DISTRIBUZIONI PARTONICHE,
QUAL È IL PROBLEMA?

STEFANO FORTE
UNIVERSITÀ DI MILANO

ROMA, 22 FEBBRAIO 2007

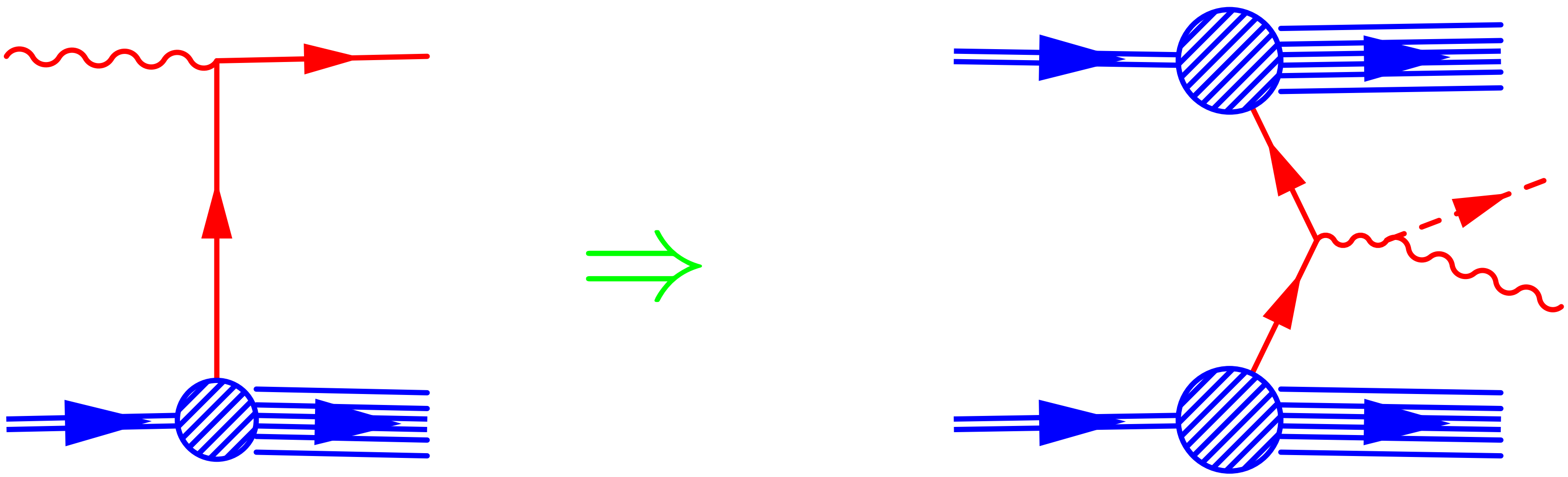
PARTONS FOR LHC:

The accurate computation of physical process at a hadron collider requires good knowledge of parton distributions of the nucleon

PARTONS FOR LHC:

The accurate computation of physical process at a hadron collider requires good knowledge of parton distributions of the nucleon

FACTORIZATION

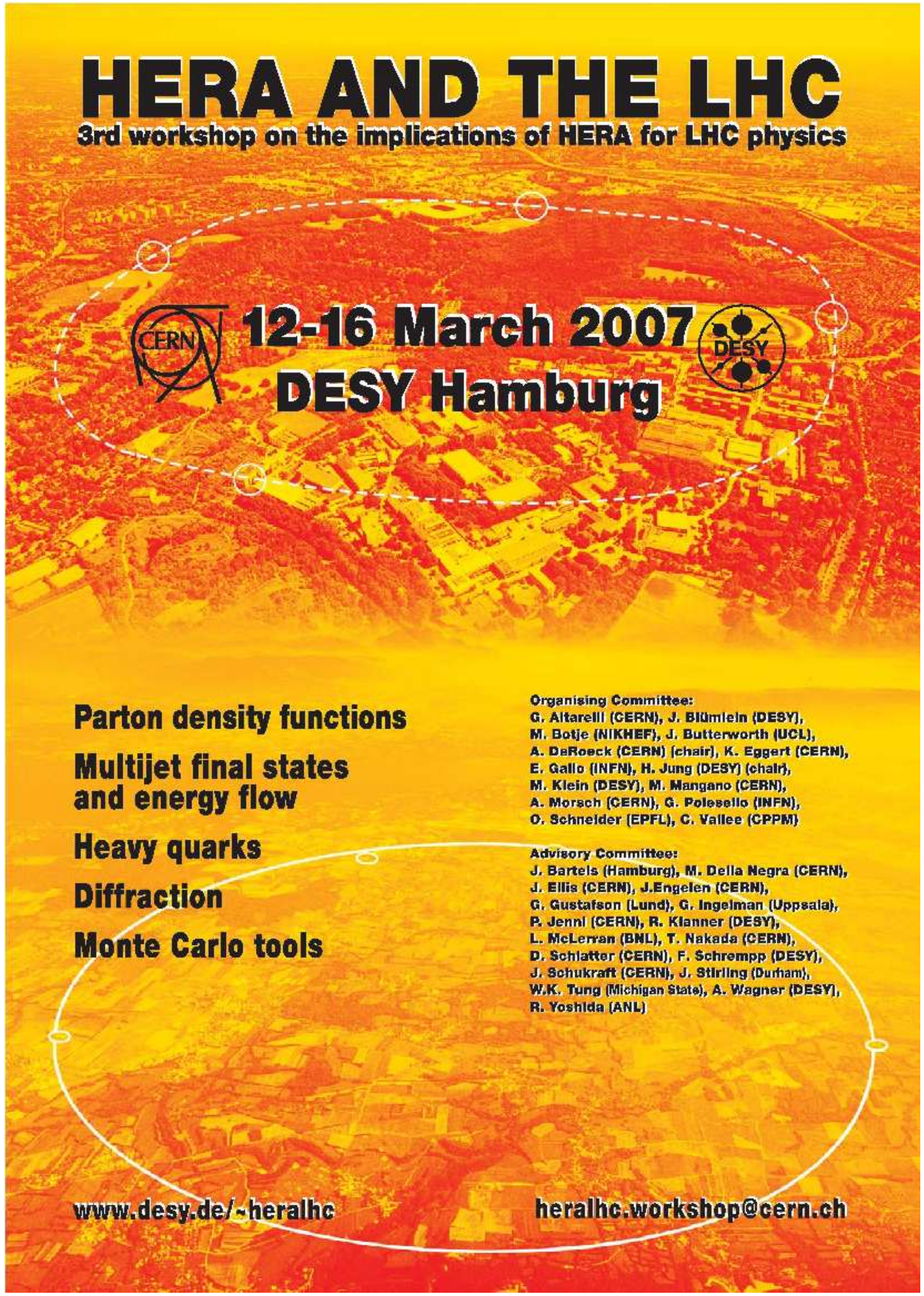


in order to extract the relevant physics signal,

we need to know the parton distributions and their uncertainty

IS THIS ASPECT OF LHC PHYSICS UNDER CONTROL?

AN ONGOING EFFORT...

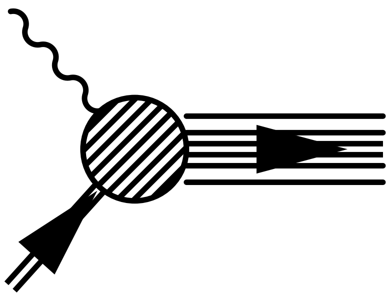


SUMMARY

- PARTON DISTRIBUTIONS: THE STATE OF THE ART
  - PARTON FITS IN THE ERA OF LHC
  - PARTON DISTRIBUTIONS WITH ERRORS: CAN WE TRUST THEM?
  - THE NEED FOR NEW IDEAS
- THE NEURAL NETWORK APPROACH TO PDFS
  - THE NEURAL MONTE CARLO
  - NEURAL NETWORKS: OVERLEARNING AND STOPPING
  - NEURAL PARTONS AND PARAMETRIZATION BIAS

DEEP-INELASTIC SCATTERING

Structure Functions. . .



Lepton fractional energy loss: $y = \frac{p\cdot q}{p\cdot k}$;

Bjorken x: $x = \frac{Q^2}{2p\cdot q}$

lepton-nucleon CM energy: $s = \frac{Q^2}{xy}$;

virtual boson-nucleon CM energy $W^2 = Q^2 \frac{1-x}{x}$;

$$\frac{d^2\sigma^{\lambda_p\lambda_\ell}(x,y,Q^2)}{dxdy} = \frac{G_F^2}{2\pi(1+Q^2/m_W^2)^2}\frac{Q^2}{xy}\left\{\left[-\lambda_\ell\, y\left(1-\frac{y}{2}\right)xF_3(x,Q^2) + (1-y)F_2(x,Q^2)\right.\right.$$

$$\left.\left. +y^2 xF_1(x,Q^2)\right] - 2\lambda_p\left[-\lambda_\ell\, y(2-y)xg_1(x,Q^2) - (1-y)g_4(x,Q^2) - y^2 xg_5(x,Q^2)\right]\right\}$$

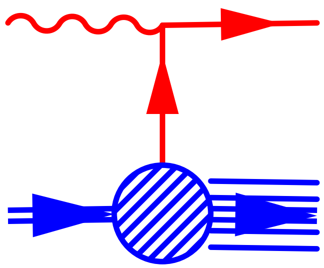
$\lambda_l \;\to$ lepton helicity

$\lambda_p \;\to$ proton helicity

	PARITY CONS.	PARITY VIOL.
UNPOL.	F_1, F_2	F_3
POL.	g_1	g_4, g_5

...AND PARTON DISTRIBUTIONS

STRUCTURE FUNCTION=HARD COEFF.$\otimes$PARTON DISTN.



$$F_2^{\rm NC}(x,Q^2) = x \sum_{\text{flav. } i} e_i^2 (q_i + \bar{q}_i) + \alpha_s \left[C_i[\alpha_s] \otimes (q_i + \bar{q}_i) + C_g[\alpha_s] \otimes g \right]$$

q_i quark, $\bar{q}_i$ antiquark, g gluon

LEADING PARTON CONTENT (up to $O[\alpha_s]$ corrections)

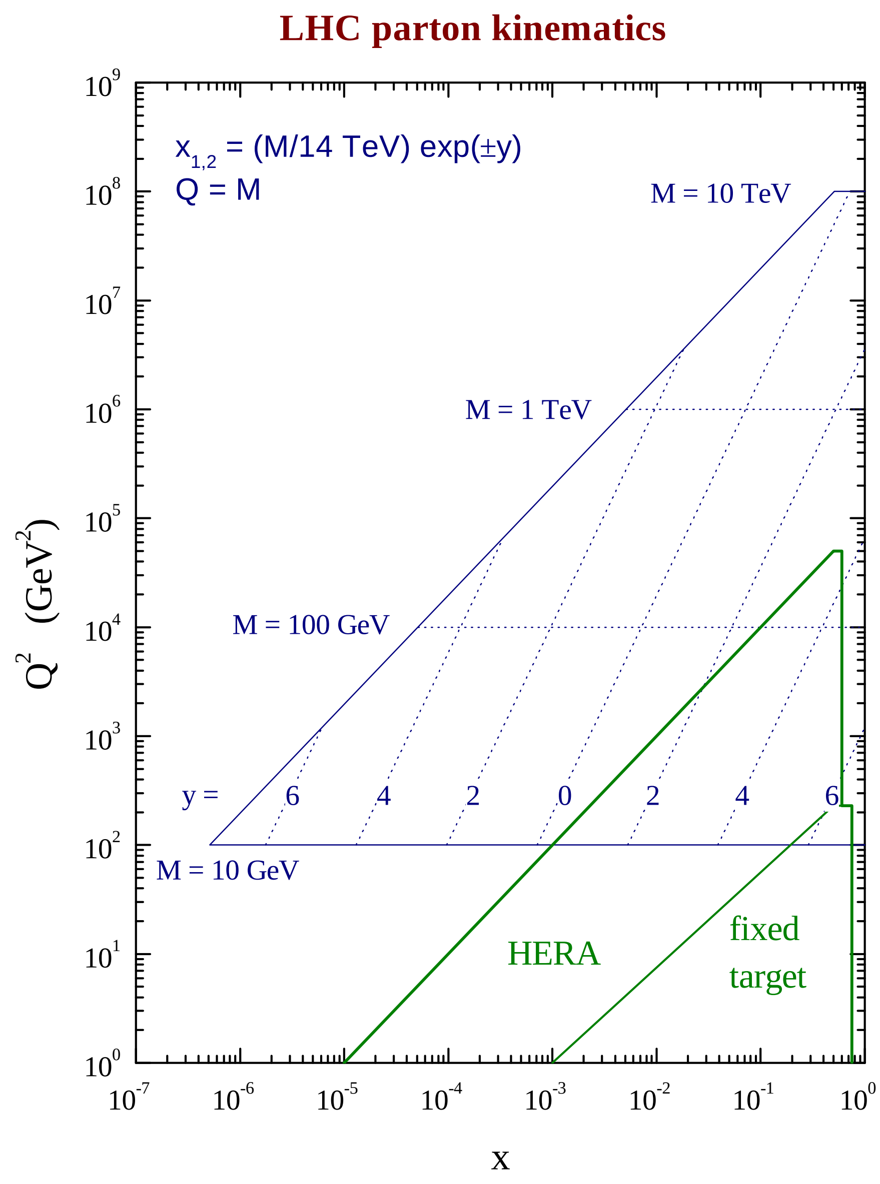
	$q_i \equiv q_i^{\uparrow\uparrow} + q_i^{\uparrow\downarrow}$	$\Delta q_i \equiv q_i^{\uparrow\uparrow} - q_i^{\uparrow\downarrow}$
NC	$F_1{}^{\gamma,\,Z} = \sum_i e_i^2\,(q_i + \bar{q}_i)$	$g_1{}^{\gamma,\,Z} = \sum_i e_i^2\,(\Delta q_i + \Delta \bar{q}_i)$
CC	$F_1^{W^+} = \bar{u} + d + s + \bar{c}$	$g_1^{W^+} = \Delta\bar{u} + \Delta d + \Delta s + \Delta\bar{c}$
CC	$-F_3^{W^+}/2 = \bar{u} - d - s + \bar{c}$	$g_5^{W^+} = \Delta\bar{u} - \Delta d - \Delta s + \Delta\bar{c}$
	$F_2 = 2xF_1$	$g_4 = 2xg_5$

$W^+ \to W^- \Rightarrow u \leftrightarrow d,\ c \leftrightarrow s$; more combinations using Isospin: $p \to n \Rightarrow u \leftrightarrow d$

FROM HERA TO LHC

AT A HADRON COLLIDER

- SCALE Q DETERMINED BY MASS OF FINAL STATE
- MOMENTUM FRACTIONS x_1 x_2 DETERMINED BY MASS & RAPIDITY OF FINAL STATE



FROM HERA TO LHC ⇒ EVOLUTION

AT A HADRON COLLIDER

- SCALE Q DETERMINED BY MASS OF FINAL STATE
- MOMENTUM FRACTIONS x_1 x_2 DETERMINED BY MASS & RAPIDITY OF FINAL STATE

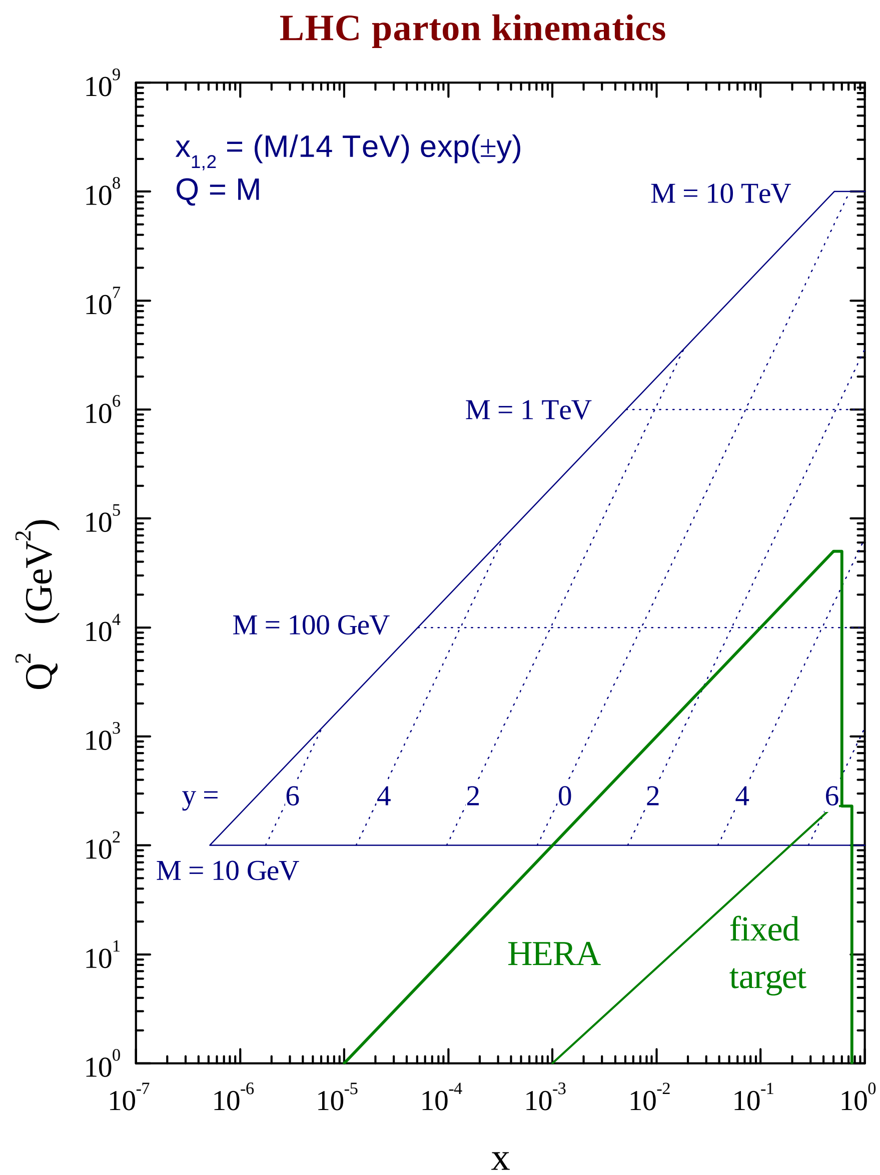
USE ALTARELLI-PARISI EQNS:

$$\frac{d}{dt}\Delta q_{NS} = \frac{\alpha_s(t)}{2\pi} P_{qq}^{NS} \otimes \Delta q_{NS},$$

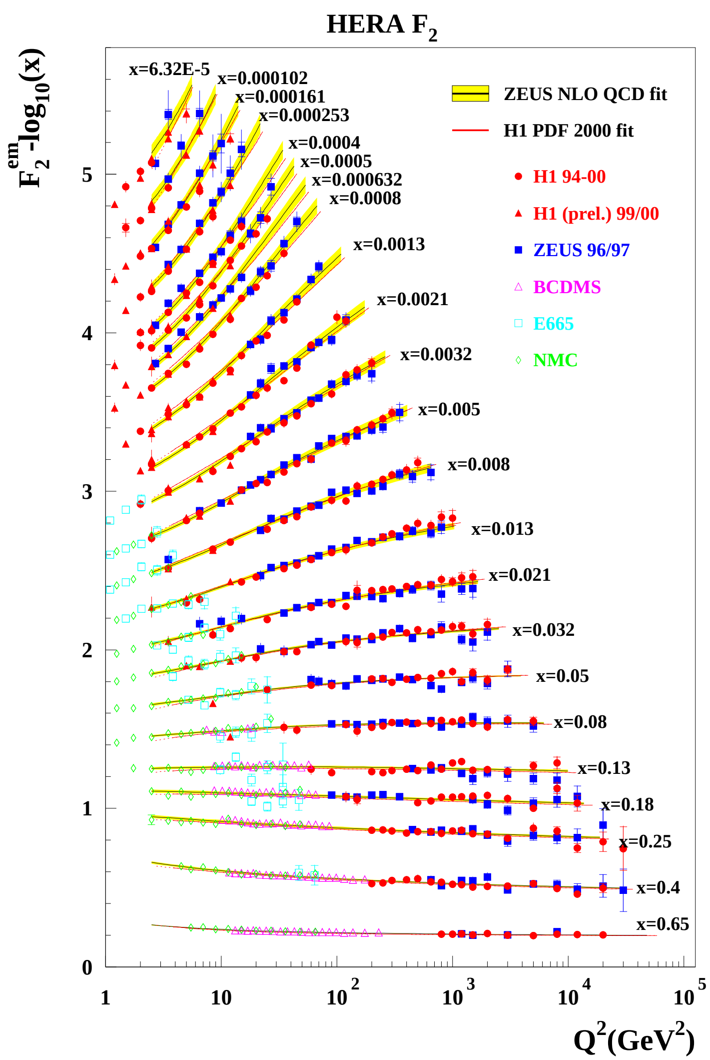
$$\frac{d}{dt}\begin{pmatrix}\Delta\Sigma \\ \Delta g\end{pmatrix} = \frac{\alpha_s(t)}{2\pi}\begin{pmatrix}P_{qq}^S & 2n_f P_{qg}^S \\ P_{gq}^S & P_{gg}^S\end{pmatrix} \otimes \begin{pmatrix}\Delta\Sigma \\ \Delta g\end{pmatrix}$$

TO EVOLVE PARTONS

FROM DIS TO LHC KINEMATICS



PARTONS WITH ERRORS

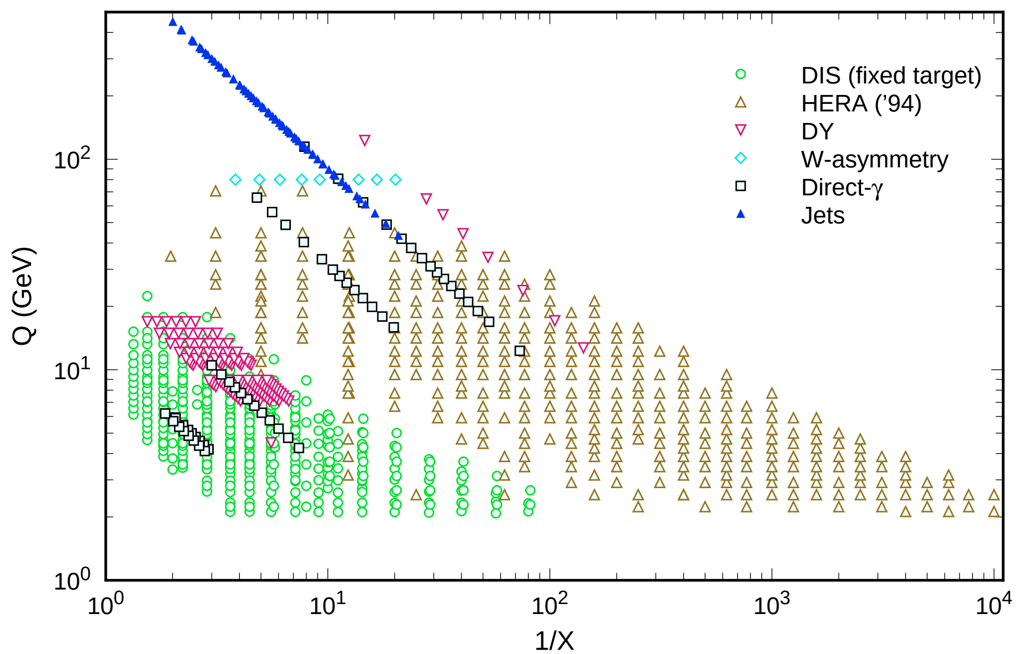


GIVEN A SET OF DATA POINTS

MUST DETERMINE A SET OF

FUNCTIONS WITH ERRORS

DATA INCLUDED IN CTEQ5 PARTON FIT



WHAT'S THE PROBLEM? D. Kosower, 1999

- FOR A SINGLE QUANTITY, WE QUOTE 1 SIGMA ERRORS: VALUE$\pm$ ERROR
- FOR A PAIR OF NUMBERS, WE QUOTE A 1 SIGMA ELLIPSE
- FOR A FUNCTION, WE NEED AN "ERROR BAR" IN A SPACE OF FUNCTIONS

MUST DETERMINE THE PROBABILITY DENSITY (MEASURE) $\mathcal{P}[f_i(x)]$

IN THE SPACE OF PARTON DISTRIBUTION FUNCTIONS $f_i(x)$ (i=quark, antiquark, gluon)

WHAT'S THE PROBLEM? D. Kosower, 1999

- FOR A SINGLE QUANTITY, WE QUOTE 1 SIGMA ERRORS: VALUE$\pm$ ERROR
- FOR A PAIR OF NUMBERS, WE QUOTE A 1 SIGMA ELLIPSE
- FOR A FUNCTION, WE NEED AN "ERROR BAR" IN A SPACE OF FUNCTIONS

MUST DETERMINE THE PROBABILITY DENSITY (MEASURE) $\mathcal{P}[f_i(x)]$

IN THE SPACE OF PARTON DISTRIBUTION FUNCTIONS $f_i(x)$ (i=quark, antiquark, gluon)

EXPECTATION VALUE OF $\sigma\left[f_i(x)\right] \Rightarrow$ FUNCTIONAL INTEGRAL

$$\left\langle \sigma\left[f_i(x)\right] \right\rangle = \int \mathcal{D}f_i\, \sigma\left[f_i(x)\right]\, \mathcal{P}[f_i],$$

WHAT'S THE PROBLEM? D. Kosower, 1999

- FOR A SINGLE QUANTITY, WE QUOTE 1 SIGMA ERRORS: VALUE$\pm$ ERROR
- FOR A PAIR OF NUMBERS, WE QUOTE A 1 SIGMA ELLIPSE
- FOR A FUNCTION, WE NEED AN "ERROR BAR" IN A SPACE OF FUNCTIONS

MUST DETERMINE THE PROBABILITY DENSITY (MEASURE) $\mathcal{P}[f_i(x)]$

IN THE SPACE OF PARTON DISTRIBUTION FUNCTIONS $f_i(x)$ (i=quark, antiquark, gluon)

EXPECTATION VALUE OF $\sigma\left[f_i(x)\right] \Rightarrow$ FUNCTIONAL INTEGRAL

$$\left\langle \sigma\left[f_i(x)\right] \right\rangle = \int \mathcal{D}f_i \, \sigma\left[f_i(x)\right] \mathcal{P}[f_i],$$

MUST DETERMINE AN INFINITE–DIMENSIONAL OBJECT FROM A FINITE SET OF DATA POINTS

THE STANDARD SOLUTION:

FUNCTIONAL PARTON FITTING

- CHOOSE A FIXED FUNCTIONAL FORM:
  - MRST: 24 PARMS., SOME FIXED $\longrightarrow$ 15 PARMS.

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

$$xg(x,Q_0^2) = A_g(1-x)^{\eta_g}(1+\epsilon_g x^{0.5}+\gamma_g x)x^{\delta_g} - A_-(1-x)^{\eta_-}x^{-\delta_-},$$

  - CTEQ: 20 PARMS.

$$x\, f(x,Q_0) = A_0\, x^{A_1}\,(1-x)^{A_2}\, e^{A_3 x}\,(1+e^{A_4}\, x)^{A_5}$$

with independent params for combinations $u_v \equiv u-\bar{u}$, $d_v \equiv d-\bar{d}$, g, and $\bar{u}+\bar{d}$, $s=\bar{s}=0.2\,(\bar{u}+\bar{d})$ at Q_0; NORM. FIXED BY SUM RULES

  - ALEKHIN: 17 PARMS.

$$xu_{\rm V}(x,Q_0) = \frac{2}{N_{\rm u}^{\rm V}}x^{a_{\rm u}}(1-x)^{b_{\rm u}}(1+\gamma_2^{\rm u}x); \quad xu_{\rm S}(x,Q_0) = \frac{A_{\rm S}}{N_{\rm S}}\eta_{\rm u}x^{a_{\rm s}}(1-x)^{b_{\rm su}}$$

$$xd_{\rm V}(x,Q_0) = \frac{1}{N_{\rm d}^{\rm V}}x^{a_{\rm d}}(1-x)^{b_{\rm d}}; \quad xd_{\rm S}(x,Q_0) = \frac{A_{\rm S}}{N^{\rm S}}x^{a_{\rm s}}(1-x)^{b_{\rm sd}},$$

$$xs_{\rm S}(x,Q_0) = \frac{A_{\rm S}}{N^{\rm S}}\eta_{\rm s}x^{a_{\rm s}}(1-x)^{(b_{\rm su}+b_{\rm sd})/2}; \quad xG(x,Q_0) = A_{\rm G}x^{a_{\rm G}}(1-x)^{b_{\rm G}}(1+\gamma_1^{\rm G}\sqrt{x}+\gamma_2^{\rm G}x),$$

THE STANDARD SOLUTION:

FUNCTIONAL PARTON FITTING

- CHOOSE A FIXED FUNCTIONAL FORM:
  - MRST: 24 PARMS., SOME FIXED $\rightarrow$ 15 PARMS.

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

$$xg(x,Q_0^2) = A_g(1-x)^{\eta_g}(1+\epsilon_g x^{0.5}+\gamma_g x)x^{\delta_g} - A_-(1-x)^{\eta_-}x^{-\delta_-},$$

  - CTEQ: 20 PARMS.

$$x\, f(x,Q_0) = A_0\, x^{A_1}\,(1-x)^{A_2}\, e^{A_3 x}\,(1+e^{A_4}\, x)^{A_5}$$

with independent params for combinations $u_v \equiv u-\bar{u}$, $d_v \equiv d-\bar{d}$, g, and $\bar{u}+\bar{d}$, $s=\bar{s}=0.2\,(\bar{u}+\bar{d})$ at Q_0; NORM. FIXED BY SUM RULES

  - ALEKHIN: 17 PARMS.

$$xu_{\rm V}(x,Q_0) = \frac{2}{N_{\rm u}^{\rm V}}x^{a_{\rm u}}(1-x)^{b_{\rm u}}(1+\gamma_2^{\rm u}x); \quad xu_{\rm S}(x,Q_0) = \frac{A_{\rm S}}{N_{\rm S}}\eta_{\rm u}x^{a_{\rm s}}(1-x)^{b_{\rm su}}$$

$$xd_{\rm V}(x,Q_0) = \frac{1}{N_{\rm d}^{\rm V}}x^{a_{\rm d}}(1-x)^{b_{\rm d}}; \quad xd_{\rm S}(x,Q_0) = \frac{A_{\rm S}}{N^{\rm S}}x^{a_{\rm s}}(1-x)^{b_{\rm sd}},$$

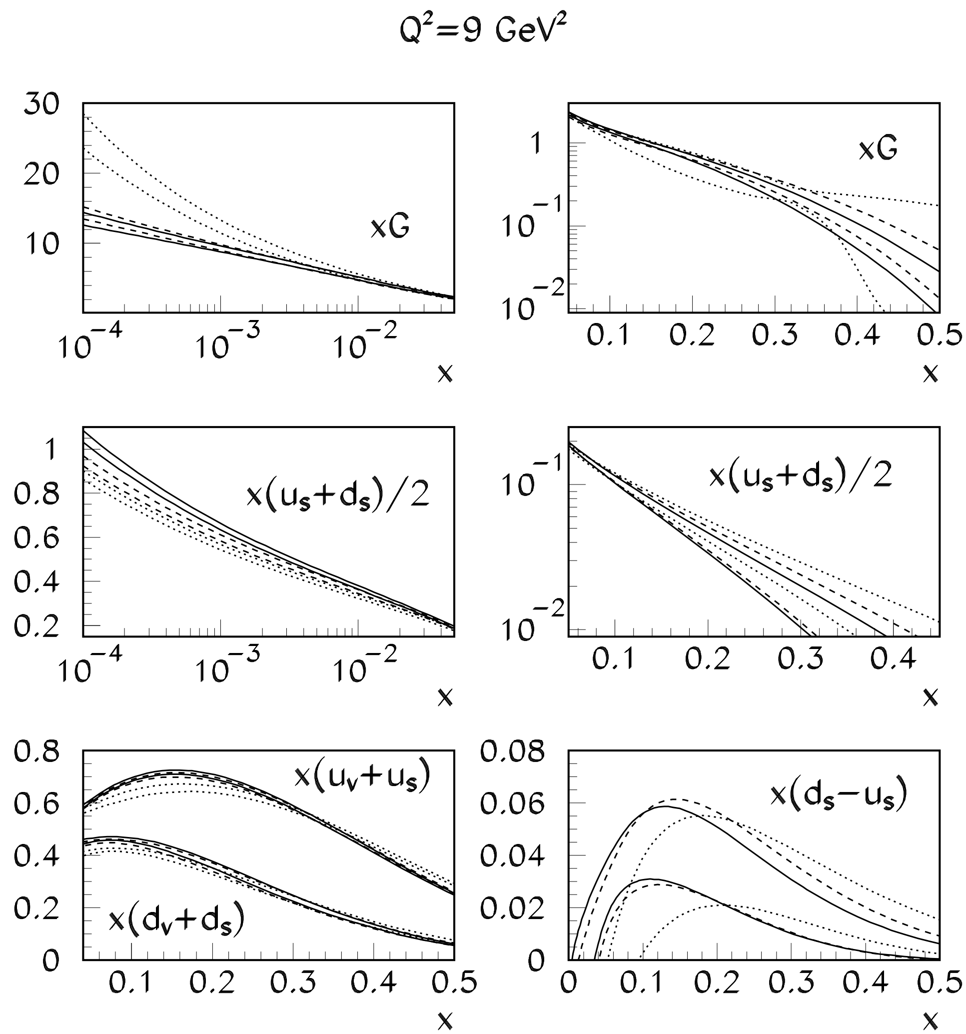
$$xs_{\rm S}(x,Q_0) = \frac{A_{\rm S}}{N^{\rm S}}\eta_{\rm s}x^{a_{\rm s}}(1-x)^{(b_{\rm su}+b_{\rm sd})/2}; \quad xG(x,Q_0) = A_{\rm G}x^{a_{\rm G}}(1-x)^{b_{\rm G}}(1+\gamma_1^{\rm G}\sqrt{x}+\gamma_2^{\rm G}x),$$

- EVOLVE TO DESIRED SCALE & COMPUTE PHYSICAL OBSERVABLES
- DETERMINE BEST-FIT VALUES OF PARAMETERS
- DETERMINE ERROR BY PROPAGATION OF ERROR ON PARMS ('HESSIAN METHOD') OR BY PARM. SCANS ('LAGRANGE MULTIPLIER METHOD')

PROBLEM PROJECTED ONTO THE FINITE–DIMENSIONAL SPACE OF PARAMETERS

HOW WELL DOES IT WORK? (DIS ONLY)

ALEKHIN 2003 PARTONS



TOTAL ERROR BANDS FOR LO (DOTS), NLO (DASHES), NNLO (SOLID) PARTON DISTRIBUTIONS

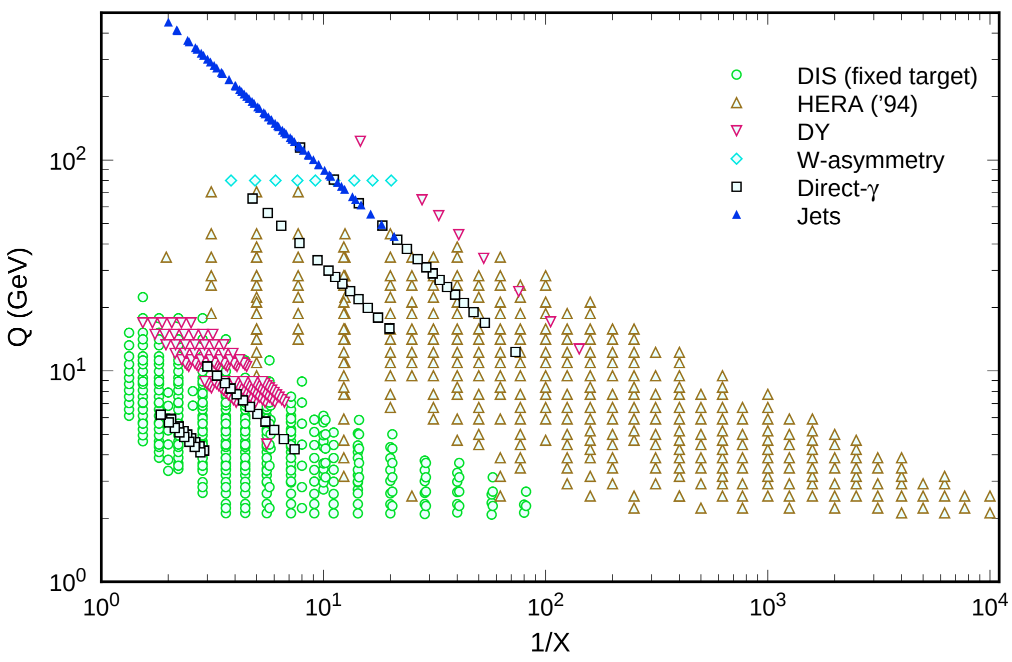
valence $u^v \equiv u - \bar{u}$, $d^v \equiv d - \bar{d}$,
sea $u^s = \bar{u}^s = d^s = \bar{d}^s$

HADRONIC CHANNELS

IN UNPOLARIZED GLOBAL FITS

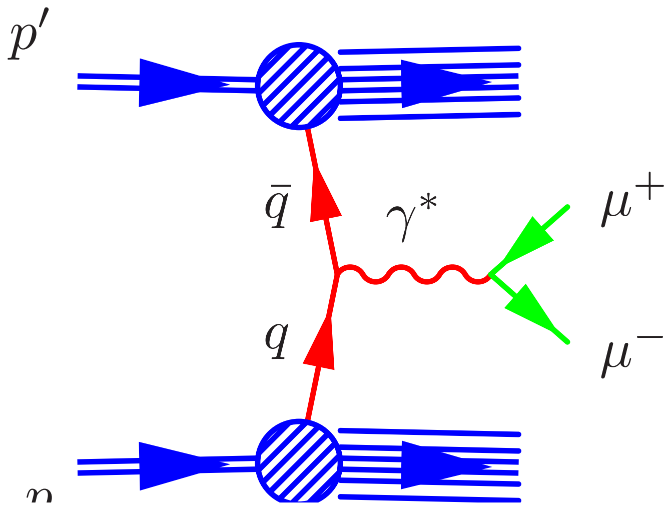
- DRELL-YAN $\Rightarrow$ $\bar{u}/\bar{d}$ ASYMMETRY
- $W^{\pm}$ $\Rightarrow$ u/d ASYMMETRY
- DIRECT γ $\Rightarrow$ GLUON (impact negligible w.r. to DIS)
- LARGE E_T JETS $\Rightarrow$ LARGE x GLUON

DATA IN CTEQ5 PARTON FIT



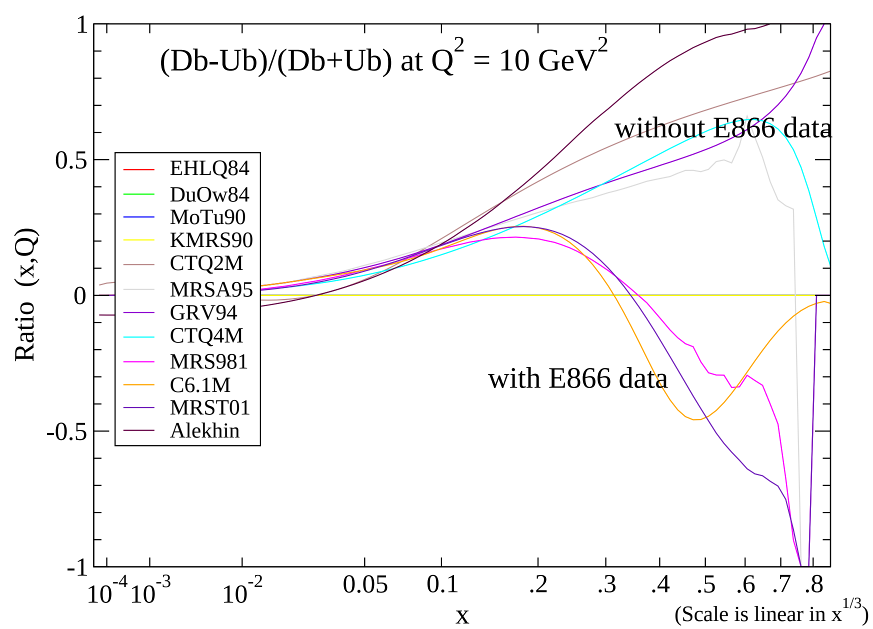
HOW WELL DOES IT WORK? (WITH HADRONIC DATA)

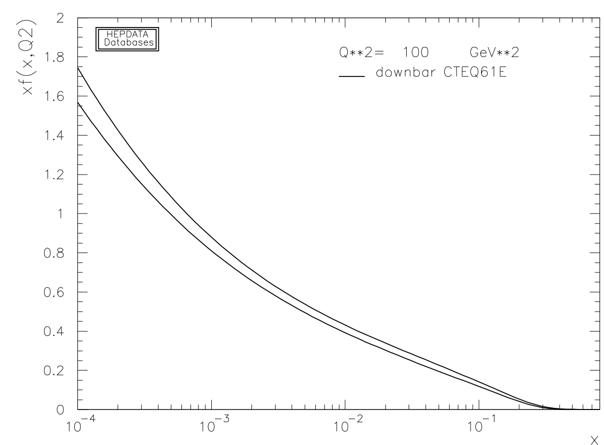
DRELL-YAN p/d ASYMMETRY

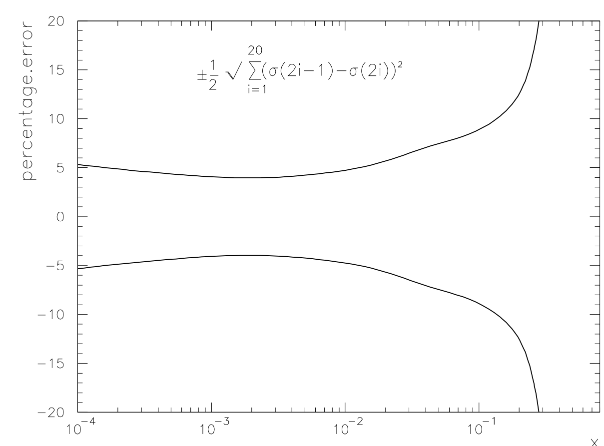


$$\left.\frac{\sigma^{pd}}{\sigma^{pp}}\right|_{x_1>>x_2} \approx \frac{1}{2}\left(1+\frac{\bar{d}(x_2)}{\bar{u}(x_2)}\right)$$

$5-10\%$ ACCURACY ON $\bar{u}$, $\bar{d}$







CAN WE TRUST GLOBAL FITS?

PARTON SETS DO NOT AGREE WITHIN RESPECTIVE ERRORS!

W PRODUCTION CROSS-SECTION
TEVATRON

PDF SET	XSEC [NB]	PDF UNCERTAINTY
ALEKHIN	2.73	$\pm$ 0.05 (TOT)
MRST2002	2.59	$\pm$ 0.03 (EXPT)
CTEQ6	2.54	$\pm$ 0.10 (EXPT)

THORNE 2003

- ALEKHIN VS. MRST/CTEQ $\rightarrow$ W PRODUCTION XSECT AT TEVATRON DO NOT AGREE WITHIN RESPECTIVE ERRORS

CAN WE TRUST GLOBAL FITS?

PARTON SETS DO NOT AGREE WITHIN RESPECTIVE ERRORS!

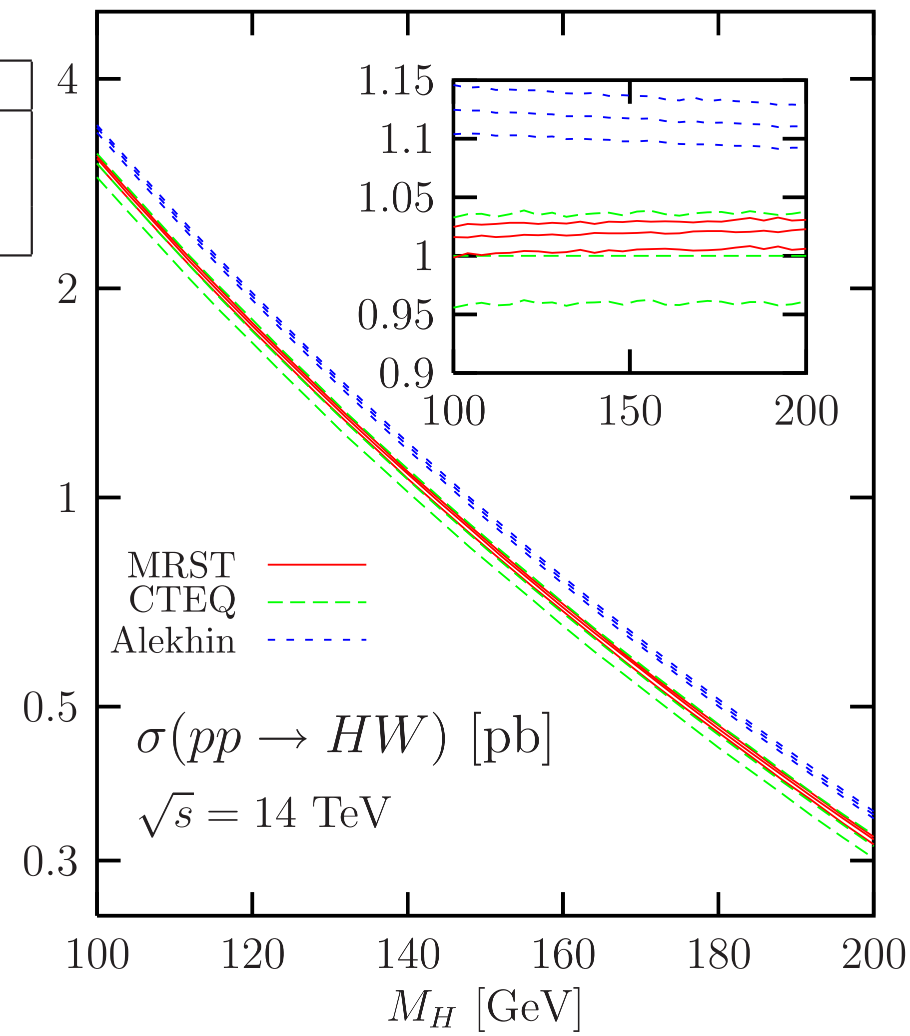
W PRODUCTION CROSS-SECTION TEVATRON

PDF SET	XSEC [NB]	PDF UNCERTAINTY
ALEKHIN	2.73	$\pm$ 0.05 (TOT)
MRST2002	2.59	$\pm$ 0.03 (EXPT)
CTEQ6	2.54	$\pm$ 0.10 (EXPT)

THORNE 2003

- ALEKHIN VS. MRST/CTEQ $\rightarrow$ W PRODUCTION XSECT AT TEVATRON DO NOT AGREE WITHIN RESPECTIVE ERRORS
- ALEKHIN VS. MRST/CTEQ $\rightarrow$ PREDICTIONS FOR ASSOCIATE HIGGS W PRODUCTION LHC DO NOT AGREE WITHIN RESPECTIVE ERRORS

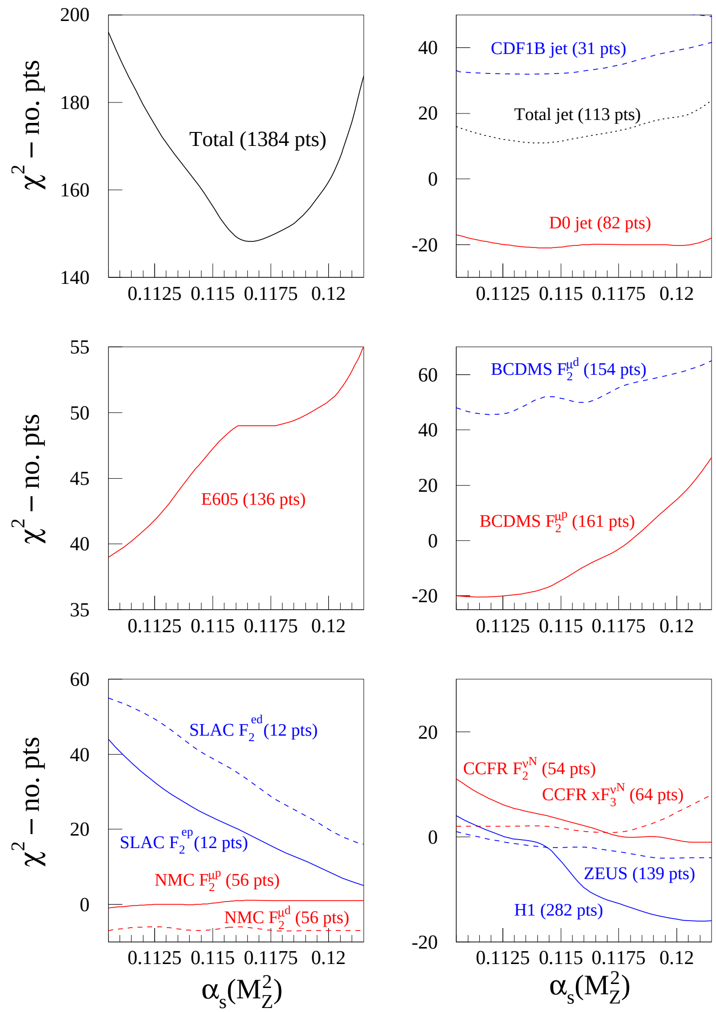
HIGGS PRODUCTION AT LHC



DJOUADI AND FERRAG, 2004

INCOMPATIBLE DATA?

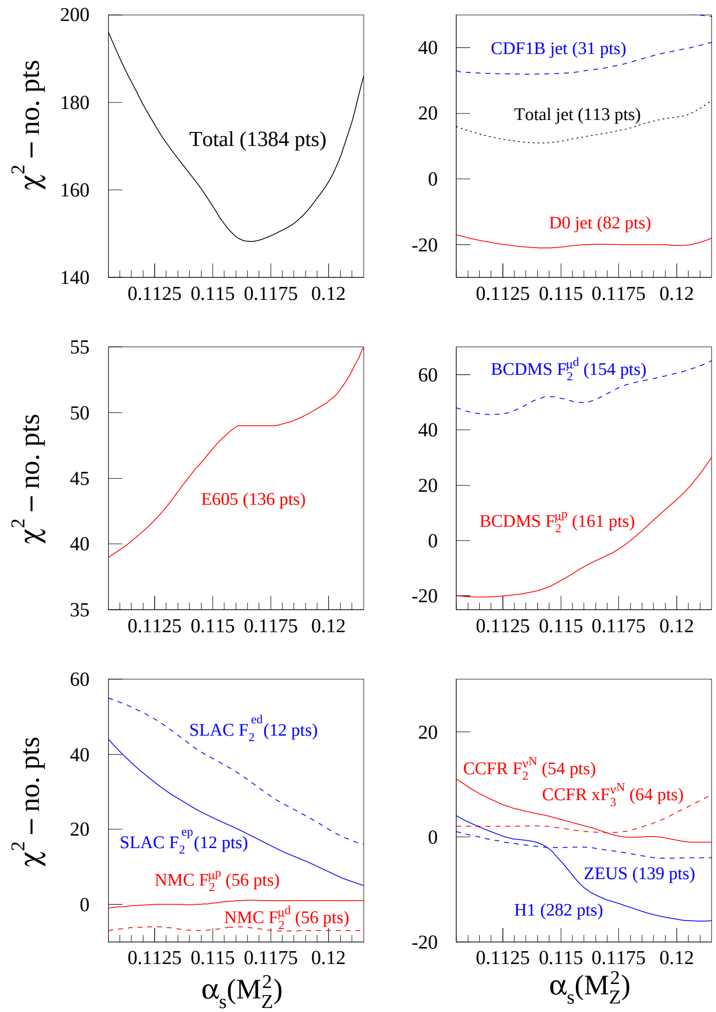
GLOBAL χ^2 MINIMUM MAY NOT CORRESPOND TO LOCAL MINIMA



MRST 2003

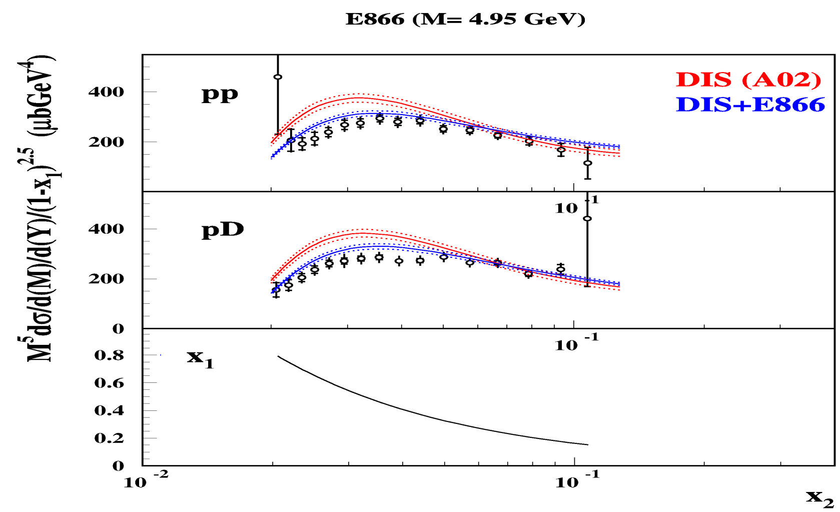
INCOMPATIBLE DATA?

GLOBAL χ^2 MINIMUM MAY NOT CORRESPOND TO LOCAL MINIMA



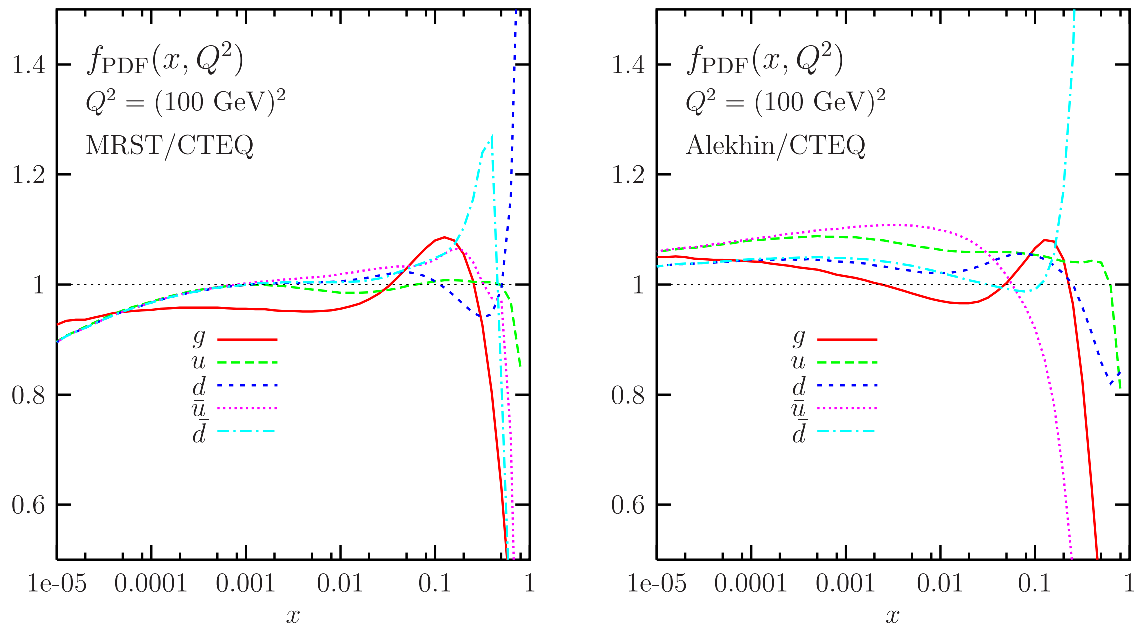
MRST 2003

E866 DY DATA DISAGREE WITH DIS DATA:
$\sigma_{DY} \sim q(x_1)q(x_2)$ DISAGREES WITH DIS QUARK AT SAME x AND Q^2



ALEKHIN 2005

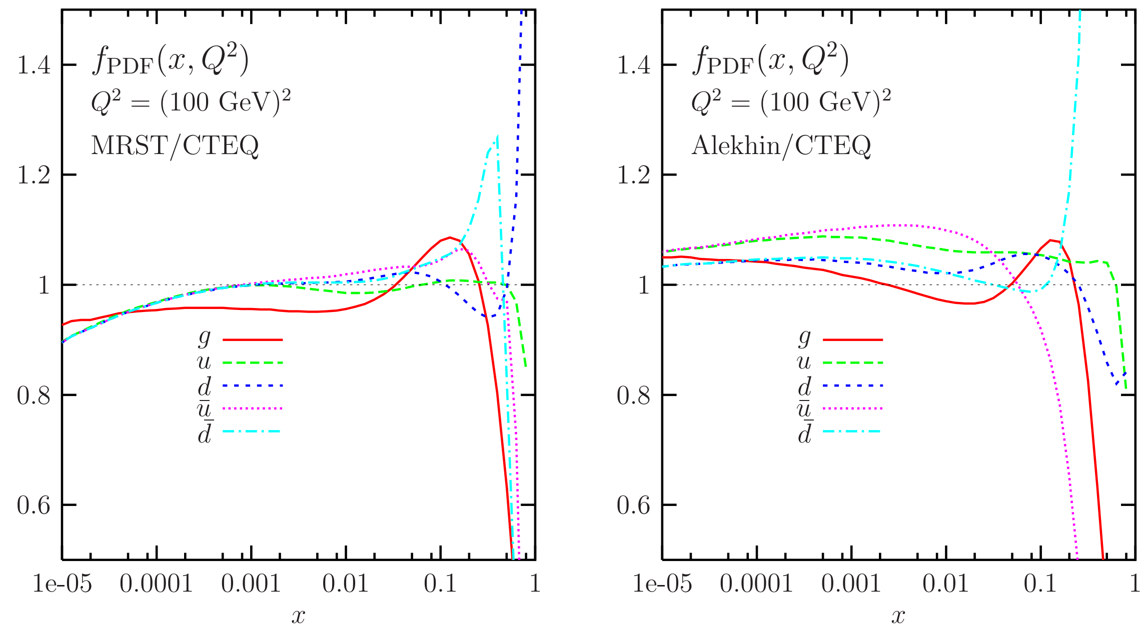
PARAMETRIZATION BIAS?



MRST & CTEQ
$\longrightarrow$ SIMILAR PARTONS

Djouadi and Ferrag 2003

PARAMETRIZATION BIAS?



MRST & CTEQ
$\longrightarrow$ SIMILAR PARTONS

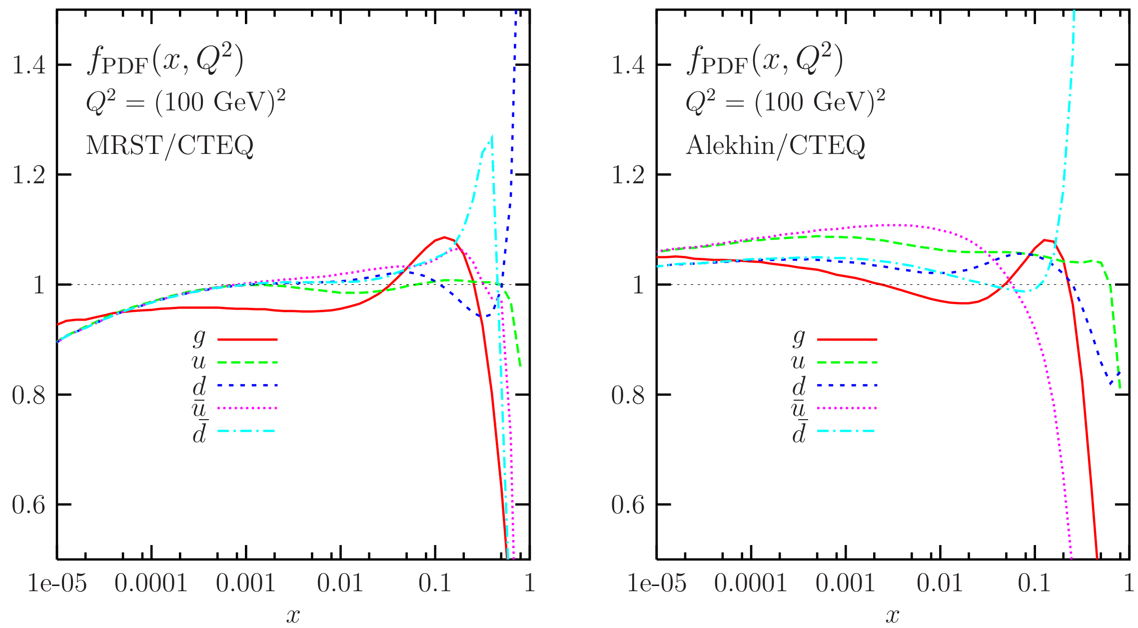
Djouadi and Ferrag 2003

SIMILAR PARTONS
$\longrightarrow$ SIMILAR RESULTS

THE W XSECT. AGAIN...

PDF SET	COMMENT	XSEC [NB]	PDF UNCERTAINTY
ALEKHIN	TEVATRON	2.73	$\pm$ 0.05 (TOT)
MRST2002	TEVATRON	2.59	$\pm$ 0.03 (EXPT)
CTEQ6	TEVATRON	2.54	$\pm$ 0.10 (EXPT)
ALEKHIN	LHC	215	$\pm$ 6 (TOT)
MRST2002	LHC	204	$\pm$ 4 (EXPT)
CTEQ6	LHC	205	$\pm$ 8 (EXPT)

PARAMETRIZATION BIAS?



MRST & CTEQ
$\longrightarrow$ SIMILAR PARTONS

Djouadi and Ferrag 2003

SIMILAR PARTONS
$\longrightarrow$ SIMILAR RESULTS

THE W XSECT. AGAIN. . .

PDF SET	COMMENT	XSEC [NB]	PDF UNCERTAINTY
ALEKHIN	TEVATRON	2.73	$\pm$ 0.05 (TOT)
MRST2002	TEVATRON	2.59	$\pm$ 0.03 (EXPT)
CTEQ6	TEVATRON	2.54	$\pm$ 0.10 (EXPT)
ALEKHIN	LHC	215	$\pm$ 6 (TOT)
MRST2002	LHC	204	$\pm$ 4 (EXPT)
CTEQ6	LHC	205	$\pm$ 8 (EXPT)

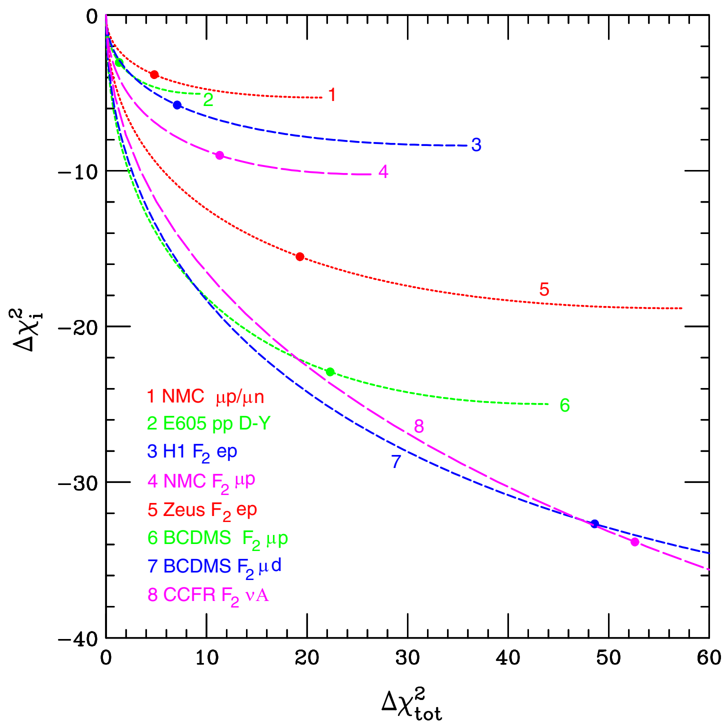
We do not seem to have the optimum parameterization for both finding the best fit and also investigating fluctuations about this best fit (...) This might then influence our error analysis...(MRST 2004)

SOLUTIONS: CTEQ TOLERANCE CRITERION

SINGLE OUT INCONSISTENT DATA

- how many parameters are significantly determined by each dataset?
- how consistent are the data from one set with the rest?

STUDY MINIMUM ALLOWED χ_i^2
FOR i-TH EXP. AS
GLOBAL χ^2 ALLOWED TO INCREASE



Collins, Pumplin 2001

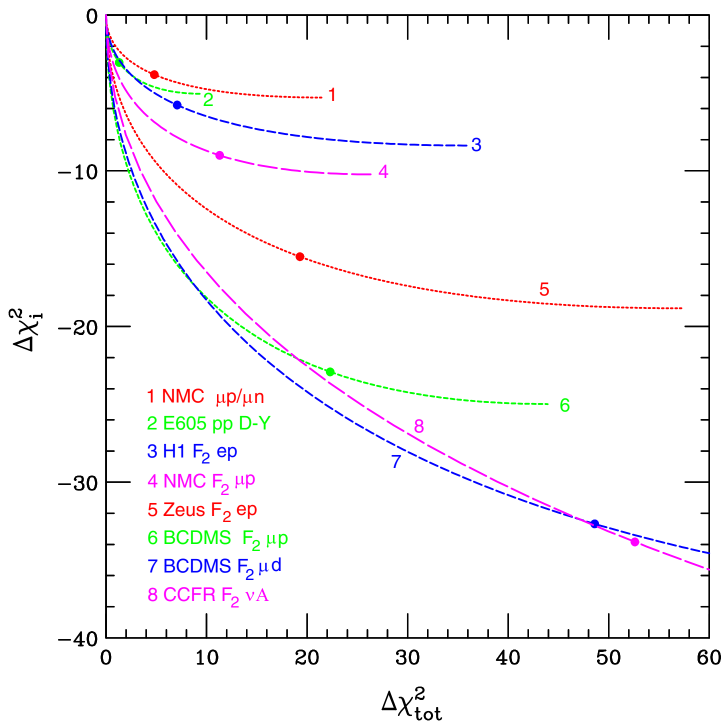
CCFR, BCDMS
INCOMPATIBLE WITH THE REST

SOLUTIONS: CTEQ TOLERANCE CRITERION

SINGLE OUT INCONSISTENT DATA

- how many parameters are significantly determined by each dataset?
- how consistent are the data from one set with the rest?

STUDY MINIMUM ALLOWED χ_i^2
FOR i-TH EXP. AS
GLOBAL χ^2 ALLOWED TO INCREASE

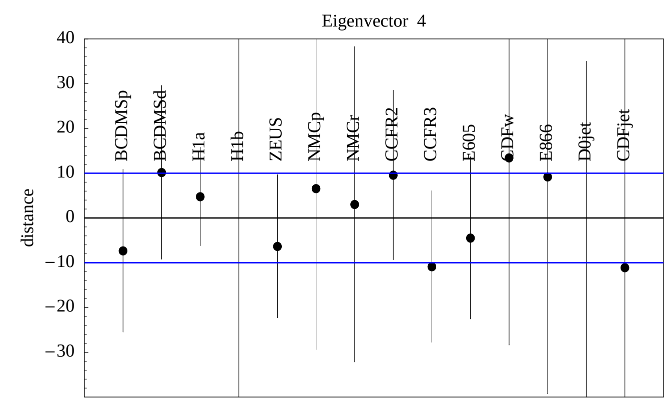


Collins, Pumplin 2001

CCFR, BCDMS
INCOMPATIBLE WITH THE REST

OPTIONS

- discard incompatible experiments
- reweight individual contributions
- INCORPORATE IN ERROR, TOLERATING FIXED MAX DEVIATION FOR EACH EXPERIMENT & EACH FIT PARAMETER

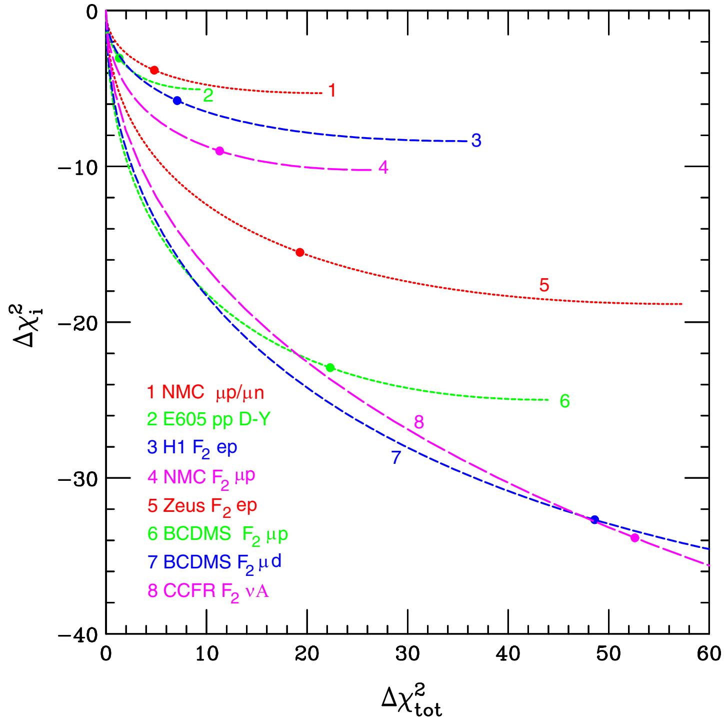


SOLUTIONS: CTEQ TOLERANCE CRITERION

SINGLE OUT INCONSISTENT DATA

- how many parameters are significantly determined by each dataset?
- how consistent are the data from one set with the rest?

STUDY MINIMUM ALLOWED χ_i^2
FOR i-TH EXP. AS
GLOBAL χ^2 ALLOWED TO INCREASE

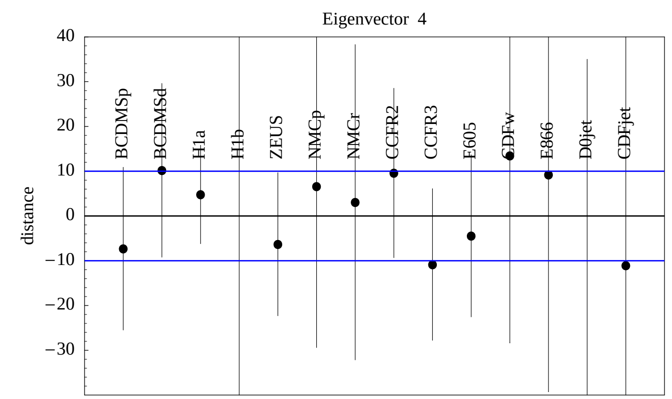


Collins, Pumplin 2001

CCFR, BCDMS
INCOMPATIBLE WITH THE REST

OPTIONS

- discard incompatible experiments
- reweight individual contributions
- INCORPORATE IN ERROR, TOLERATING FIXED MAX DEVIATION FOR EACH EXPERIMENT & EACH FIT PARAMETER

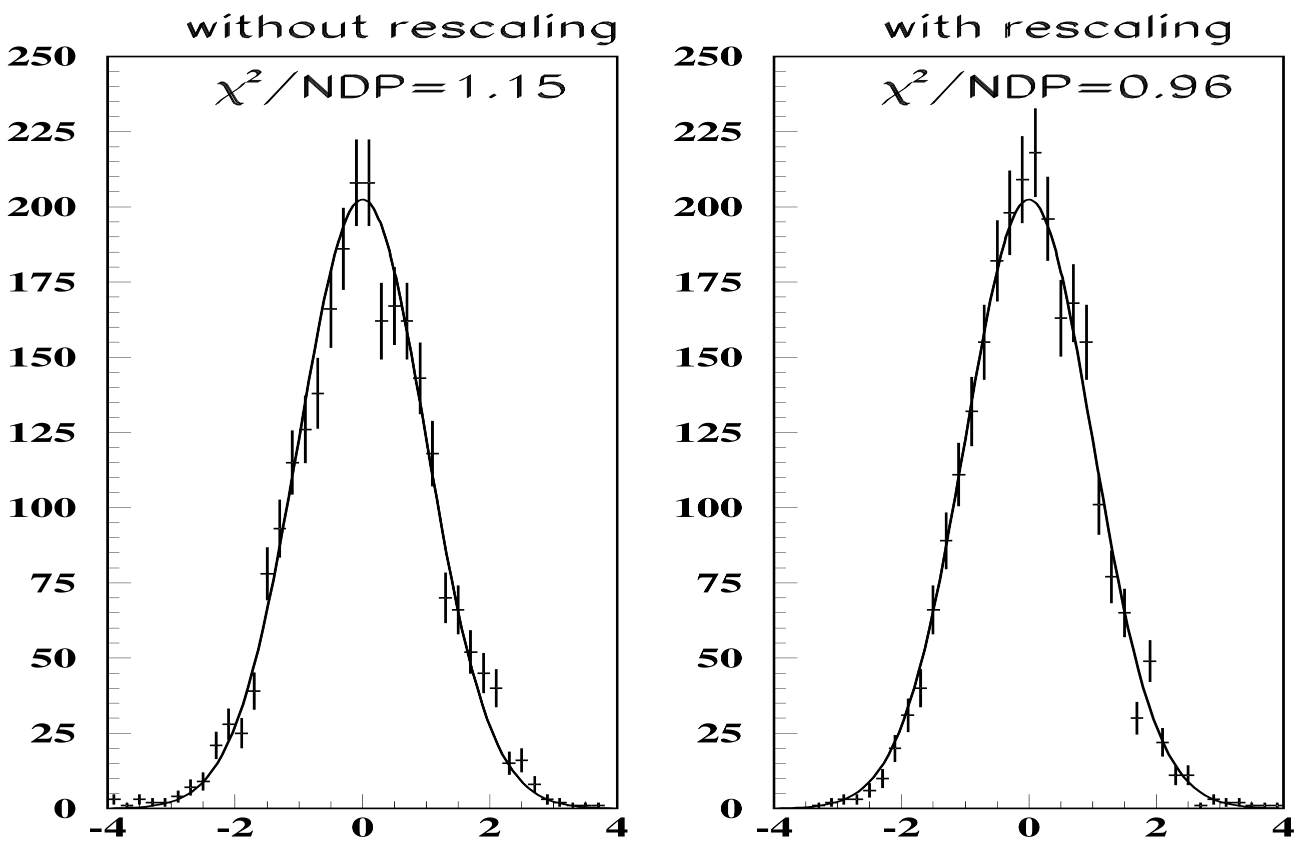


$\Delta\chi^2 = 100$ (CTEQ6)

SOLUTIONS: ERROR RESCALING

HOW CAN DATA FROM INCONSISTENT SETS BE INCLUDED?

ASSUME INCONSISTENCY DUE TO UNDERESTIMATED (SYST.) ERROR:



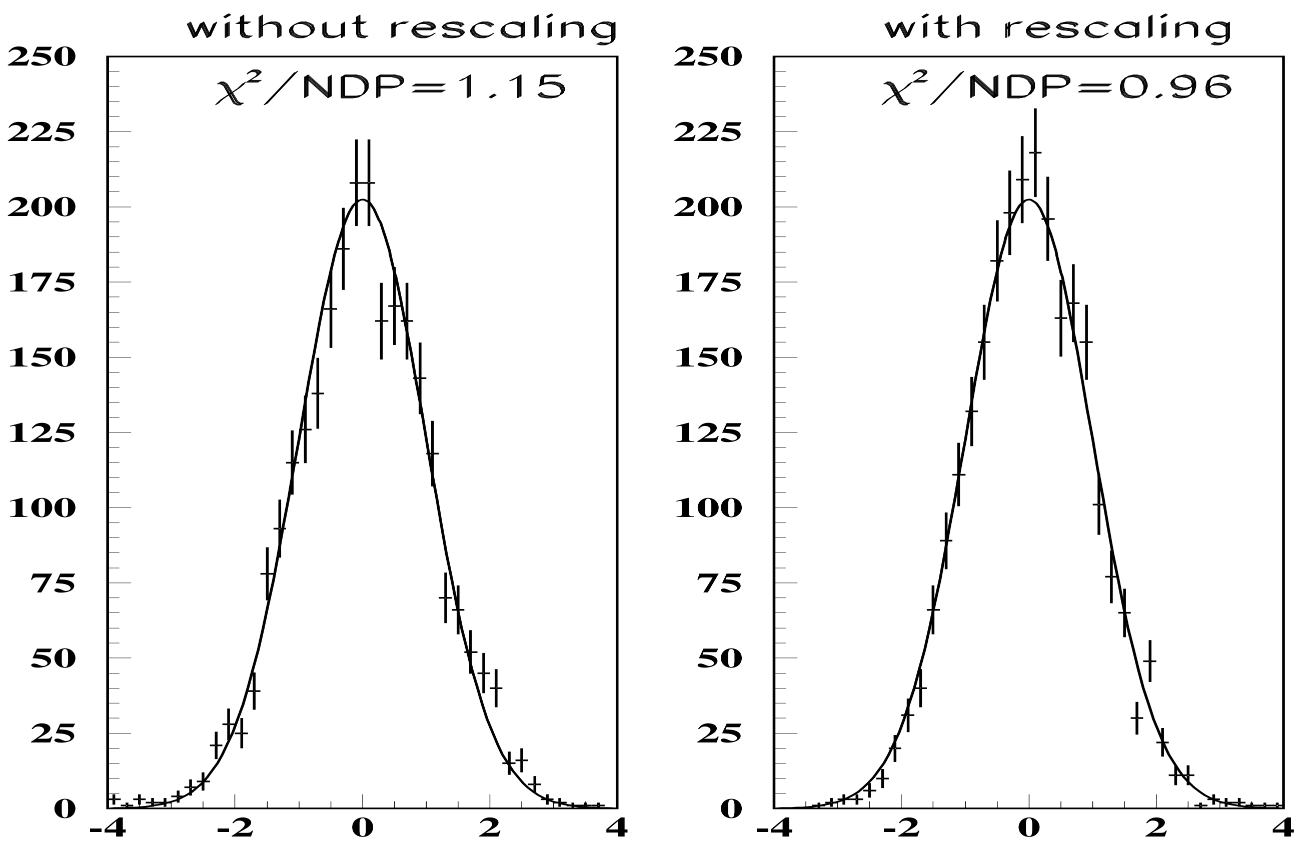
For the experiments with $\chi^2 > 1$ the statistical errors in data were rescaled in order to get $\chi^2 = 1$

ALEKHIN 2005

SOLUTIONS: ERROR RESCALING

HOW CAN DATA FROM INCONSISTENT SETS BE INCLUDED?

ASSUME INCONSISTENCY DUE TO UNDERESTIMATED (SYST.) ERROR:



For the experiments with $\chi^2 > 1$ the statistical errors in data were rescaled in order to get $\chi^2 = 1$

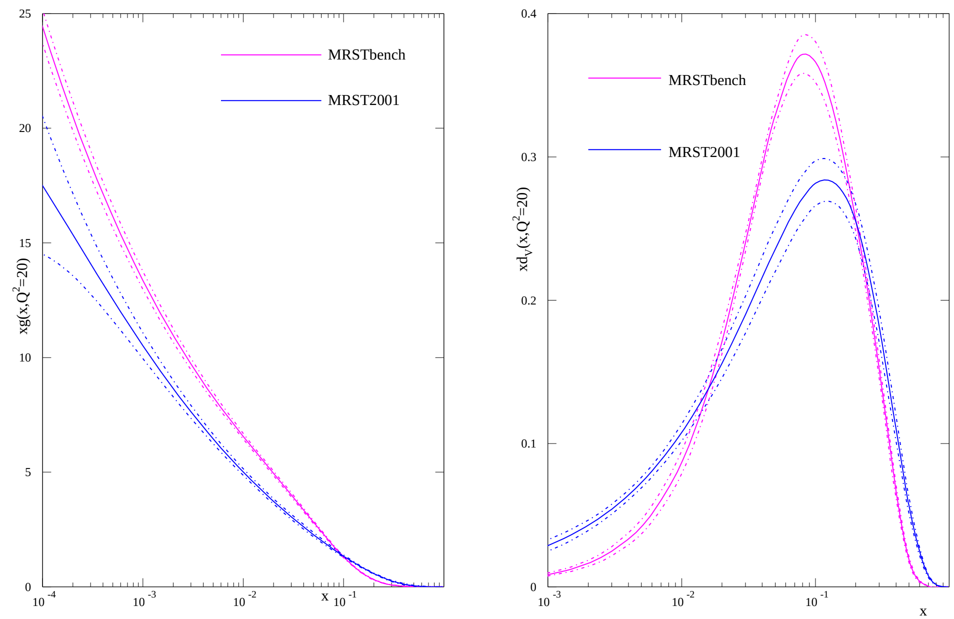
ALEKHIN 2005

ALTERNATIVE (ALEKHIN 2006): DISCARD INCONSISTENT DATA, RETAIN SUBSET

THE HERA-LHC BENCHMARK: AN IMPASSE

GLUON AND d_V: MRST VS. BENCH

HERA-LHC
BENCHMARK PARTONS
OBTAINED FROM NC DIS
DATA ONLY, $Q^2 > 9\ \mathrm{GeV}^2$

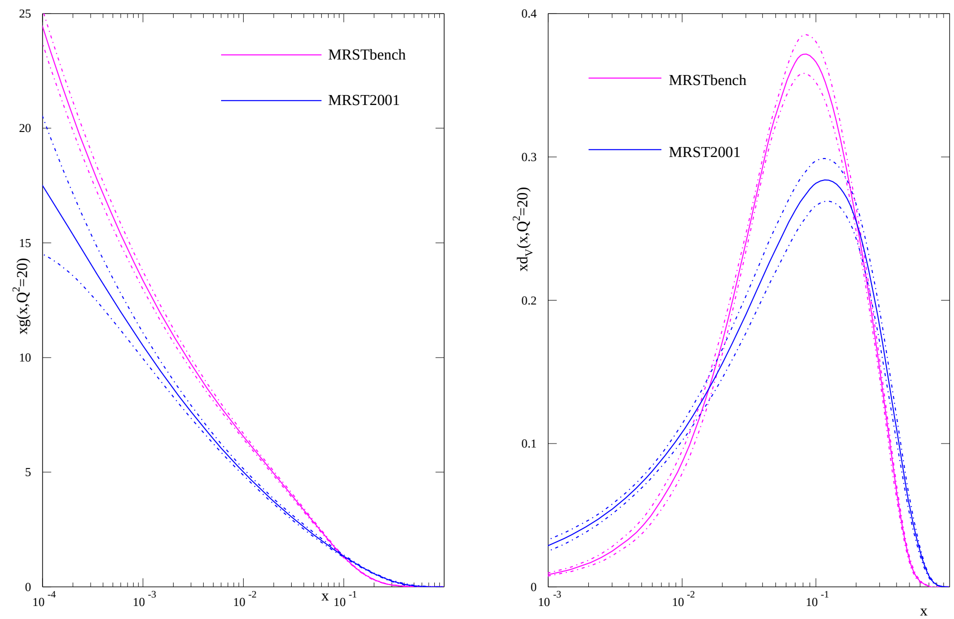


- IT IS UNSURPRIZING THAT CENTRAL VALUES DEPEND STRONGLY ON THE DATASET

THE HERA-LHC BENCHMARK: AN IMPASSE

GLUON AND d_V: MRST VS. BENCH

HERA-LHC
BENCHMARK PARTONS
OBTAINED FROM NC DIS
DATA ONLY, $Q^2 > 9\ \mathrm{GeV}^2$

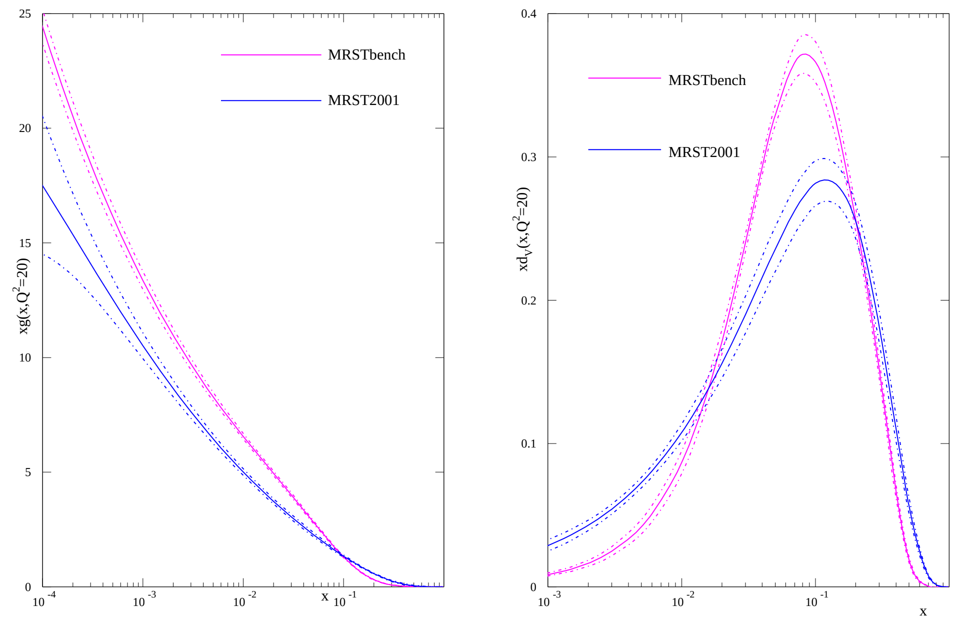


- IT IS UNSURPRIZING THAT CENTRAL VALUES DEPEND STRONGLY ON THE DATASET
- BUT IT IS VERY WORRISOME THAT THE RESULT WITH THE FULL DATA SET IS NOT WITHIN THE ERROR BAND OF THE RESULT FROM A DATA SUBSET

THE HERA-LHC BENCHMARK: AN IMPASSE

GLUON AND d_V: MRST VS. BENCH

HERA-LHC
BENCHMARK PARTONS
OBTAINED FROM NC DIS
DATA ONLY, $Q^2 > 9\ \mathrm{GeV}^2$



- IT IS UNSURPRIZING THAT CENTRAL VALUES DEPEND STRONGLY ON THE DATASET
- BUT IT IS VERY WORRISOME THAT THE RESULT WITH THE FULL DATA SET IS NOT WITHIN THE ERROR BAND OF THE RESULT FROM A DATA SUBSET

A NEW IDEA IS NEEDED!

THE BAYESIAN MONTE CARLO (Giele, Kosower, Keller 2001)

- generate a Monte-Carlo sample of fcts. with "reasonable" prior distn. (e.g. an available parton set) $\rightarrow$ representation of probability functional $\mathcal{P}[f_i]$
- calculate observables with functional integral
- update probability using Bayesian inference on MC sample: better agreement with data $\rightarrow$ more functions in sample
- iterate until convergence achieved

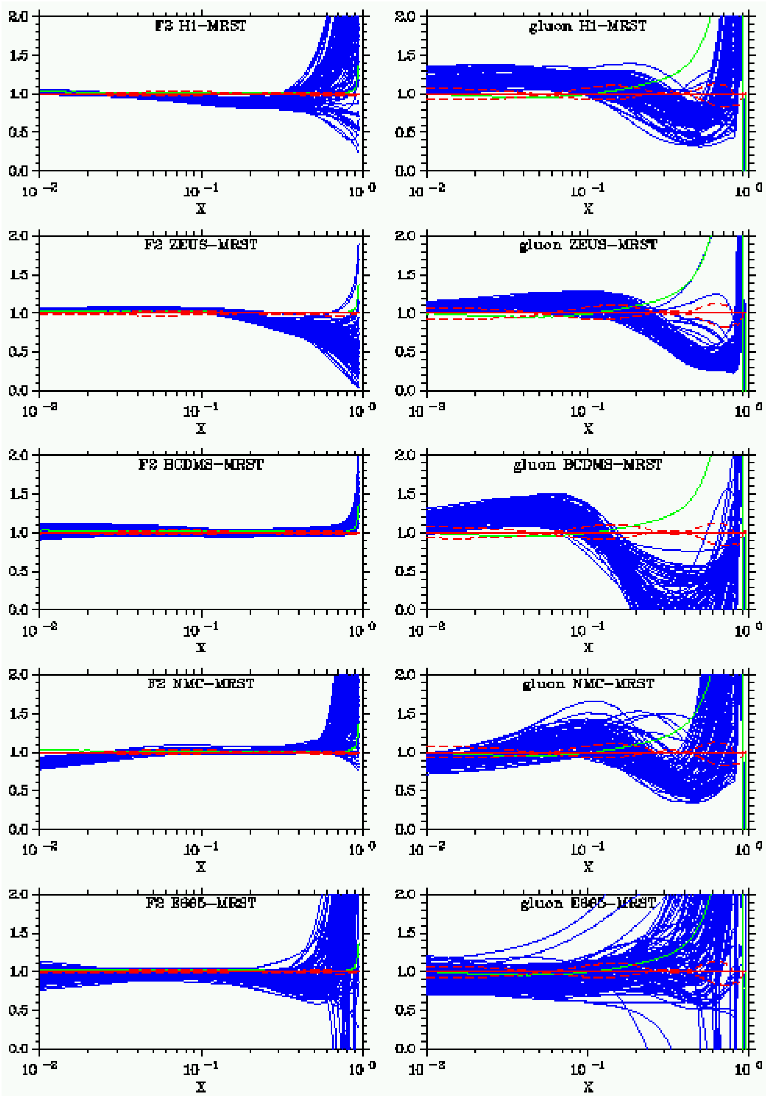
THE BAYESIAN MONTE CARLO (Giele, Kosower, Keller 2001)

- generate a Monte-Carlo sample of fcts. with "reasonable" prior distn. (e.g. an available parton set) $\rightarrow$ representation of probability functional $\mathcal{P}[f_i]$
- calculate observables with functional integral
- update probability using Bayesian inference on MC sample: better agreement with data $\rightarrow$ more functions in sample
- iterate until convergence achieved

PROBLEM IS MADE FINITE-DIMENSIONAL BY THE CHOICE OF PRIOR, BUT RESULT DO NOT DEPEND ON THE CHOICE IF SUFFICIENTLY GENERAL

HARD TO HANDLE "FLAT DIRECTIONS"
(Monte Carlo replicas which lead to same agreement with data);
COMPUTATIONALLY VERY INTENSIVE;
DIFFICULT TO ACHIEVE INDEP. FROM PRIOR

RESULT: FERMI PARTONS



F_2^{singlet} AND GLUON RATIOS FERMI/MRST

ONLY SUBSET OF DATA FITTED (H1, E665, BCDMS DIS DATA)

GOOD AGREEMENT WITH TEVATRON W XSECT

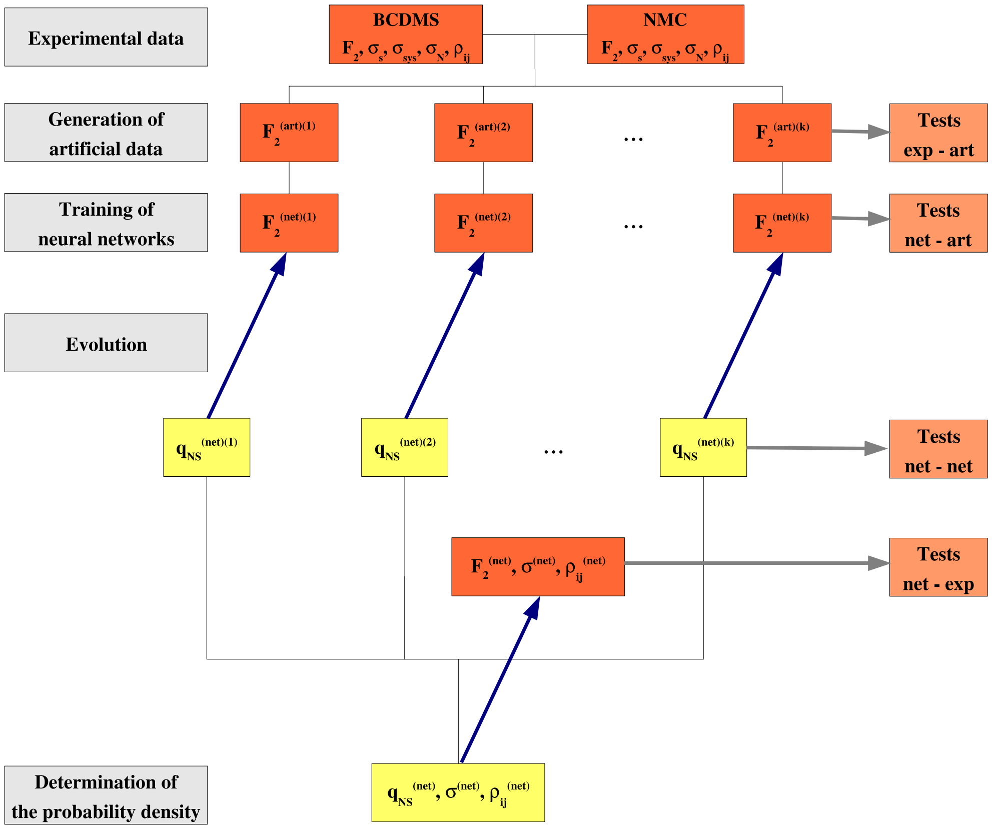
TROUBLE WITH VALUE OF α_s

THE NEURAL MONTE CARLO

THE NNPDF COLLABORATION

(2004: Del Debbio, SF, Latorre, Piccione, Rojo; 2007: +Ball, Guffanti, Ubiali)

BASIC IDEA: USE NEURAL NETWORKS AS UNIVERSAL UNBIASED INTERPOLANTS



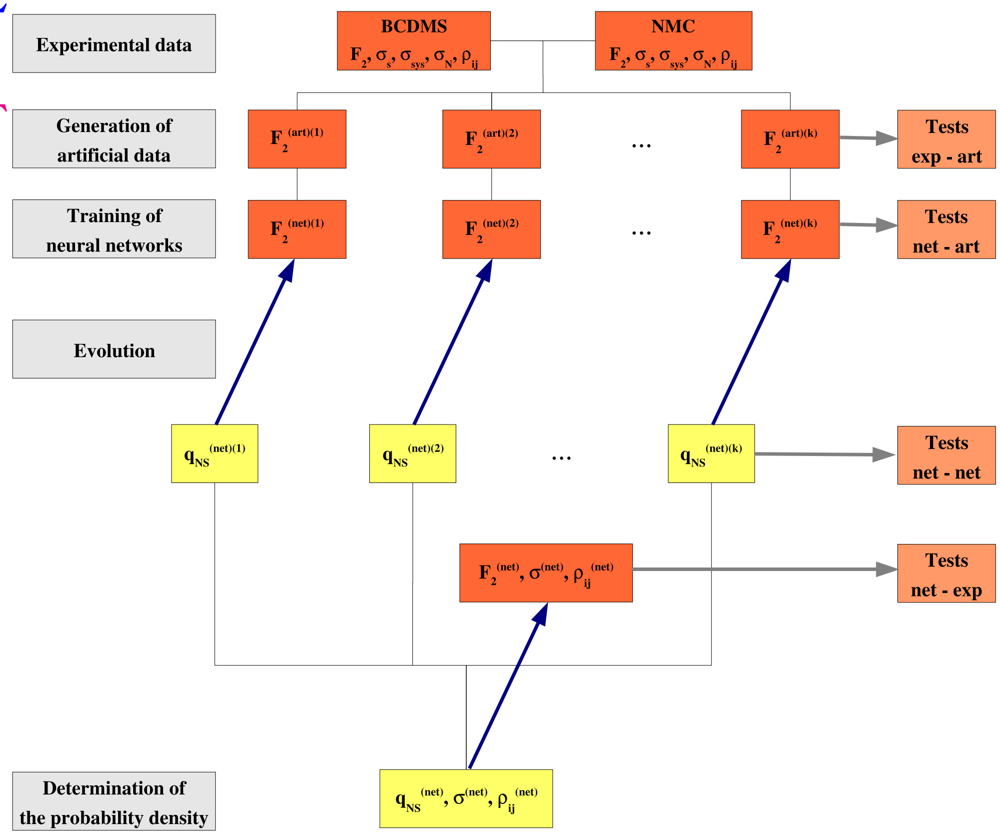
THE NEURAL MONTE CARLO

THE NNPDF COLLABORATION

(2004: Del Debbio, SF, Latorre, Piccione, Rojo; 2007: +Ball, Guffanti, Ubiali)

BASIC IDEA: USE NEURAL NETWORKS AS UNIVERSAL UNBIASED INTERPOLANTS

- GENERATE A SET OF MONTE CARLO REPLICAS $\sigma^{(k)}(p_i)$ OF THE ORIGINAL DATASET $\sigma^{(\mathrm{data})}(p_i)$
  $\Rightarrow$ REPRESENTATION OF $\mathcal{P}[\sigma(p_i)]$ AT DISCRETE SET OF POINTS p_i



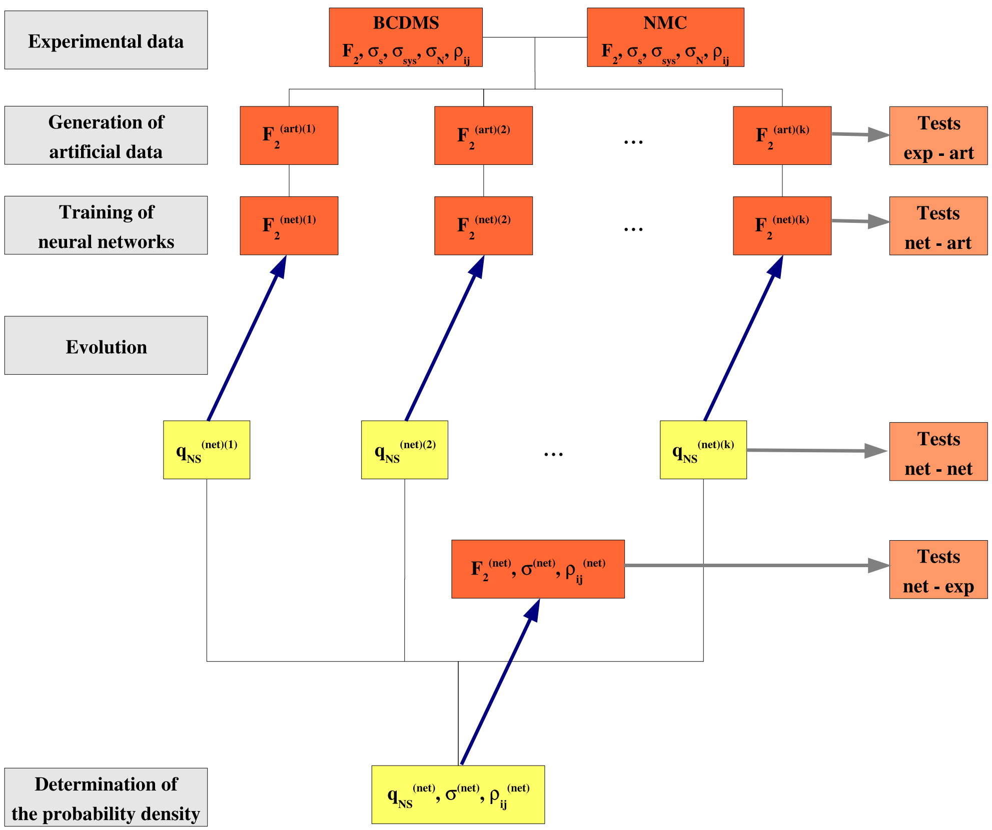
THE NEURAL MONTE CARLO

THE NNPDF COLLABORATION

(2004: Del Debbio, SF, Latorre, Piccione, Rojo; 2007: +Ball, Guffanti, Ubiali)

BASIC IDEA: USE NEURAL NETWORKS AS UNIVERSAL UNBIASED INTERPOLANTS

- GENERATE A SET OF MONTE CARLO REPLICAS $\sigma^{(k)}(p_i)$ OF THE ORIGINAL DATASET $\sigma^{(\mathrm{data})}(p_i)$
  $\Rightarrow$ REPRESENTATION OF $\mathcal{P}[\sigma(p_i)]$ AT DISCRETE SET OF POINTS p_i
- TRAIN A NEURAL NET FOR EACH PDF ON EACH REPLICA, THUS OBTAINING A NEURAL REPRESENTATION OF THE PDFS $f_i^{(net),(k)}$



THE NEURAL MONTE CARLO

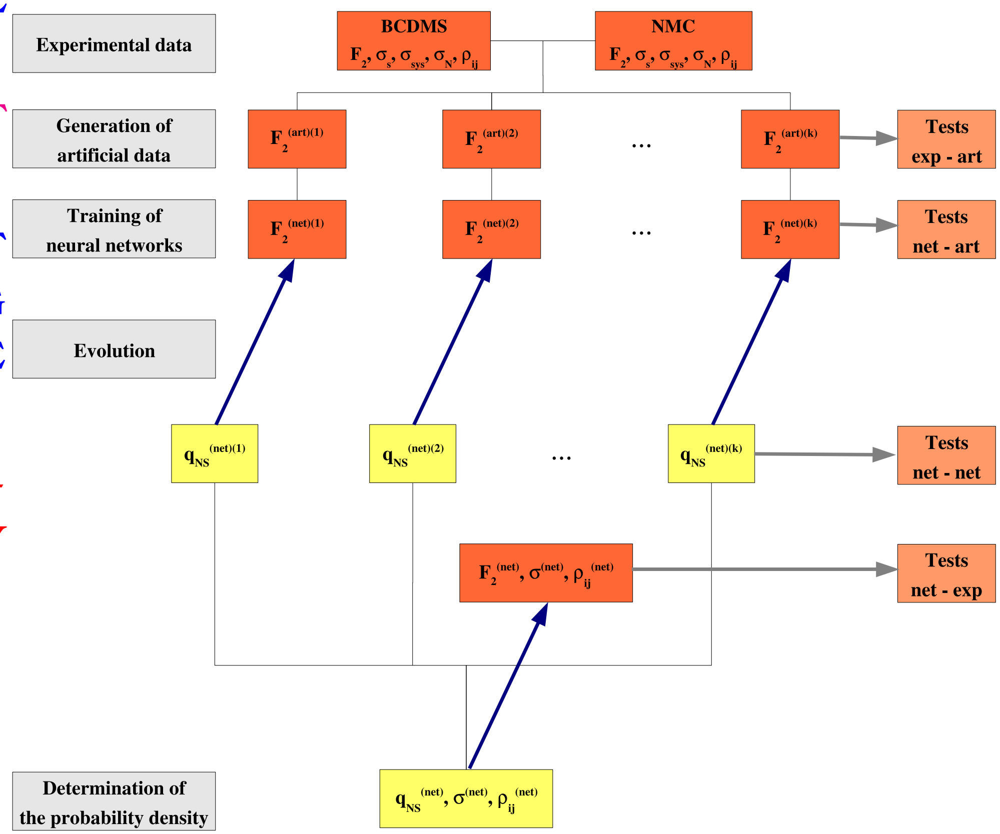
THE NNPDF COLLABORATION

(2004: Del Debbio, SF, Latorre, Piccione, Rojo; 2007: +Ball, Guffanti, Ubiali)

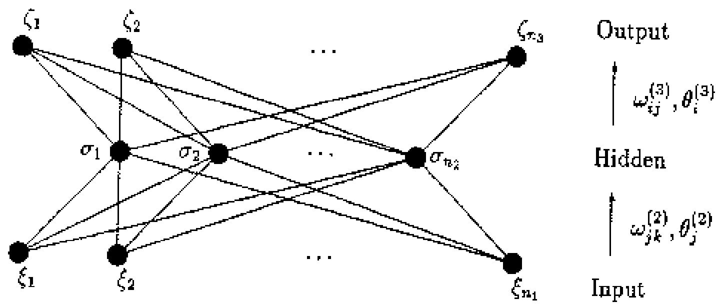
BASIC IDEA: USE NEURAL NETWORKS AS UNIVERSAL UNBIASED INTERPOLANTS

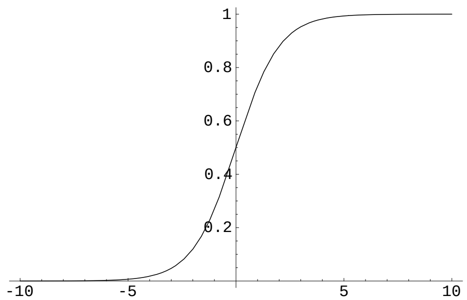
- GENERATE A SET OF MONTE CARLO REPLICAS $\sigma^{(k)}(p_i)$ OF THE ORIGINAL DATASET $\sigma^{(\mathrm{data})}(p_i)$
  $\Rightarrow$ REPRESENTATION OF $\mathcal{P}[\sigma(p_i)]$ AT DISCRETE SET OF POINTS p_i
- TRAIN A NEURAL NET FOR EACH PDF ON EACH REPLICA, THUS OBTAINING A NEURAL REPRESENTATION OF THE PDFS $f_i^{(net),(k)}$
- THE SET OF NEURAL NETS IS A REPRESENTATION OF THE PROBABILITY DENSITY:

$$\left\langle \sigma\left[f_i\right]\right\rangle = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} \sigma\left[f_i{}^{(net)(k)}\right]$$



WHAT ARE NEURAL NETWORKS?

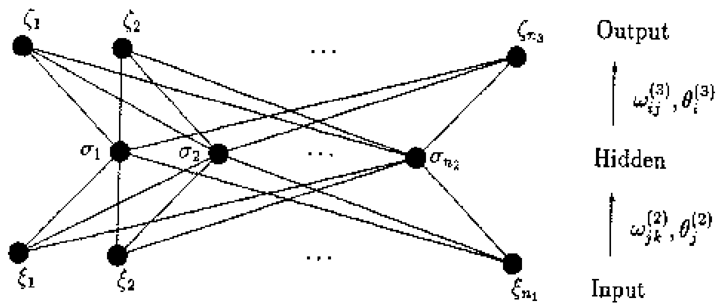


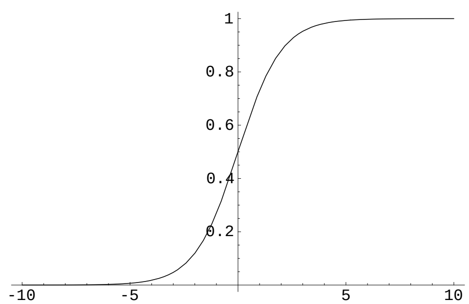


Multilayer feed-forward networks

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

WHAT ARE NEURAL NETWORKS?





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

JUST ANOTHER SET OF BASIS FUNCTIONS!

A 1-2-1 NN: $\xi_1^{(3)}(\xi_1^{(1)}) = \dfrac{1}{1+e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)} - \xi_1^{(1)}\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)} - \xi_1^{(1)}\omega_{21}^{(1)}}}}}$

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

WHAT ARE NEURAL NETWORKS GOOD FOR?

WHAT ARE NEURAL NETWORKS GOOD FOR?

IN A STANDARD FIT, ONE LOOKS FOR MINIMUM χ^2 WITH GIVEN FINITE PARM.

- IF THE BASIS IS TOO LARGE, THE FIT NEVER CONVERGES
- IF THE BASIS IS TOO SMALL, THE FIT IS BIASED

WHAT ARE NEURAL NETWORKS GOOD FOR?

IN A STANDARD FIT, ONE LOOKS FOR MINIMUM χ^2 WITH GIVEN FINITE PARM.

- IF THE BASIS IS TOO LARGE, THE FIT NEVER CONVERGES
- IF THE BASIS IS TOO SMALL, THE FIT IS BIASED

Q: HOW CAN ONE BE SURE THAT THE COMPROMISE IS UNBIASED?

WHAT ARE NEURAL NETWORKS GOOD FOR?

IN A STANDARD FIT, ONE LOOKS FOR MINIMUM χ^2 WITH GIVEN FINITE PARM.

- IF THE BASIS IS TOO LARGE, THE FIT NEVER CONVERGES
- IF THE BASIS IS TOO SMALL, THE FIT IS BIASED

Q: HOW CAN ONE BE SURE THAT THE COMPROMISE IS UNBIASED?

IN A NEURAL FIT, SMOOTHNESS DECREASES AS FIT QUALITY IMPROVES:

WHAT ARE NEURAL NETWORKS GOOD FOR?

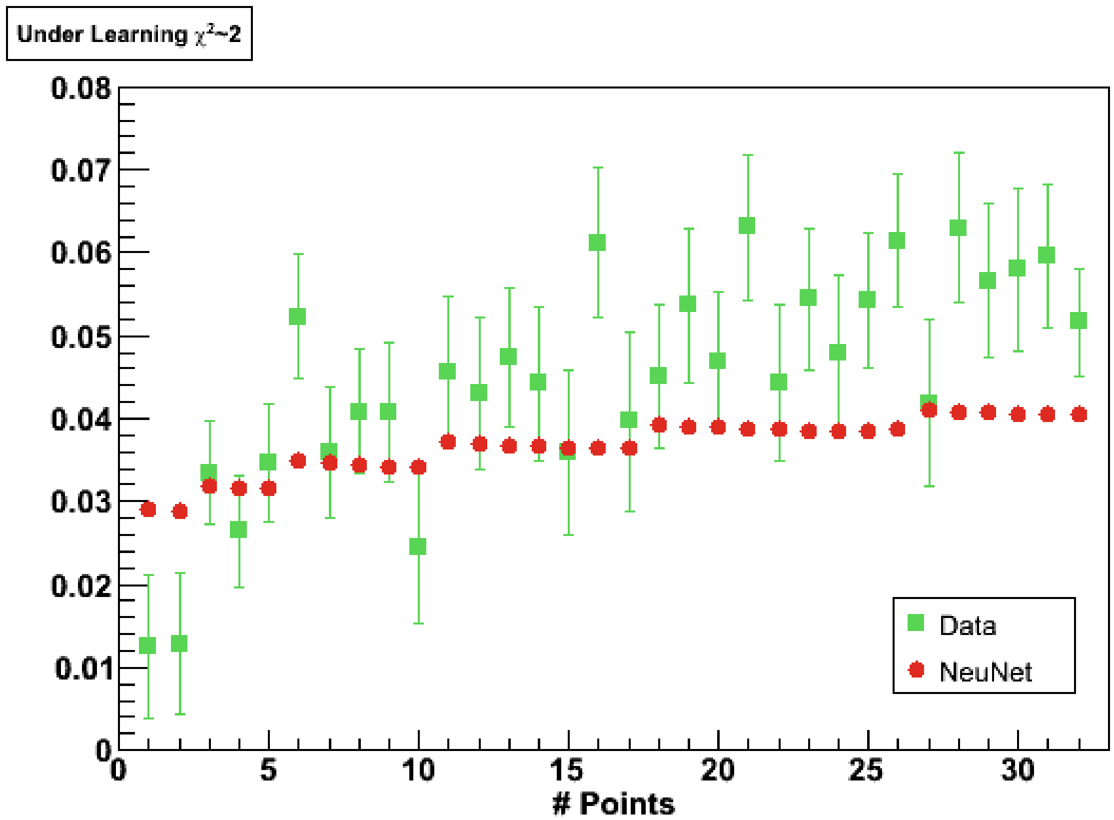
IN A STANDARD FIT, ONE LOOKS FOR MINIMUM χ^2 WITH GIVEN FINITE PARM.

- IF THE BASIS IS TOO LARGE, THE FIT NEVER CONVERGES
- IF THE BASIS IS TOO SMALL, THE FIT IS BIASED

Q: HOW CAN ONE BE SURE THAT THE COMPROMISE IS UNBIASED?

IN A NEURAL FIT, SMOOTHNESS DECREASES AS FIT QUALITY IMPROVES:

UNDERLEARNING



WHAT ARE NEURAL NETWORKS GOOD FOR?

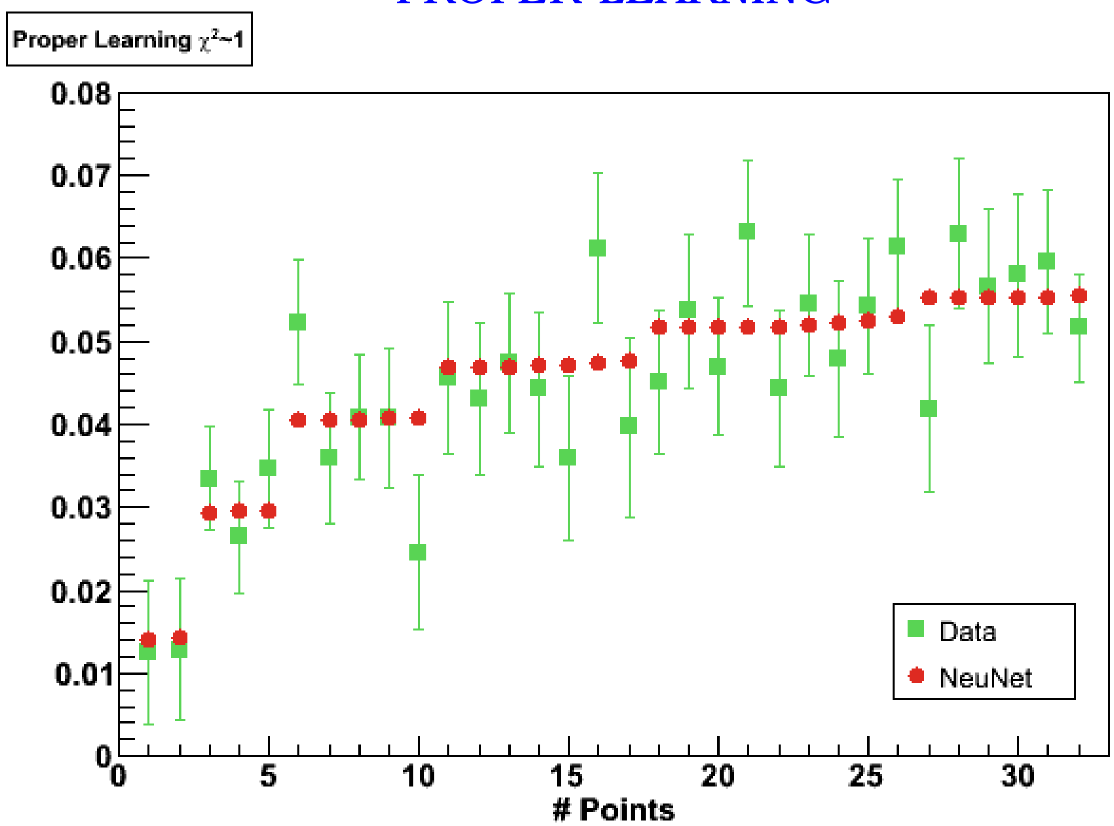
IN A STANDARD FIT, ONE LOOKS FOR MINIMUM χ^2 WITH GIVEN FINITE PARM.

- IF THE BASIS IS TOO LARGE, THE FIT NEVER CONVERGES
- IF THE BASIS IS TOO SMALL, THE FIT IS BIASED

Q: HOW CAN ONE BE SURE THAT THE COMPROMISE IS UNBIASED?

IN A NEURAL FIT, SMOOTHNESS DECREASES AS FIT QUALITY IMPROVES:

PROPER LEARNING



WHAT ARE NEURAL NETWORKS GOOD FOR?

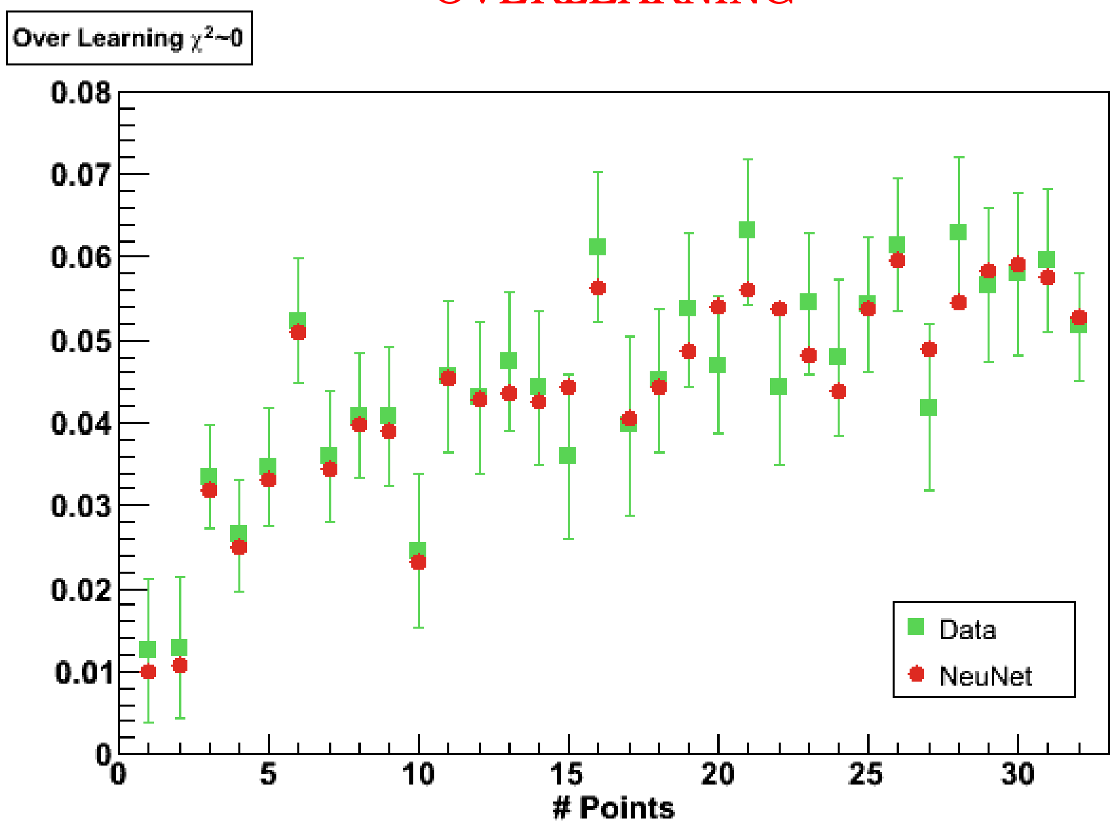
IN A STANDARD FIT, ONE LOOKS FOR MINIMUM χ^2 WITH GIVEN FINITE PARM.

- IF THE BASIS IS TOO LARGE, THE FIT NEVER CONVERGES
- IF THE BASIS IS TOO SMALL, THE FIT IS BIASED

Q: HOW CAN ONE BE SURE THAT THE COMPROMISE IS UNBIASED?

IN A NEURAL FIT, SMOOTHNESS DECREASES AS FIT QUALITY IMPROVES:

OVERLEARNING



WHAT ARE NEURAL NETWORKS GOOD FOR?

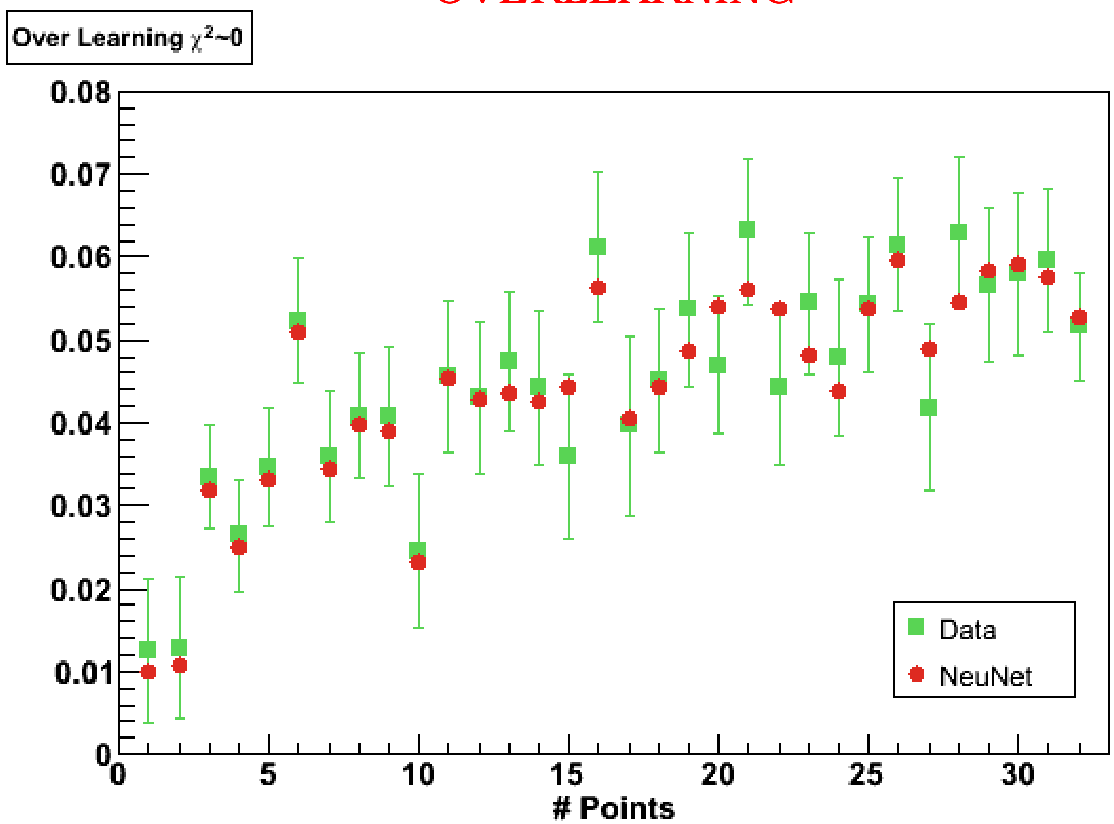
IN A STANDARD FIT, ONE LOOKS FOR MINIMUM χ^2 WITH GIVEN FINITE PARM.

- IF THE BASIS IS TOO LARGE, THE FIT NEVER CONVERGES
- IF THE BASIS IS TOO SMALL, THE FIT IS BIASED

Q: HOW CAN ONE BE SURE THAT THE COMPROMISE IS UNBIASED?

IN A NEURAL FIT, SMOOTHNESS DECREASES AS FIT QUALITY IMPROVES:

OVERLEARNING

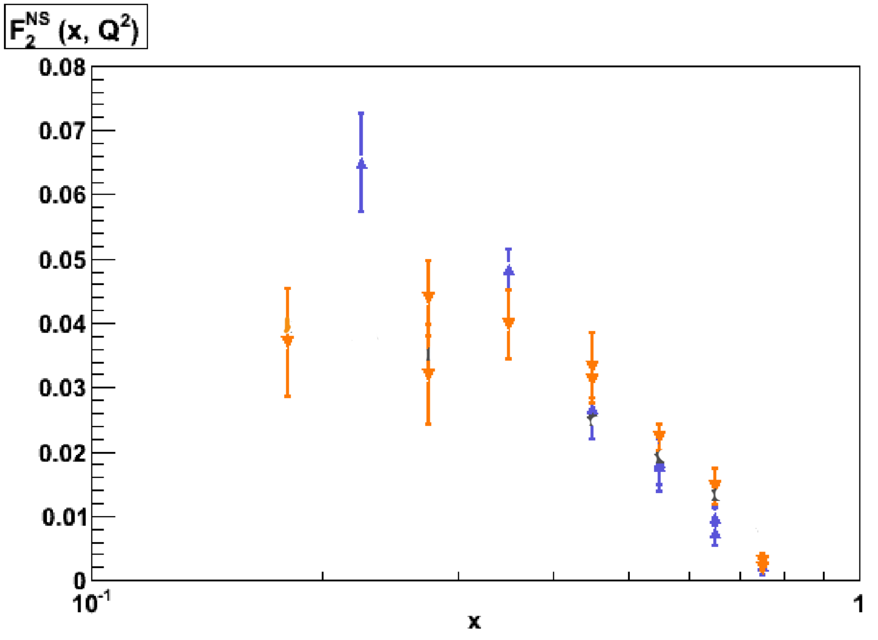


A: STOP THE FIT BEFORE OVERLEARNING SETS IN!

THE STOPPING CRITERION

MINIMIZE BY GENETIC ALGORITHM:
AT EACH GENERATION, THE χ^2 EITHER UNCHANGED OR DECREASING

- DIVIDE THE DATA IN TWO SETS: TRAINING AND VALIDATION
- MINIMIZE THE χ^2 OF THE DATA IN THE TRAINING SET
- AT EACH ITERATION, COMPUTE THE χ^2 FOR THE DATA IN THE VALIDATION SET (NOT USED FOR FITTING)
- WHEN THE VALIDATION χ^2 STOPS DECREASING, STOP THE FIT

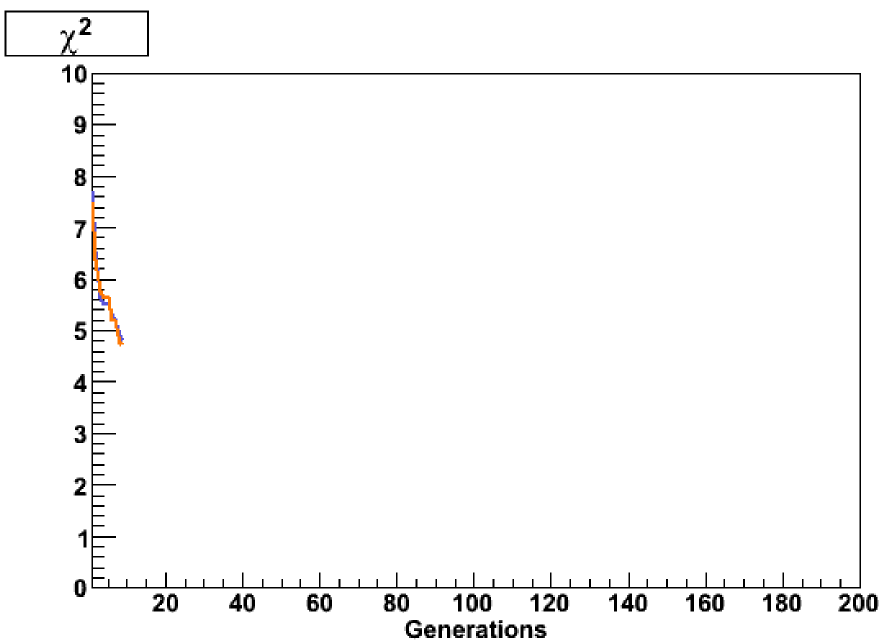


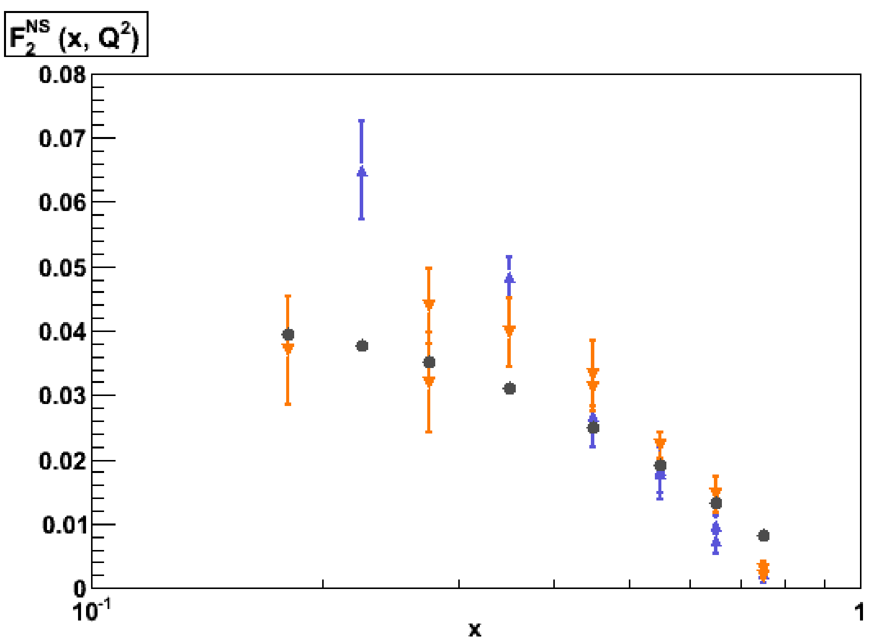
THE STOPPING CRITERION

MINIMIZE BY GENETIC ALGORITHM:
AT EACH GENERATION, THE χ^2 EITHER UNCHANGED OR DECREASING

- DIVIDE THE DATA IN TWO SETS: TRAINING AND VALIDATION
- MINIMIZE THE χ^2 OF THE DATA IN THE TRAINING SET
- AT EACH ITERATION, COMPUTE THE χ^2 FOR THE DATA IN THE VALIDATION SET (NOT USED FOR FITTING)
- WHEN THE VALIDATION χ^2 STOPS DECREASING, STOP THE FIT

GO!



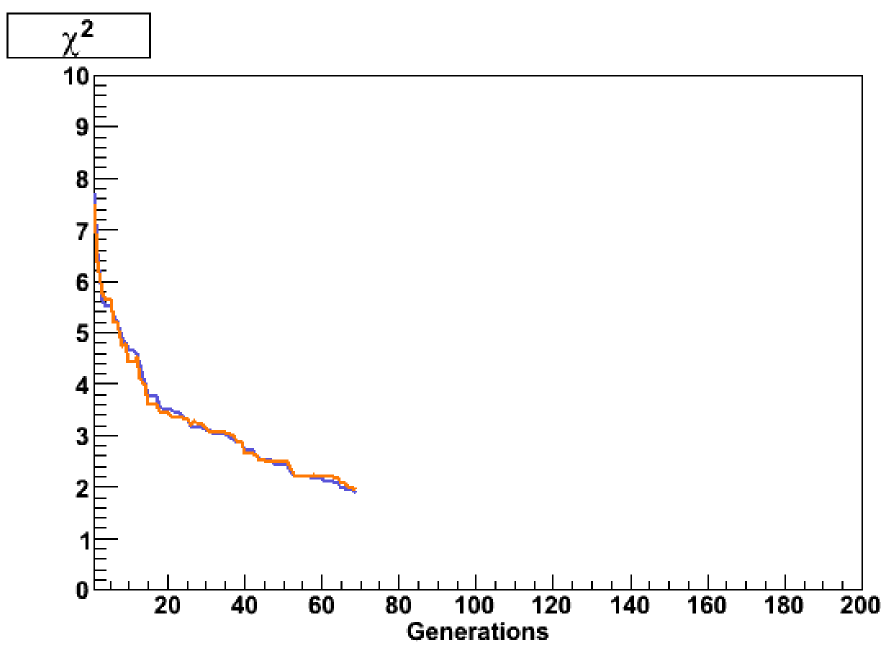


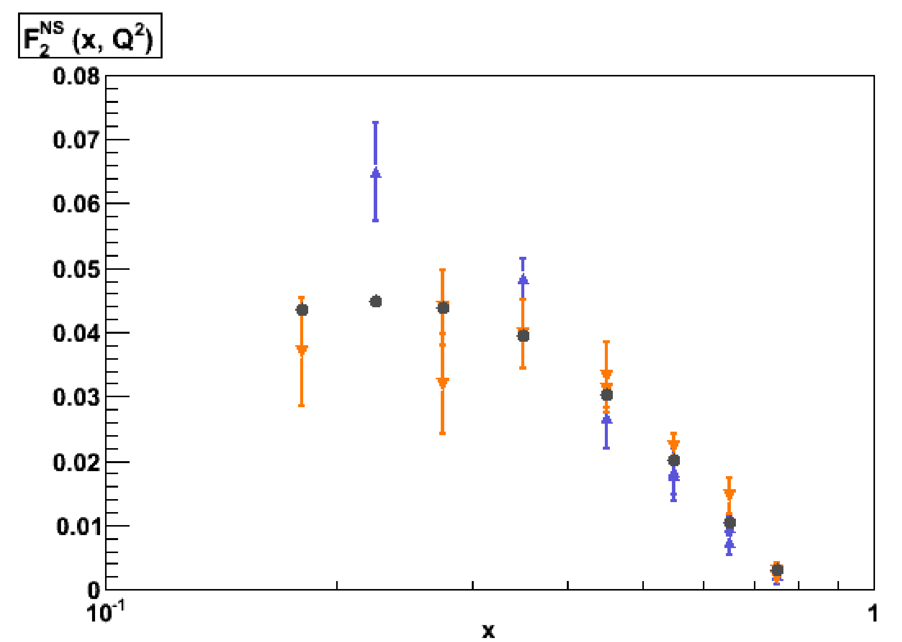
THE STOPPING CRITERION

MINIMIZE BY GENETIC ALGORITHM:
AT EACH GENERATION, THE χ^2 EITHER UNCHANGED OR DECREASING

- DIVIDE THE DATA IN TWO SETS: TRAINING AND VALIDATION
- MINIMIZE THE χ^2 OF THE DATA IN THE TRAINING SET
- AT EACH ITERATION, COMPUTE THE χ^2 FOR THE DATA IN THE VALIDATION SET (NOT USED FOR FITTING)
- WHEN THE VALIDATION χ^2 STOPS DECREASING, STOP THE FIT

STOP!



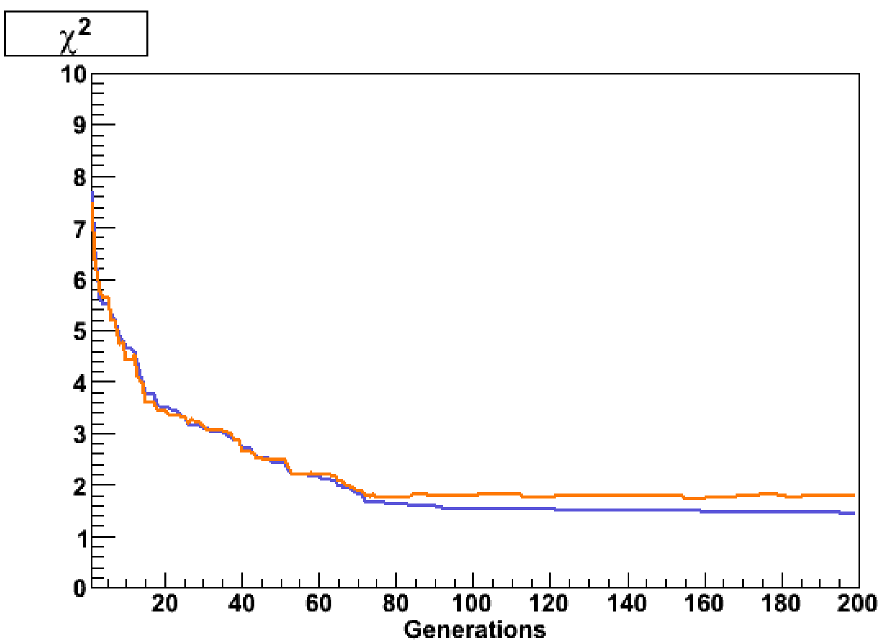


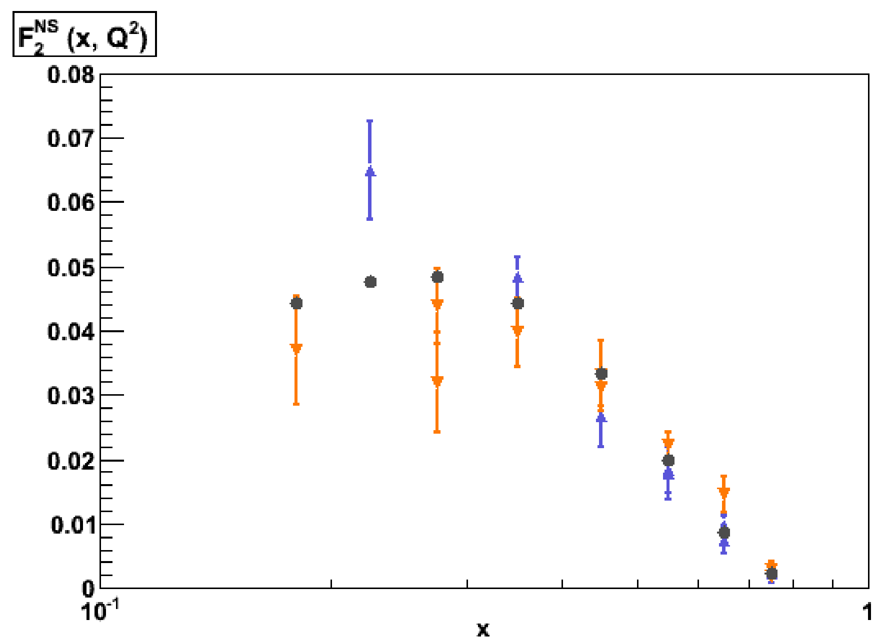
THE STOPPING CRITERION

MINIMIZE BY GENETIC ALGORITHM:
AT EACH GENERATION, THE χ^2 EITHER UNCHANGED OR DECREASING

- DIVIDE THE DATA IN TWO SETS: TRAINING AND VALIDATION
- MINIMIZE THE χ^2 OF THE DATA IN THE TRAINING SET
- AT EACH ITERATION, COMPUTE THE χ^2 FOR THE DATA IN THE VALIDATION SET (NOT USED FOR FITTING)
- WHEN THE VALIDATION χ^2 STOPS DECREASING, STOP THE FIT

TOO LATE!



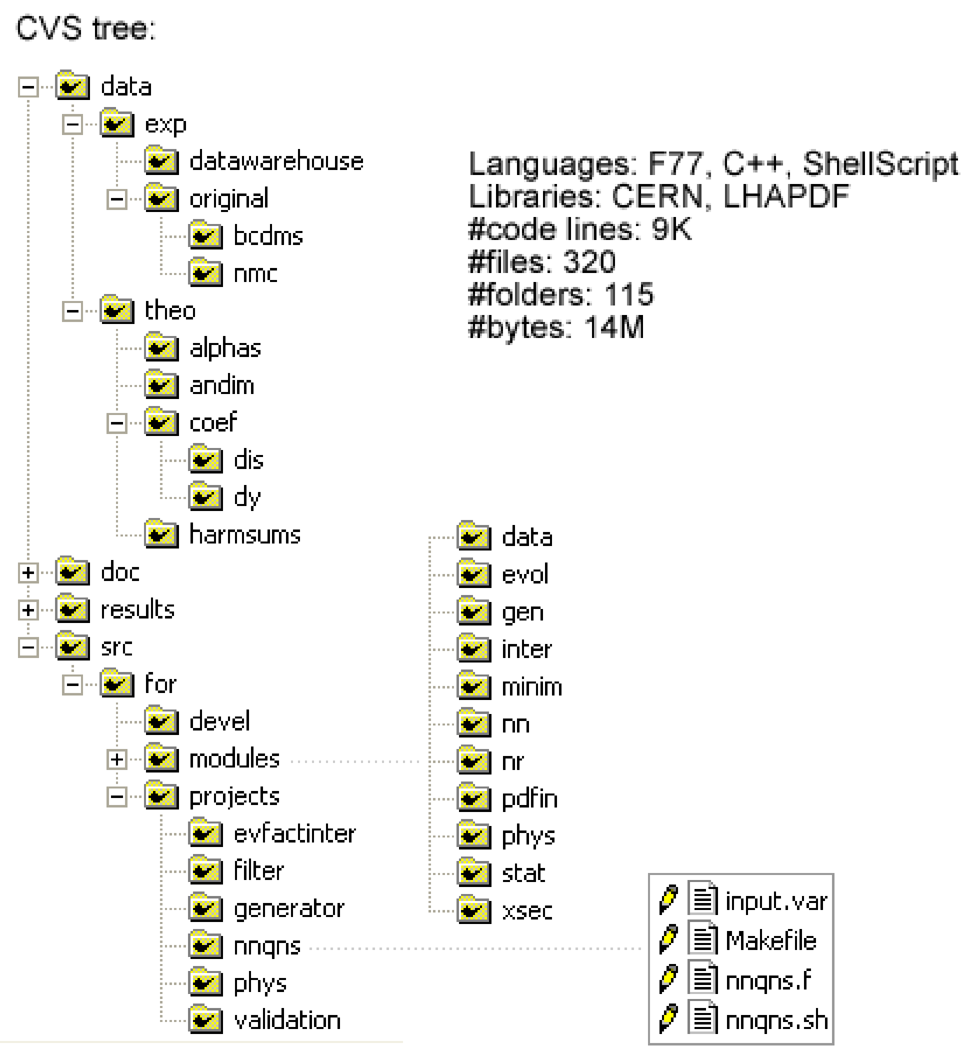


A FIT OF THE ISOTRIPLET QUARK DISTRIBUTION

hep-ph/0701127

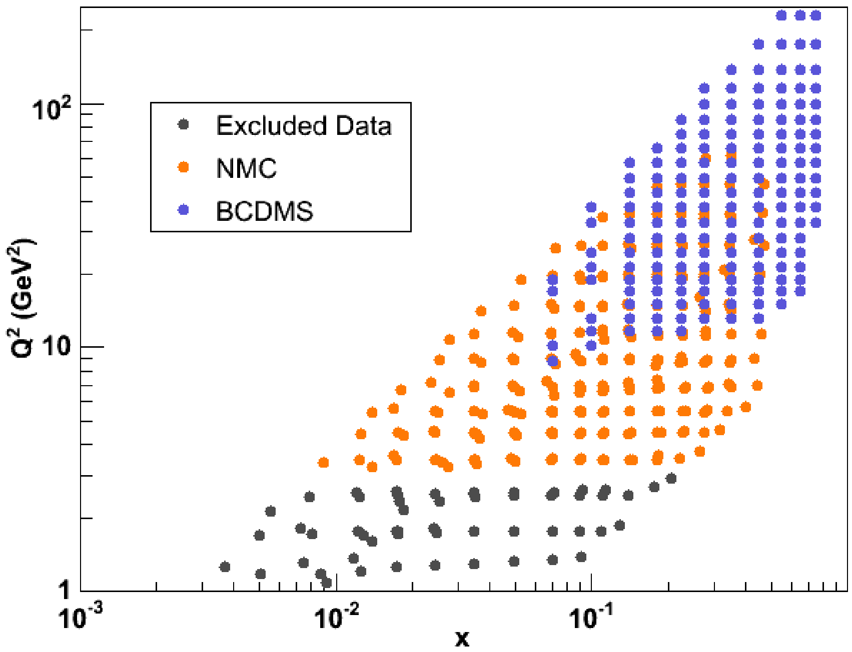
THE CVS TREE STRUCTURE

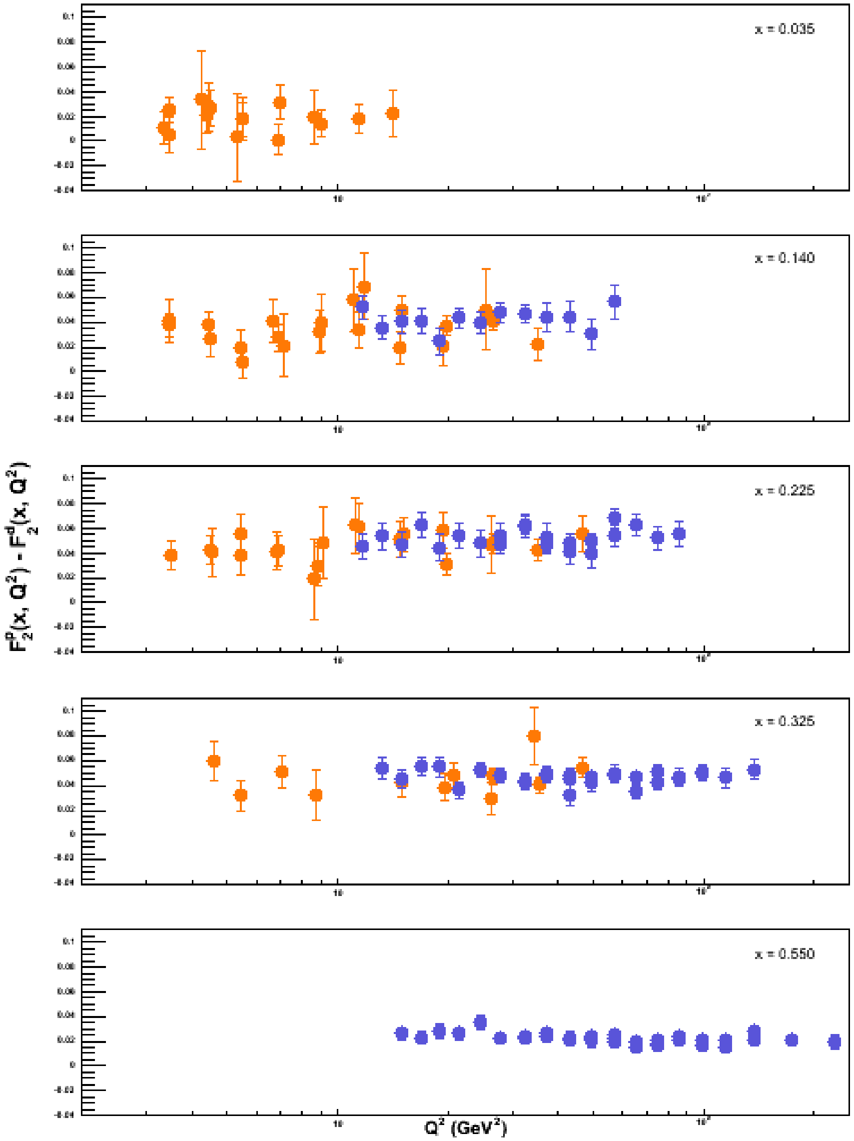
(now switched to SVN)



- FIRST FULL TEST OF THE METHOD AND ITS IMPLEMENTATION
- NONTRIVIAL ISSUES OF CODE DESIGN AND TASK COORDINATION
- OBJECT-ORIENTED & MODULAR STRUCTURE OF THE CODE

THE DATA





- PHYSICAL OBSERVABLE: $F_2^p - F_2^d$
- 229 NMC & 254 BCDMS DATA WITH $Q^2 > 3\ \mathrm{GeV}^2$
- DETERMINATION OF $q^{\mathrm{NS}}(x) = u(x) + \bar{u}(x) - \left(d(x) + \bar{d}(x)\right)$ AT REF. SCALE $Q_0^2 = 3\ \mathrm{GeV}^2$ AT LO, NLO, NNLO

Monte Carlo data generation

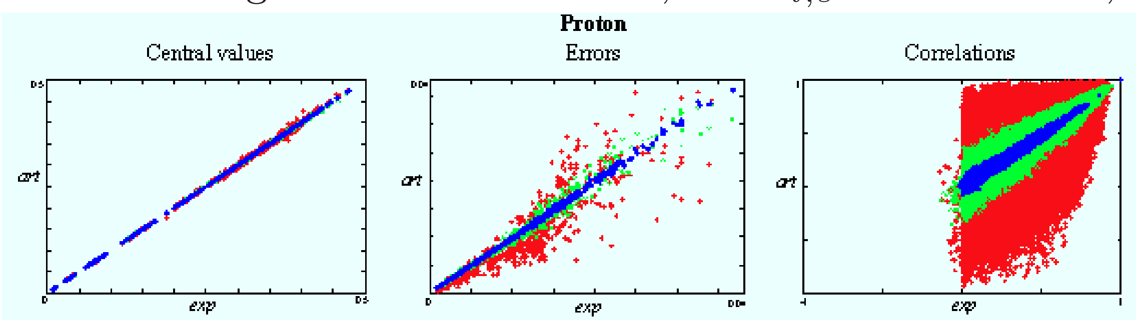
- BCDMS+ NMC proton & deuteron F_2 data (full correlated systematics available), taken at 4 beam energies
- on top of stat. errors, 4 systematics + 1 normalization (NMC) or 6 systematics + 1 absolute & 2 relative normalizations (BCDMS), with various forms of correlation (full, or for each target, or 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

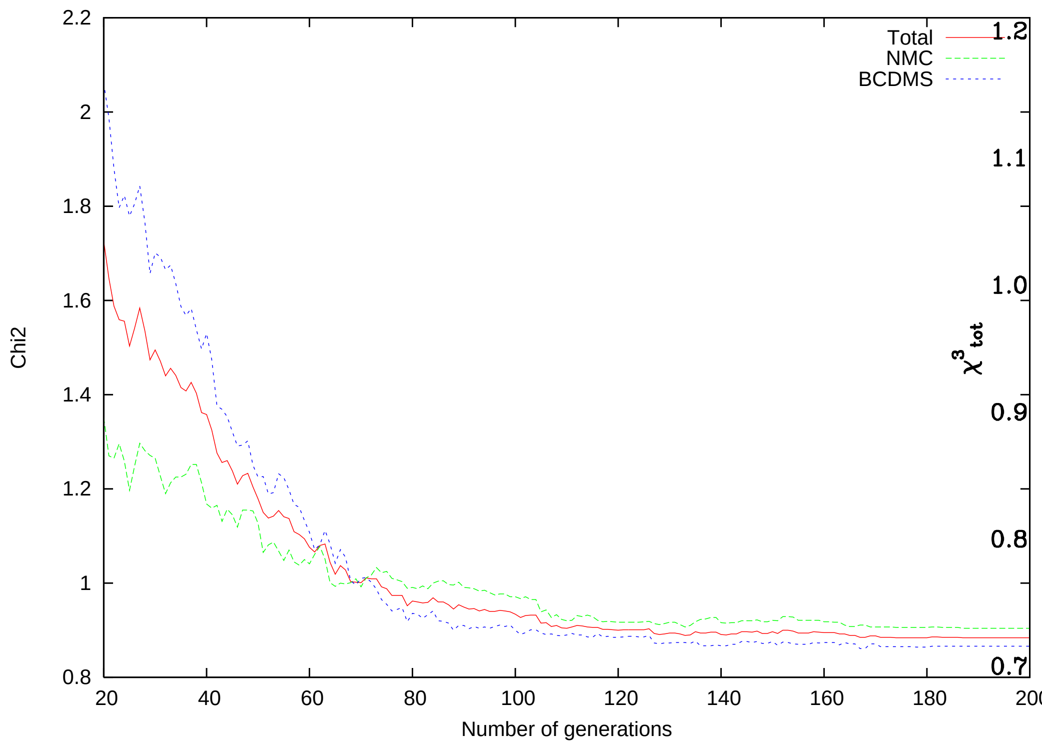
PERTURBATIVE EVOLUTION

- PARAMETRIZE INITIAL PDFS AS A FUNCTION OF x
- DETERMINE GREEN'S FUNCTION FOR ALTARELLI-PARISI EVOLUTION $\Gamma(x, \alpha_s\left(Q^2\right), \alpha_s\left(Q_0^2\right))$ (note it is a distribution)
- DETERMINE EVOLVED PDF AS $q(x, Q^2) = Gq(x, Q_0^2) + \int_x^1 \frac{dy}{y}\Gamma^{(+)}(y, \alpha_s\left(Q^2\right), \alpha_s\left(Q_0^2\right))q\left(\frac{x}{y}, Q_0^2\right)$
- GREEN FUNCTION CAN BE INTERPOLATED OR COMPUTED ON A GRID AND STORED
- EVOLUTION AND INTERPOLATION FULLY BENCHMARKED

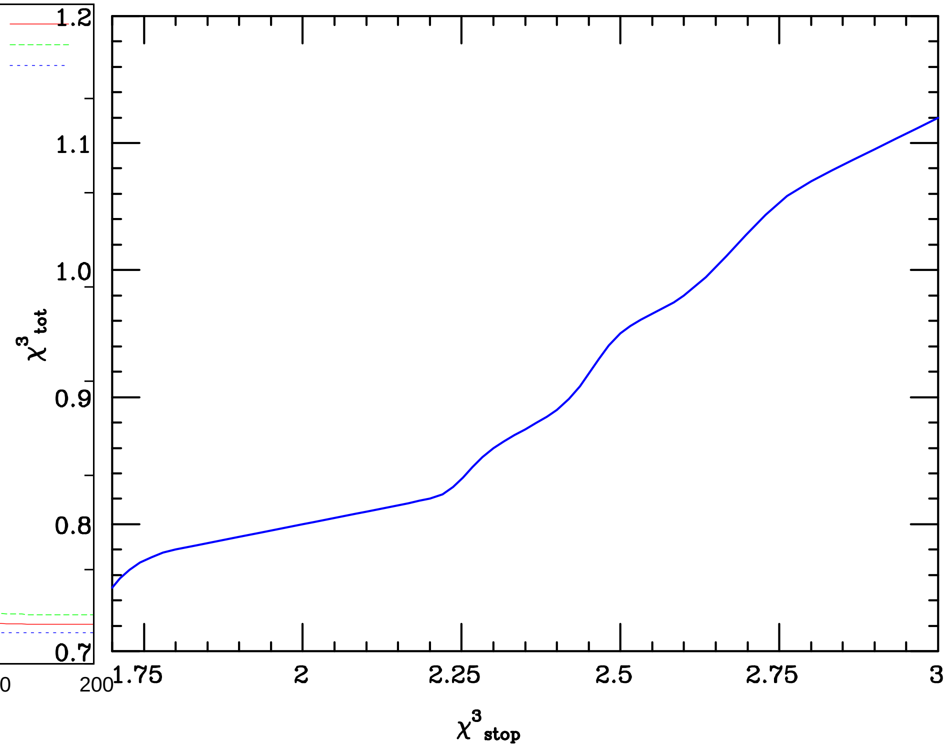
TRAINING...

- EACH NEURAL NET IS FITTED TO A PSEUDODATA REPLICA BY MINIMIZING ITS χ^2
- MINIMIZATION THROUGH GENETIC ALGORITHM + REWEIGHTING OF EXPERIMENTS
- QUALITY OF FIT MEASURED BY χ^2 OF AVERAGE OF NN COMPARED TO DATA

χ^2 OF BEST FIT



χ^2 OF BEST FIT VS. AVERAGE χ^2

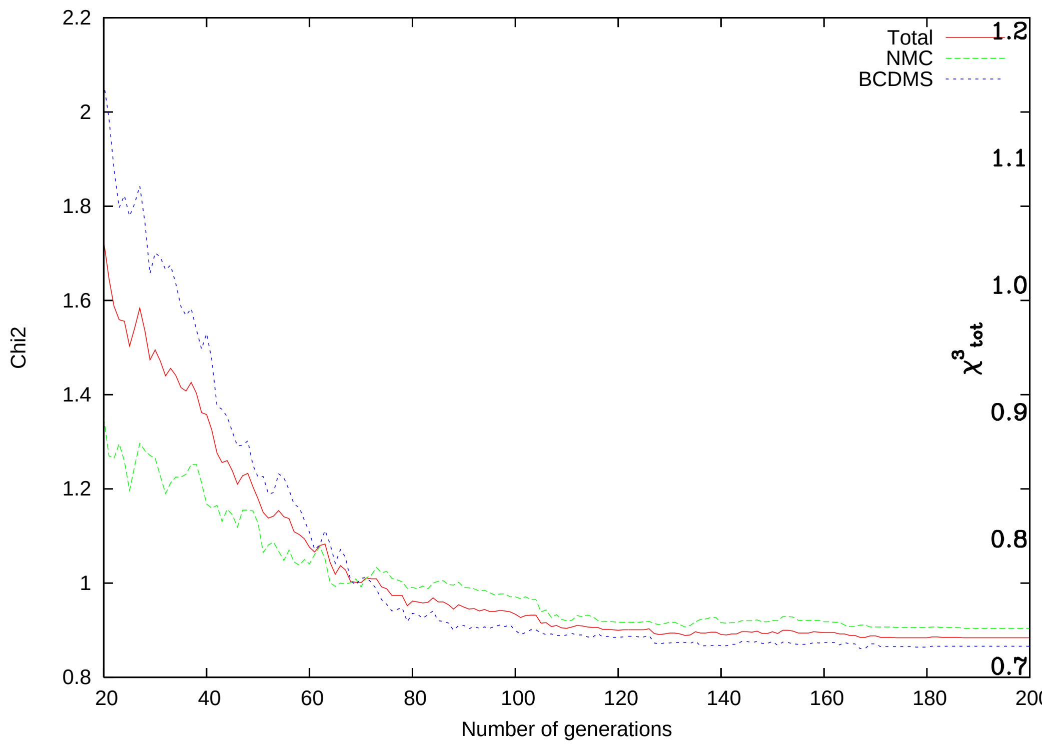


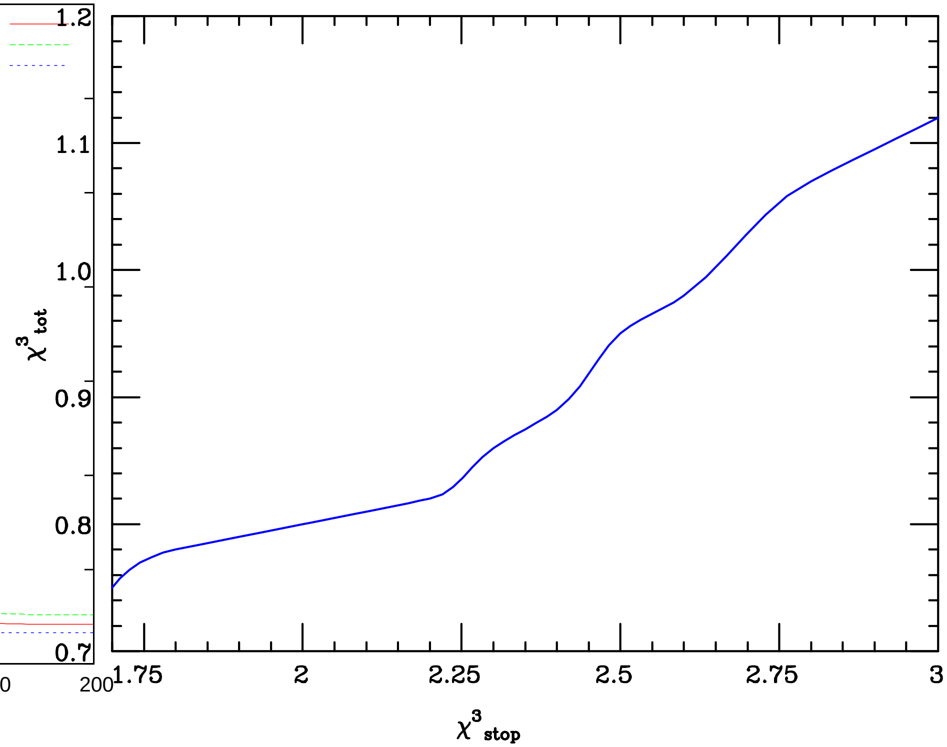
TRAINING...

- EACH NEURAL NET IS FITTED TO A PSEUDODATA REPLICA BY MINIMIZING ITS χ^2
- MINIMIZATION THROUGH GENETIC ALGORITHM + REWEIGHTING OF EXPERIMENTS
- QUALITY OF FIT MEASURED BY χ^2 OF AVERAGE OF NN COMPARED TO DATA

χ^2 OF BEST FIT

χ^2 OF BEST FIT VS. AVERAGE χ^2



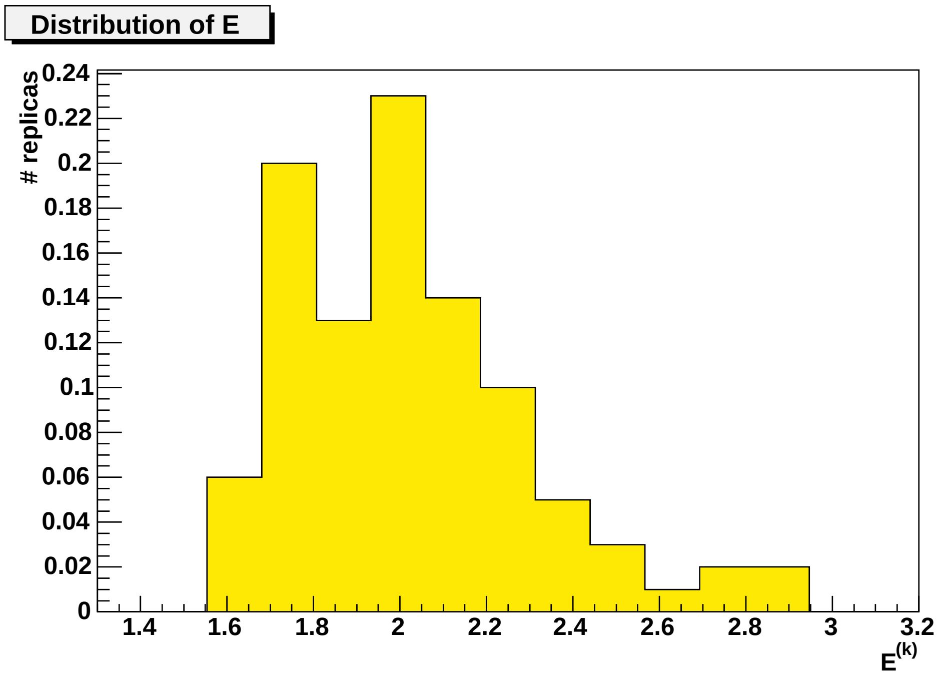


- IF NO STOPPING IMPLEMENTED, χ^2 OF THE AVERAGE DECREASES AS A FUNCTION OF AVERAGE χ^2 OF REPLICAS
- AT BEST FIT, AVERAGE χ^2 OF REPLICAS ~ 2; χ^2 OF AVERAGE TO DATA ~ 1

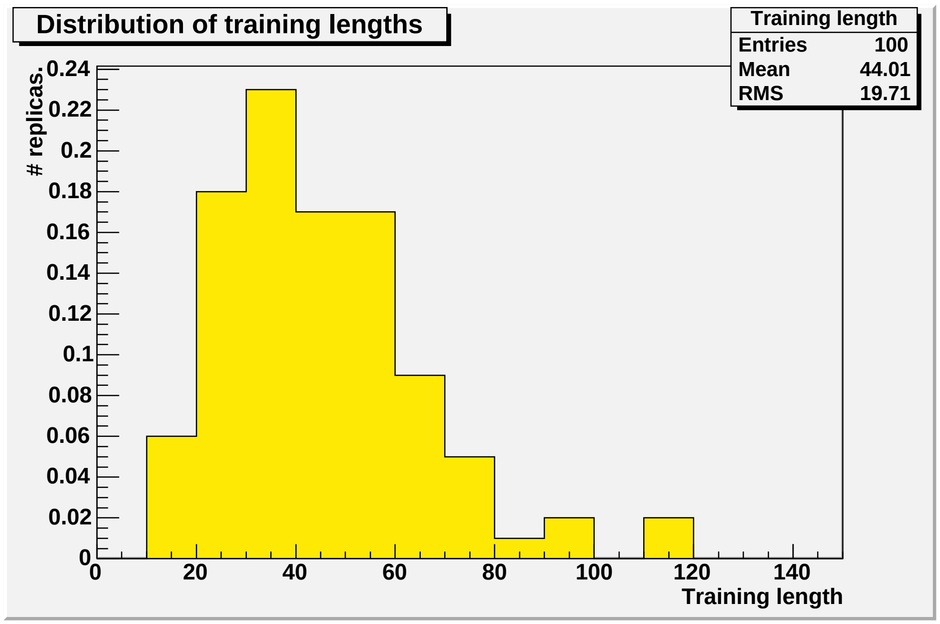
…AND STOPPING

AFTER STOPPING CRITERION IMPLEMENTED

DISTRIBUTION OF χ^2 AT STOPPING



DISTRIBUTION OF TRAINING LENGTHS



- POISSONIAN DISTRIBUTION OF TRAINING LENGTHS
- BEST FIT $\chi^2 = 0.75$ (BCDMS: 0.75, NMC: 0.72):
  EXPT. ERRORS SOMEWHAT OVERESTIMATED?

STABILITY

CAN CHECK STABILITY BY COMPARING RESULTS IF THE WHOLE PROCEDURE IS REPEATED WITH A DIFFERENT SET OF REPLICAS

DEFINE R.M.S. DISTANCE $\langle d[q]\rangle = \sqrt{\left\langle \frac{\left(\langle q_i\rangle_{(1)} - \langle q_i\rangle_{(2)}\right)^2}{\sigma^2[q_i^{(1)}] + \sigma^2[q_i^{(2)}]} \right\rangle_{\rm dat}}$

NOTE $\sigma \Rightarrow$ ERROR ON AVERAGE = (ERROR ON q_i)/$\sqrt{N}$

$\Rightarrow$ TESTS BOTH ACCURACY OF CENTRAL VALUE & ERRORS

SELF-STABILITY:
DIFFERENT SETS OF 100 REPLICAS

$\langle d\,[q]\rangle_{\rm dat}$	0.96
$\langle d\,[q]\rangle_{\rm extra}$	0.99
$\langle d\,[\sigma_q]\rangle_{\rm dat}$	0.88
$\langle d\,[\sigma_q]\rangle_{\rm extra}$	0.97

CHANGE OF ARCHITECTURE:
2-4-3-1 VS. 2-5-3-1

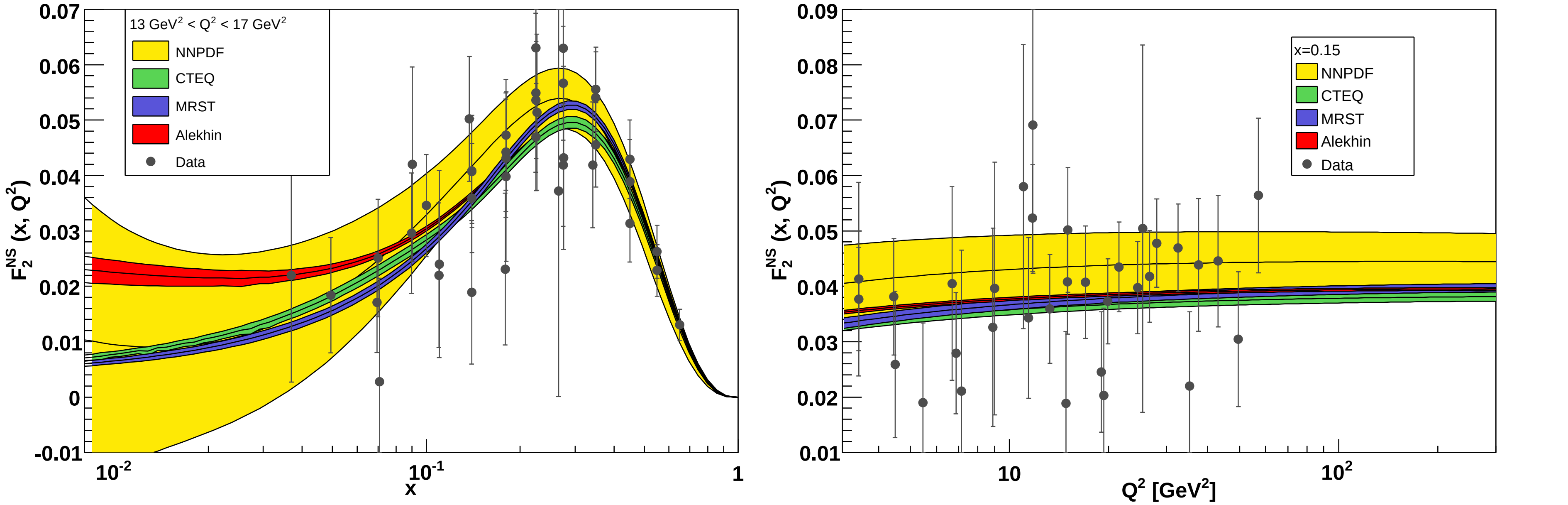
$\langle d\,[q]\rangle_{\rm dat}$	0.9
$\langle d\,[q]\rangle_{\rm extra}$	0.9
$\langle d\,[\sigma_q]\rangle_{\rm dat}$	0.9
$\langle d\,[\sigma_q]\rangle_{\rm extra}$	1.4

DISTANCE COMPUTED FOR 14 POINTS LINEARLY SPACED IN THE DATA REGION ($0.05 \leq x \leq 0.75$)
& 14 POINTS LOG SPACED IN THE EXTRAPOLATION REGION ($10^{-3} \leq x \leq 10^{-2}$)

RESULTS & COMPARISON TO OTHER APPROACHES

NLO RESULTS: THE STRUCTURE FUNCTION $F_2^{\mathrm{NS}}(x, Q^2)$

VS x AT $Q^2 = 15$ GEV^2 VS Q^2 AT $x = 0.15$

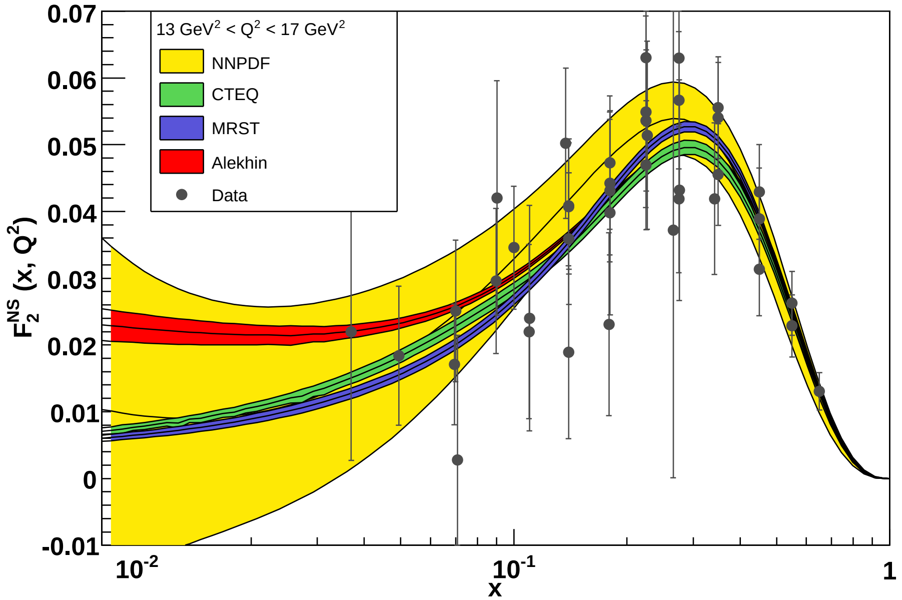


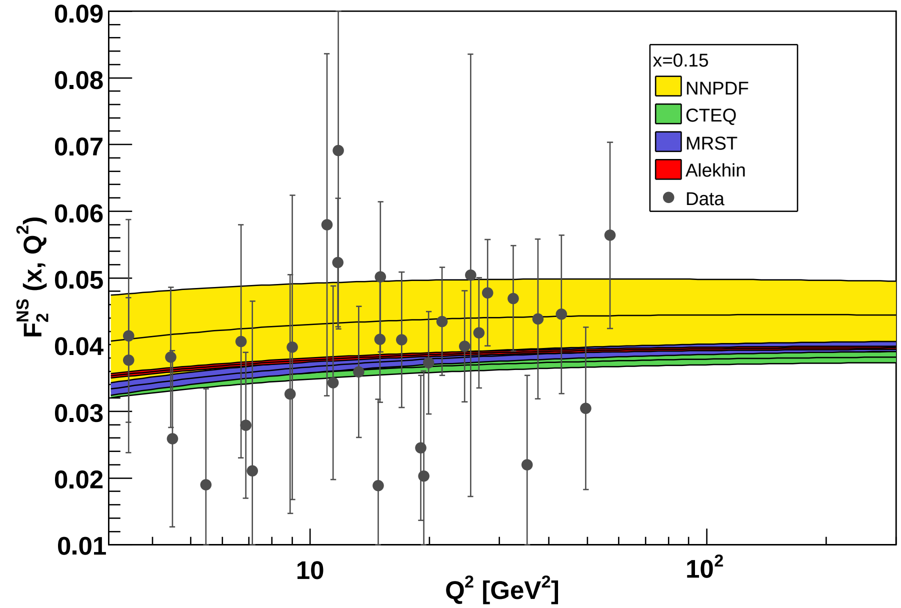
- COMPATIBLE WITH EXISTING FITS WITHIN ERROR
  (even when they disagee with each other)

RESULTS & COMPARISON TO OTHER APPROACHES

NLO RESULTS: THE STRUCTURE FUNCTION $F_2^{\mathrm{NS}}(x, Q^2)$

VS x AT $Q^2 = 15$ GEV^2 VS Q^2 AT $x = 0.15$



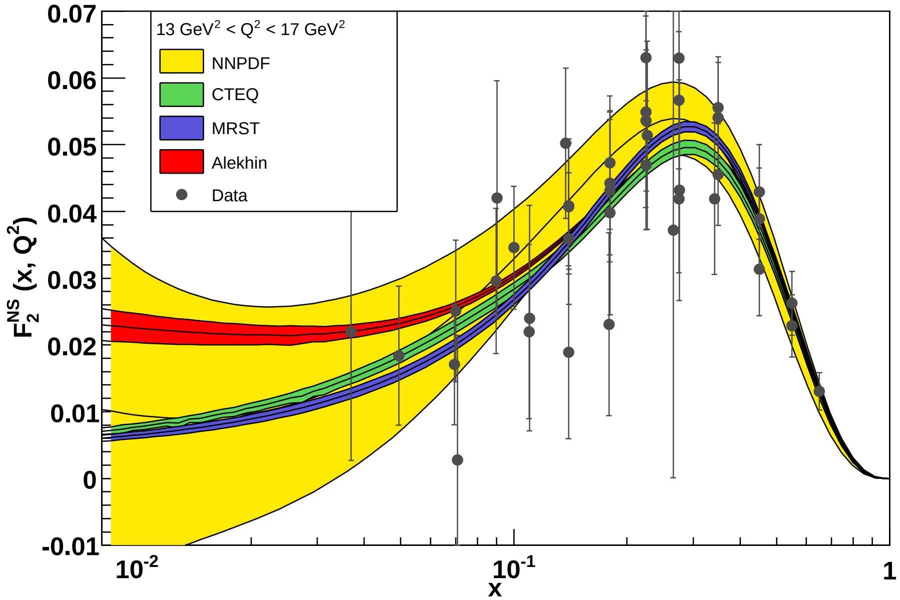


- COMPATIBLE WITH EXISTING FITS WITHIN ERROR
  (even when they disagee with each other)
- UNCERTAINTY MUCH LARGER IN EXTRAPOLATION BUT ALSO IN DATA REGION
  (note no other global fit data constrain q_{NS})

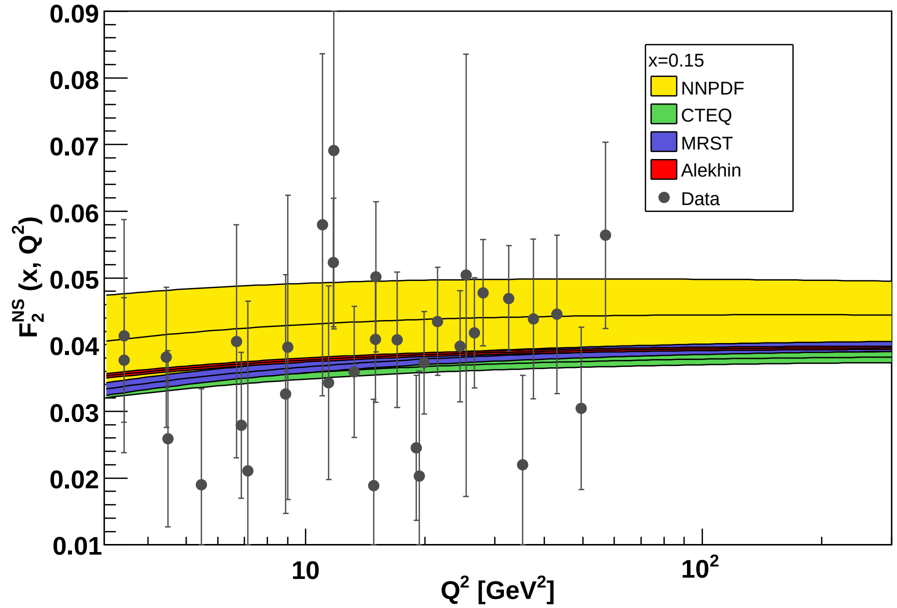
RESULTS & COMPARISON TO OTHER APPROACHES

NLO RESULTS: THE STRUCTURE FUNCTION $F_2^{\mathrm{NS}}(x, Q^2)$

VS x AT $Q^2 = 15$ GEV^2



VS Q^2 AT $x = 0.15$



- COMPATIBLE WITH EXISTING FITS WITHIN ERROR
  (even when they disagee with each other)

- UNCERTAINTY MUCH LARGER IN EXTRAPOLATION BUT ALSO IN DATA REGION
  (note no other global fit data constrain q_{NS})

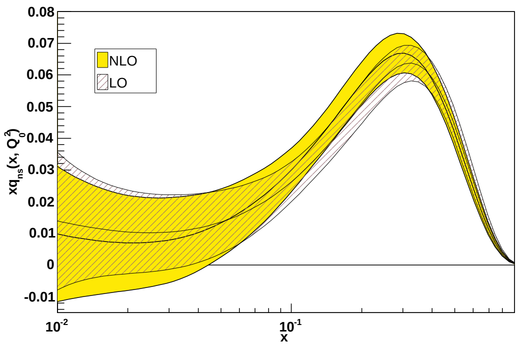
- CENTRAL FIT DISAGREES WITH EXISTING FITS IN VALENCE REGION
  $0.1 \le x \le 0.3$

RESULTS:

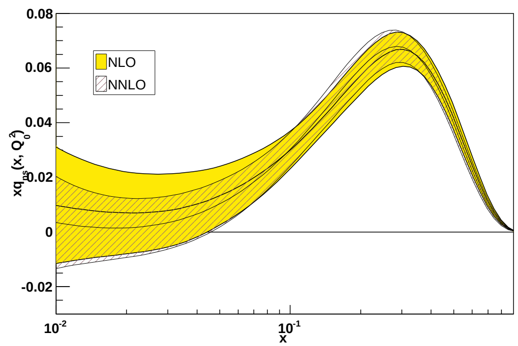
THE NONSINGLET QUARK PDF $q^{\mathrm{NS}}(x, Q^2)$

LO, NLO & NNLO

LO VS. NLO



NLO VS. NNLO



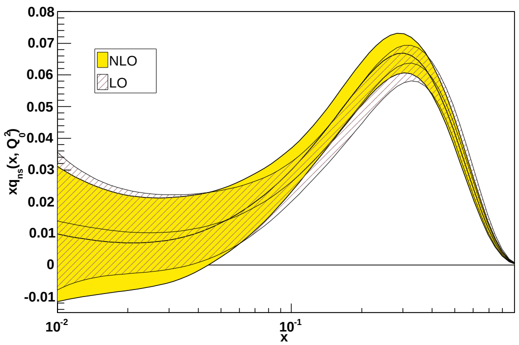
- quality of fit (χ^2) same at LO, NLO, NNLO
- NLO & NNLO agree within one σ
  NNLO terms negligible within errors
- LO & NLO agree within three σ
  NLO terms absorbed in b.c.

RESULTS:

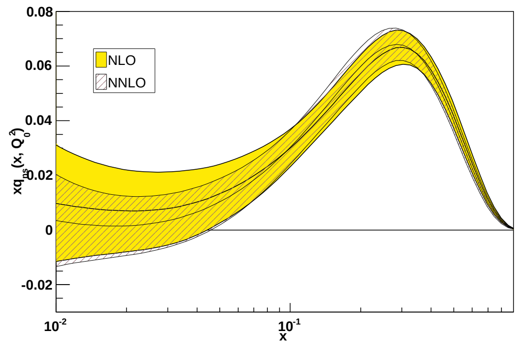
THE NONSINGLET QUARK PDF $q^{\mathrm{NS}}(x, Q^2)$

LO, NLO & NNLO

LO VS. NLO



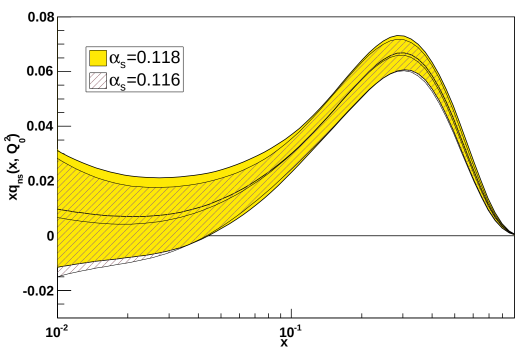
NLO VS. NNLO



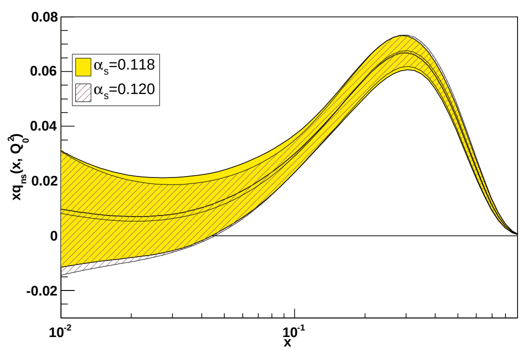
- quality of fit (χ^2) same at LO, NLO, NNLO
- NLO & NNLO agree within one σ
  NNLO terms negligible within errors
- LO & NLO agree within three σ
  NLO terms absorbed in b.c.

VARIATION OF α_s

LOW $\alpha_s = 0.116$



HIGH $\alpha_s = 0.120$

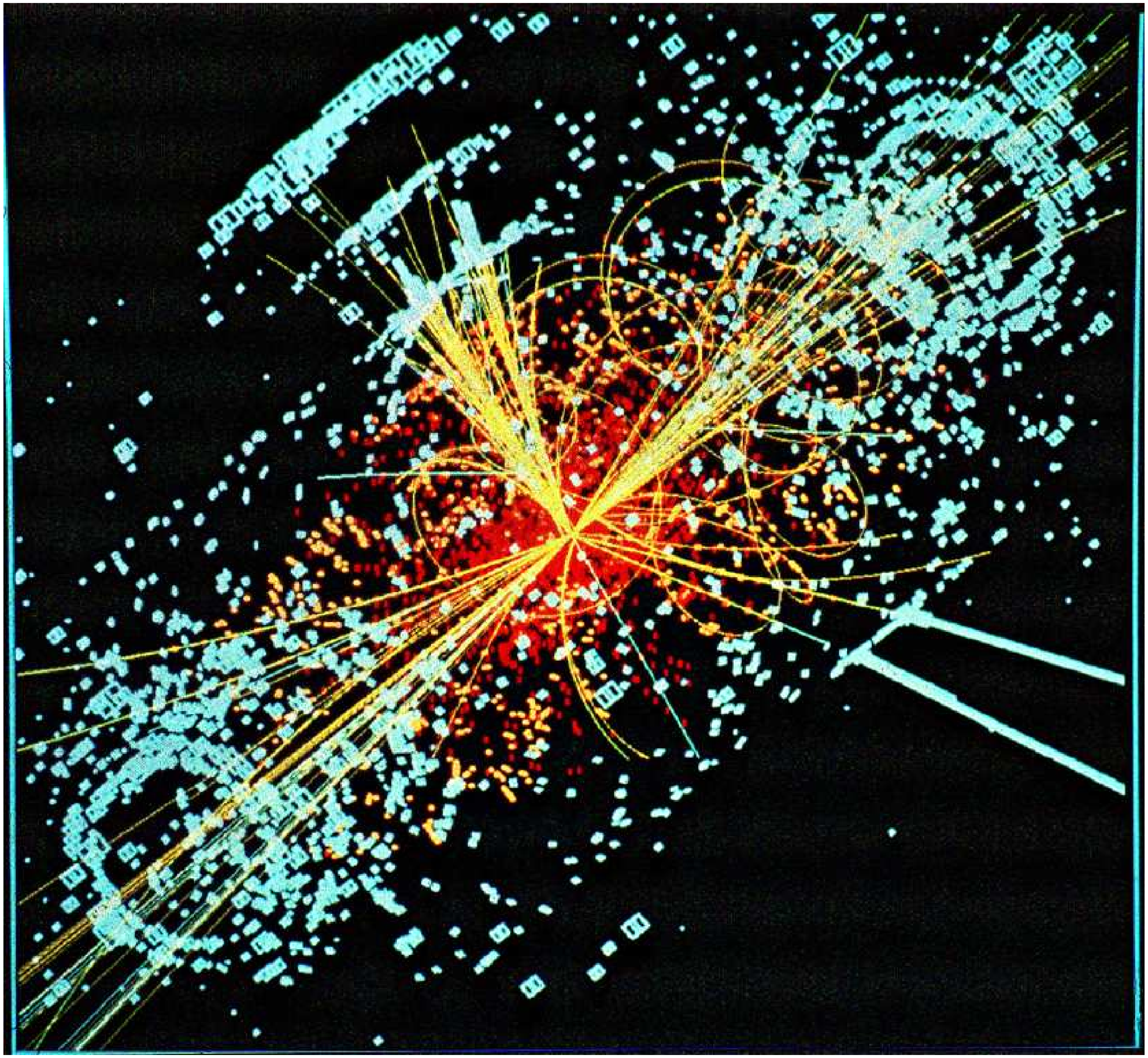


- quality of fit (χ^2) unchanged with $\alpha_s = 0.118 \pm 0.002$
- all fits agree within one σ
  $\Rightarrow$ α_s cannnot be determined with good accuracy

CONCLUSIONS

- STANDARD METHODS OF PDF DETERMINATION ARE STRETCHED TO THEIR LIMIT BY THE NEEDS OF PRECISION PHENOMENOLOGY AT THE LHC
- NEURAL PARTON DISTRIBUTIONS ARE

...BEHIND THE CORNER



Higgs decay in $e^+e^- + 2$ jets at CMS