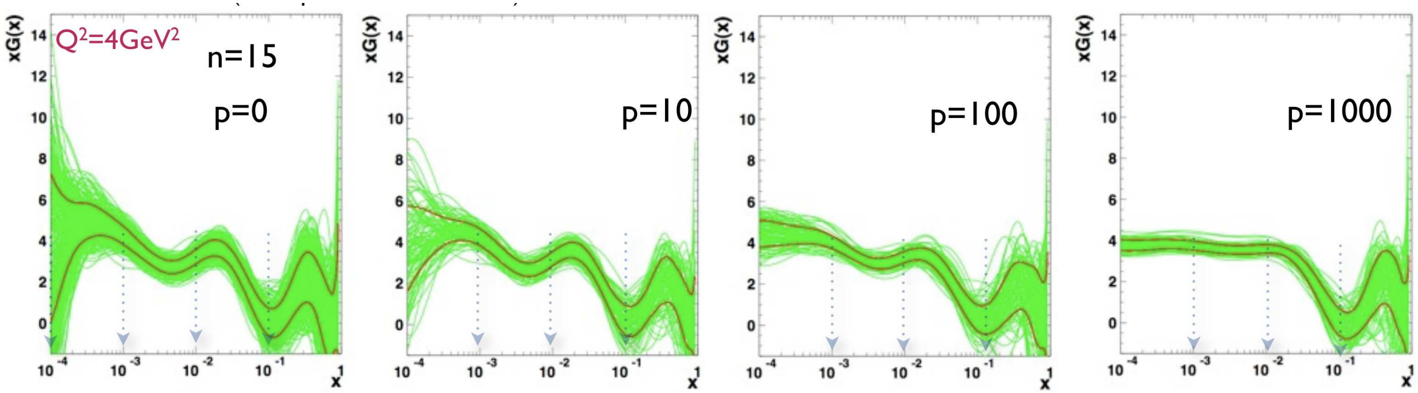
- IN HESSIAN APPROACH CAN VARY THE FUNCTIONAL FORM, ASSUMPTIONS, STARTING SCALE
- DONE IN THE HERAPDF1.0 FIT: VARIATION OF STRANGENESS FRACTION, LARGE x BEHAVIOUR, HIGHER ORDER POLYNOMIAL TERMS
- NO TOLERANCE ($\Delta\chi^2 = 1$), UNCERTAINTY DOUBLED



- 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



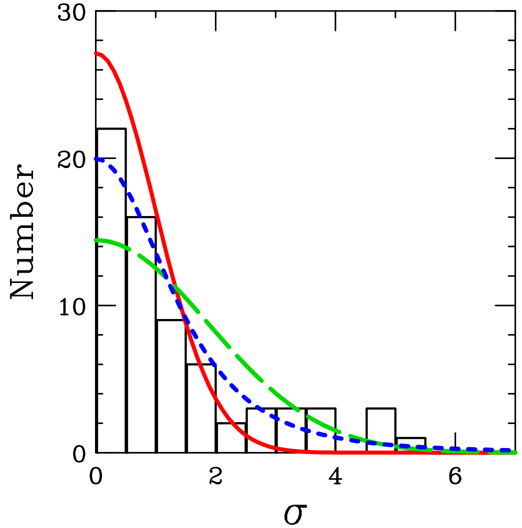
(Glazov, Radescu, 2009)

WHERE IS THE UNCERTAINTY COMING FROM?

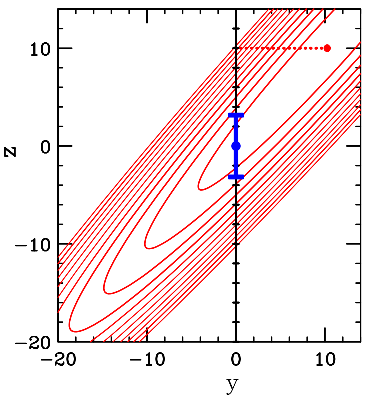
WHY DOES ONE NEED LARGE TOLERANCES?

DATA INCOMPATIBILITY (Pumplin, 2009)

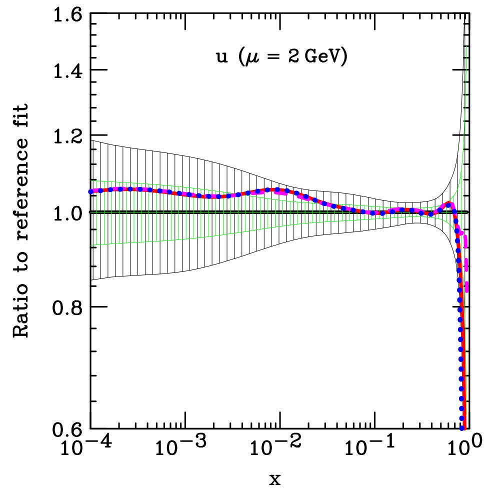
- CAN "REDIAGONALIZE": DIAGONALIZE SIMULTANEOUSLY χ^2 FOR TOTAL AND i–TH EXPT $\Rightarrow$ COMPATIBILITY OF EACH EXPT WITH GLOBAL FIT
- STUDY DISTRIBUTION OF DISCREPANCIES
- APPROX. GAUSSIAN WITH UNCERTAINTIES RESCALED BY 2 $\Rightarrow$ $\Delta\chi^2 \sim 10$ FOR 90%C.L.



FUNCTIONAL BIAS (Pumplin, 2010)



- 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. COMPATIBLE WITH $\Delta\chi^2 = 100$ RANGE OF CT10 PARM.



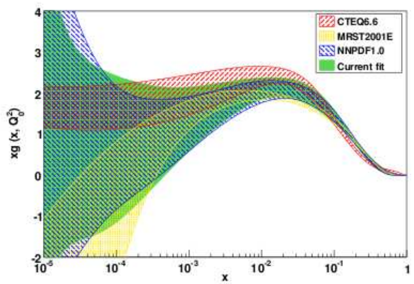
WHERE IS THE UNCERTAINTY COMING FROM?

FIT TO REPLICAS VS RANDOM SUBSET OF CENTRAL VAL.S

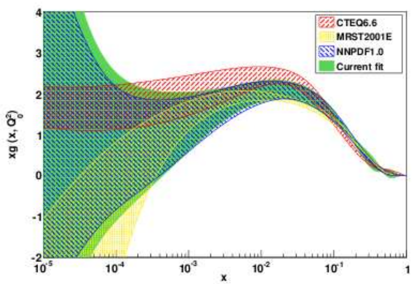
	REPLICAS	CENTRAL V.
χ^2	1.32	1.32
$\langle\chi^2\rangle_{\rm rep}$	2.79 ± 0.24	1.65 ± 0.20
$\langle\sigma^{\rm dat}\rangle$	0.039	0.035

GLUE

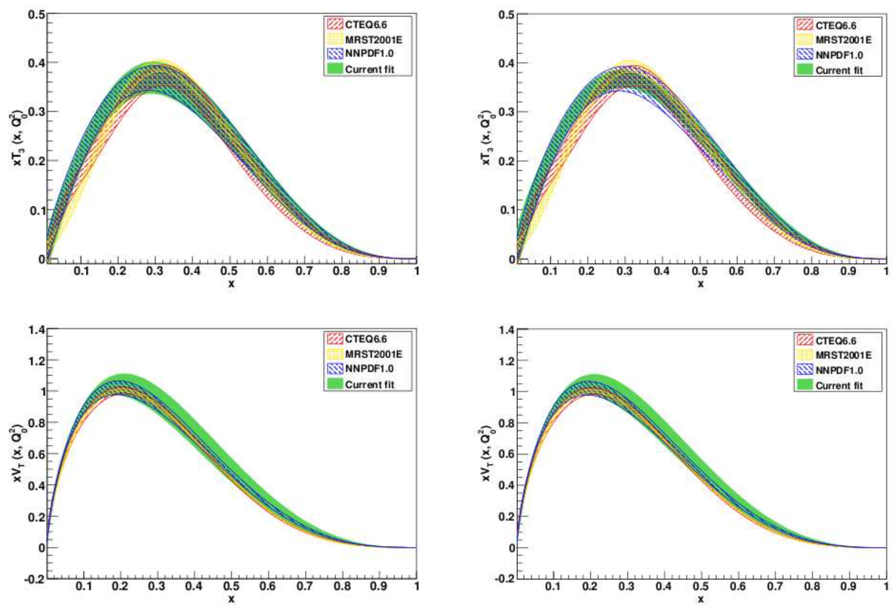
replicas



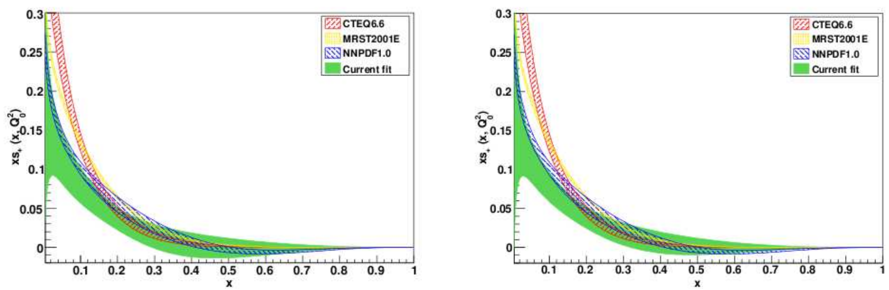
c. vals.



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

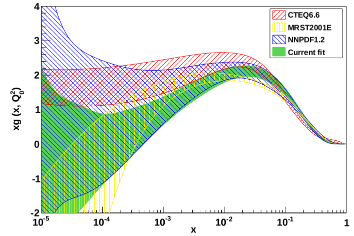
WHERE IS THE UNCERTAINTY COMING FROM?

CENTRAL VALUES: VARYING PARTITION VS FIXED PARTITION

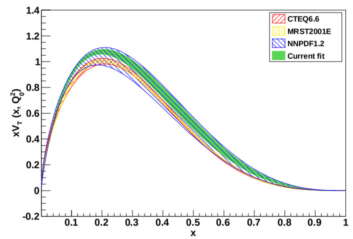
	REPLICAS	CENTRAL VALUE	FIXED PARTITION
χ^2	1.32	1.32	$\sim$1.3
$\langle\chi^2\rangle_{\rm rep}$	2.79 ± 0.24	1.65 ± 0.20	$\sim 1.6 \pm 0.2$
$\langle\sigma^{\rm dat}\rangle$	0.039	0.035	$\sim$0.03

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

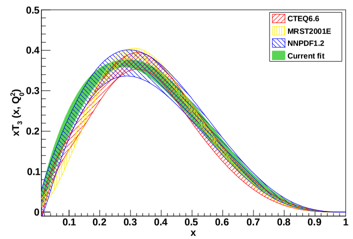
GLUE



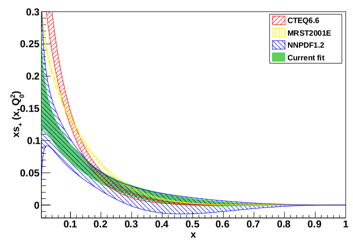
VALENCE



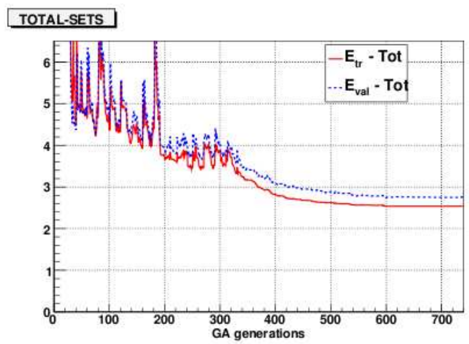
TRIPLET

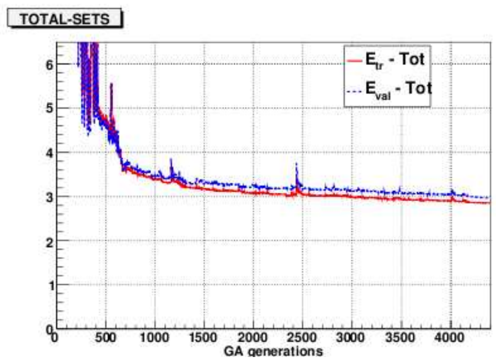


STRANGE



- 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.