NNPDF:
RESULTS AND DEPENDENCE ON PARAMETRIZATION

STEFANO FORTE
UNIVERSITÀ DI MILANO
FOR THE NNPDF COLLABORATION: R. D. BALL, L. DEL DEBBIO,
S.F., A. GUFFANTI, J. I. LATORRE, A. PICCIONE, J. ROJO, M. UBIALI

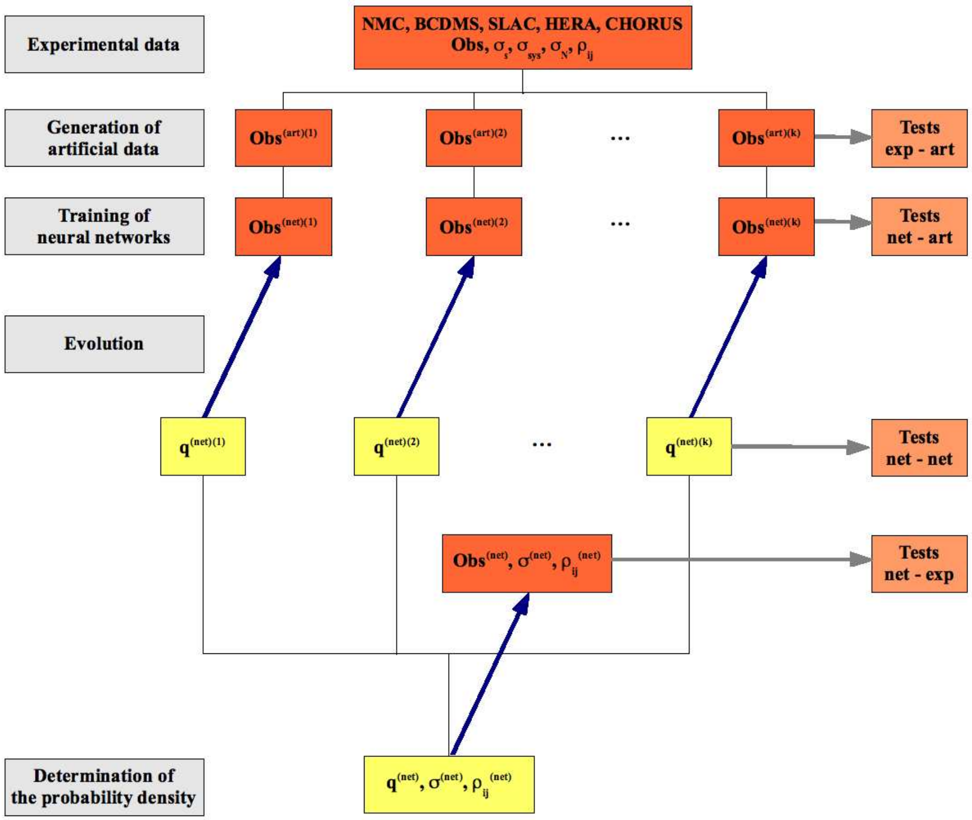
PDF4LHC WORKSHOP

CERN, JULY 14, 2008

THE NEURAL MONTE CARLO

- 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:

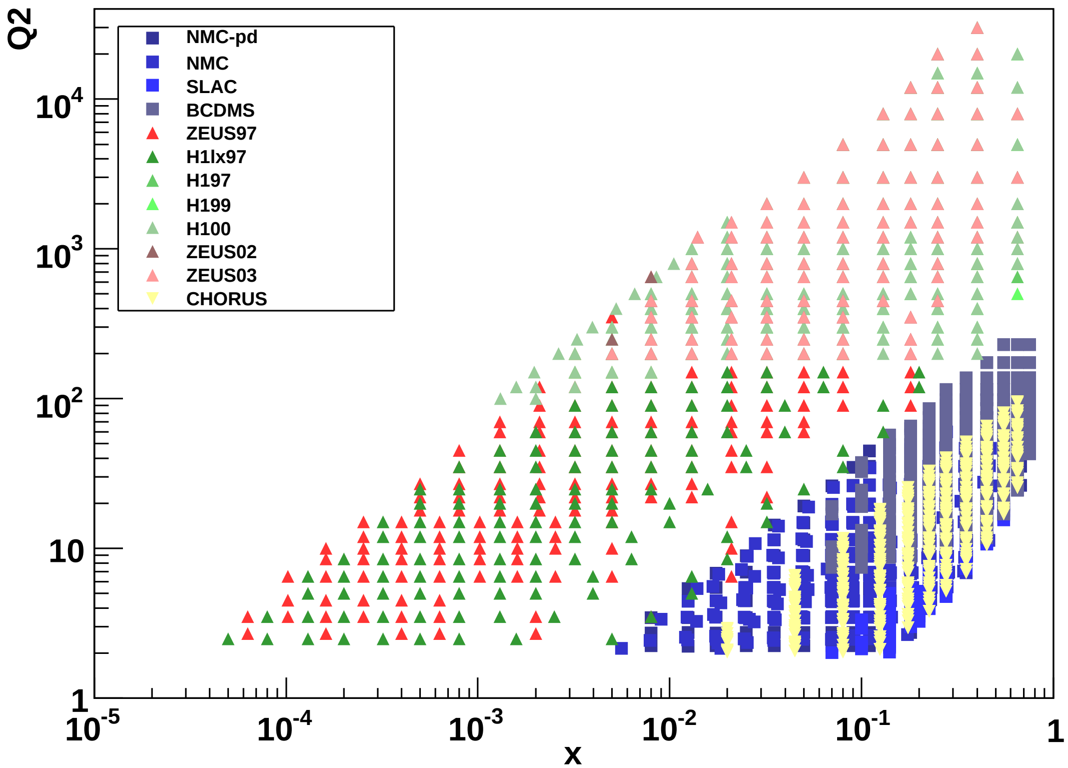
$$\left\langle \sigma\left[f_i\right] \right\rangle = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} \sigma\left[f_i^{(net)(k)}\right]$$



FEATURES OF THE FIT

THE DATASET

$Q^2 > 2\ \mathrm{GeV}^2;\ W^2 > 12.5\ \mathrm{GeV}^2$



Name	Data points	Target
NMC_pd	153	F_2^d / F_2^p
NMC	245	F_2^p
SLAC	47 (47)	$F_2^{p(d)}$
BCDMS	333 (248)	$F_2^{p(d)}$
ZEUS97	240 (29)	$\tilde{\sigma}^{NC(CC),+}$
ZEUS02	92 (26)	$\tilde{\sigma}^{NC(CC),-}$
ZEUS03	90 (30)	$\tilde{\sigma}^{NC(CC),+}$
H1lx97	135	$\tilde{\sigma}^{NC,+}$
H197	130 (25)	$\tilde{\sigma}^{NC(CC),+}$
H199	139 (28)	$\tilde{\sigma}^{NC(CC),-}$
H100	147 (28)	$\tilde{\sigma}^{NC(CC),+}$
H108	8	F_L
CHORUS	471 (471)	$\tilde{\sigma}^{\nu(\bar{\nu})}$
Total	3161	

FEATURES OF THE FIT

THEORY

- NLO EVOLUTION ($N-$ SPACE, EXPANDED)
- ZM-VFN SCHEME FOR THRESHOLDS
- $\alpha_s(M_z) = 0.119$, PDFS GIVEN AT $Q_0^2 = 2$ GEV^2
- TARGET-MASS CORRECTIONS INCLUDED UP TO TWIST FOUR

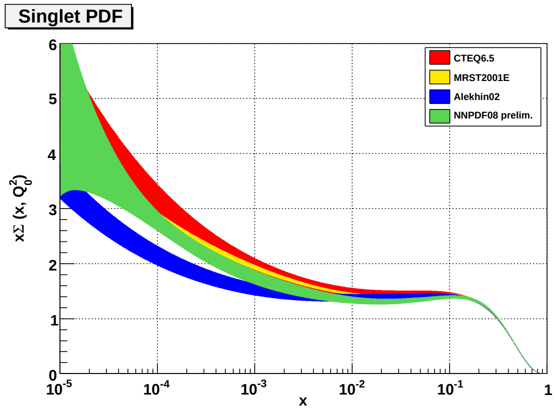
BASIS FUNCTIONS AND PARAMETRIZATION

- FIVE INDEPENDENT PDFS: SINGLET, GLUON, TOTAL VALENCE, TRIPLET, $\bar{d} - \bar{u}$.
- SYMMETRIC STRANGE SEA $s(x) = \bar{s}(x)$,
  PROPORTIONAL TO NON-STRANGE SEA, $\bar{s}(x) = \frac{C}{2}(\bar{u}(x) + \bar{d}(x))$, (C = 0.5)
- ALL PDFS PARAMETRIZED BY A 2-5-3-1 NEURAL NETWORK:
  $37 \times 5 = 185$ PARAMETERS
- MOMENTUM AND VALENCE SUM RULES ENFORCED STRICTLY
- POSITIVITY OF F_L ENFORCED for $x \geq 10^{-7}$, $Q^2 \geq 2$ GeV^2

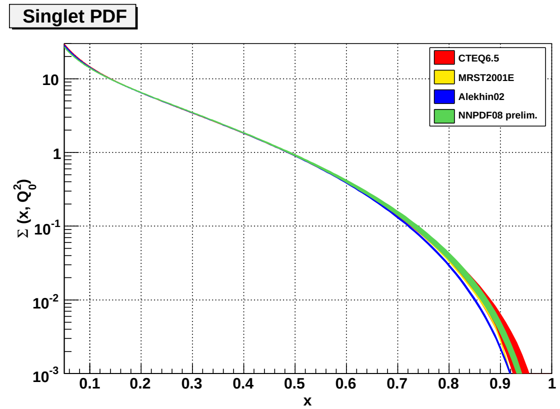
RESULTS

PARTON DISTRIBUTIONS

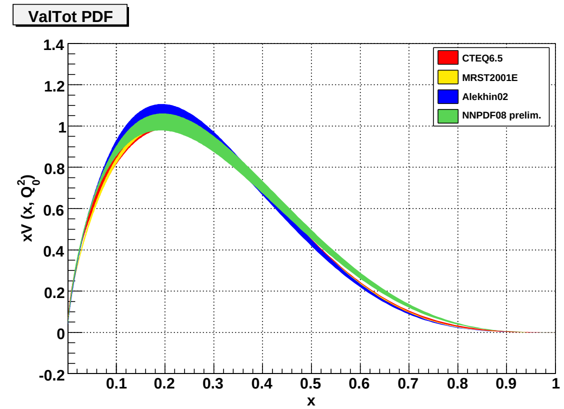
SINGLET



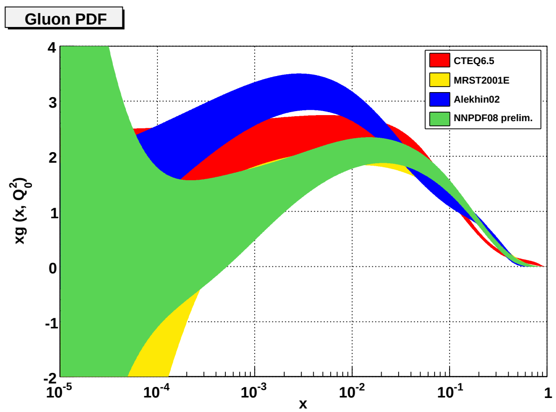
SINGLET



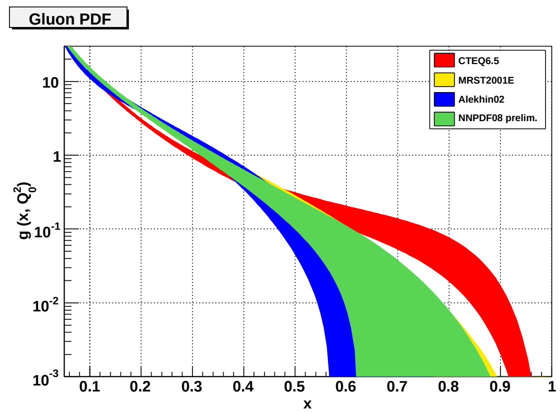
VALENCE



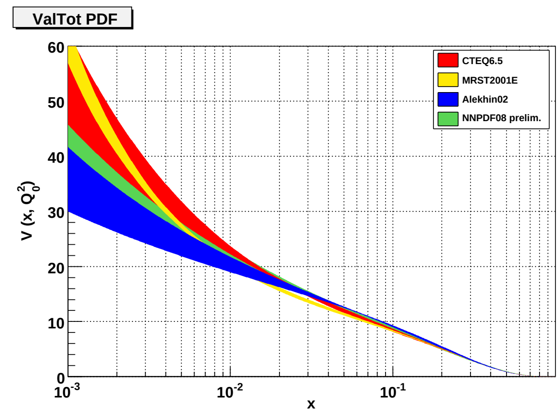
GLUON



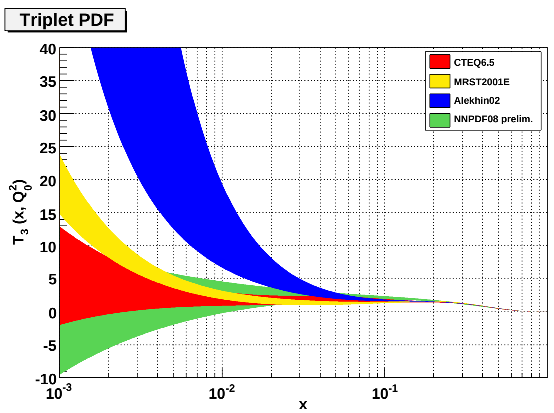
GLUON



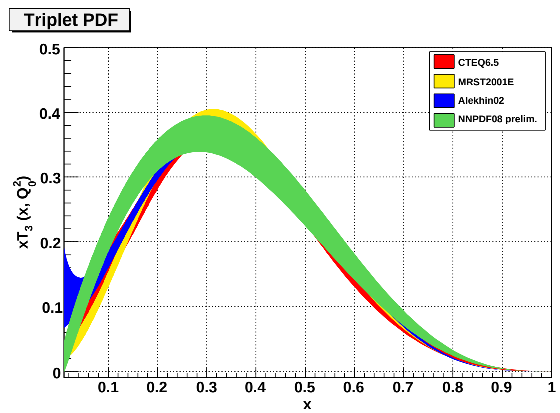
VALENCE



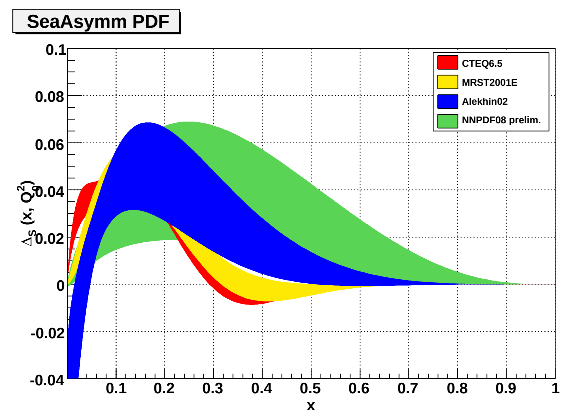
TRIPLET



TRIPLET



SEA ASYM.

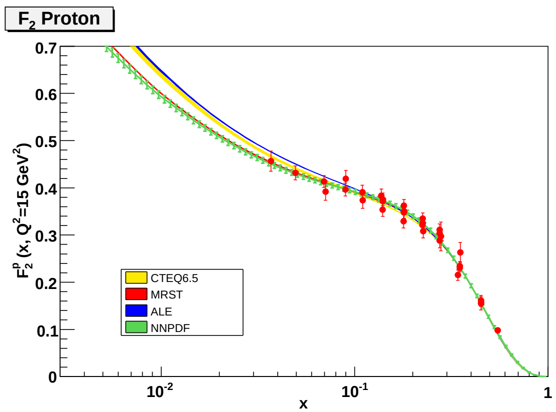


RESULTS

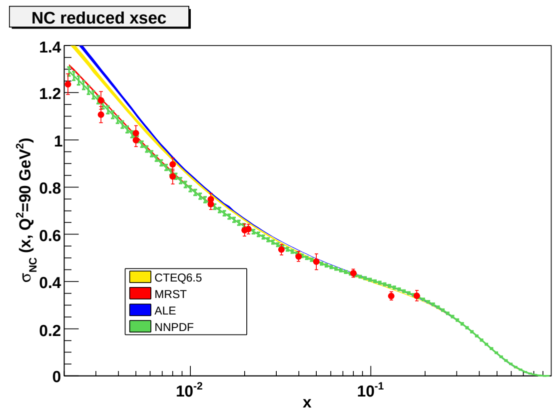
PHYSICAL OBSERVABLES

STRUCTURE FUNCTIONS

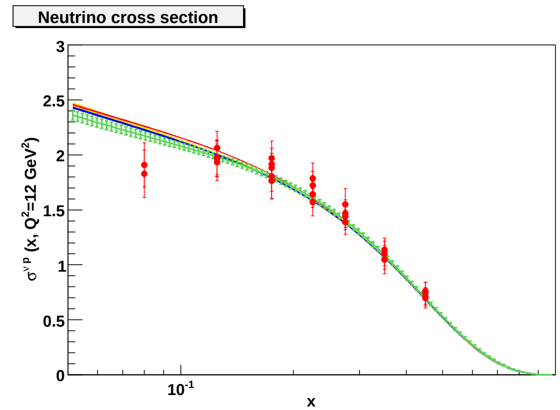
PROTON F_2



NC XSECT.

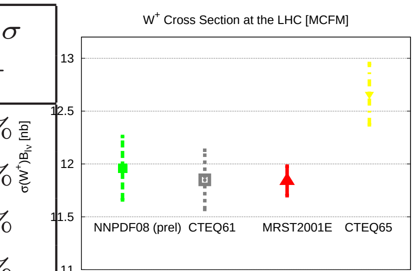


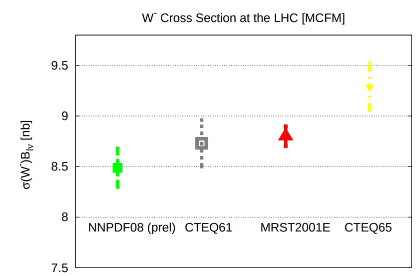
NEUTRINO XSECT.



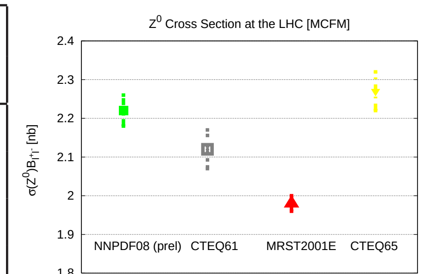
TOTAL CROSS-SECTIONS AT LHC, NLO FROM MCFM

-	$\sigma_{W^+}\mathcal{B}_{l^+\nu}$ [nb]	$\Delta\sigma/\sigma$ W^+	$\sigma_{W^-}\mathcal{B}_{l^-\nu}$ [nb]	$\Delta\sigma/\sigma$ W^-
NNPDF08	11.96 ± 0.30	2.5%	8.49 ± 0.19	2.3%
CTEQ6.5	12.66 ± 0.29	2.3%	9.29 ± 0.23	2.5%
CTEQ6.1	11.85 ± 0.28	2.4%	8.73 ± 0.23	2.6%
MRST01	11.84 ± 0.14	1.2%	8.80 ± 0.10	1.1%





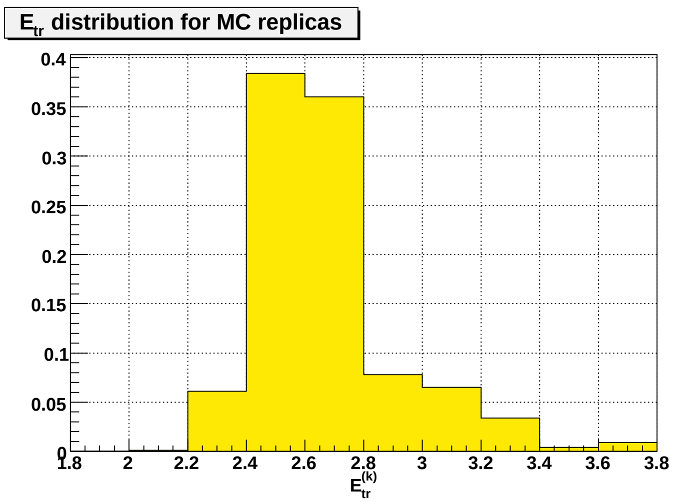
	$\sigma_Z\mathcal{B}_{l^+l^-}$ [nb]	$\Delta\sigma/\sigma$ Z	$\sigma_{t\bar{t}}$ [pb]	$\Delta\sigma/\sigma$ $t\bar{t}$	σ_H [pb]	$\Delta\sigma/\sigma$ H
NNPDF08	2.22 ± 0.04	2.0%	1014 ± 24	2.3%	35.79 ± 1.04	3.0%
CTEQ6.5	2.27 ± 0.05	2.2%	942 ± 19	2.0%	37.51 ± 0.80	2.2%
CTEQ6.1	2.12 ± 0.05	2.3%	970 ± 18	1.9%	38.50 ± 0.85	2.2%
MRST01	1.98 ± 0.02	1.0%	1013 ± 13	1.3%	37.52 ± 0.40	1.1%



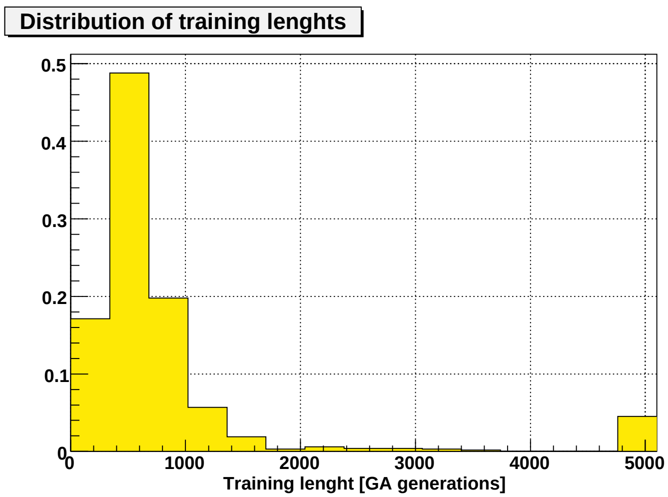
RESULTS

GENERAL STATISTICAL FEATURES

χ^2 AT STOPPING



TRAINING LENGTHS



$\chi^2_{\rm tot}$	1.34
$\langle E \rangle$	2.71
$\langle E_{\rm tr} \rangle$	2.68
$\langle E_{\rm val} \rangle$	2.72
$\langle {\rm TL} \rangle$	824.23
$N_{\rm rep}$	1000

- POISSONIAN DISTRIBUTION OF TRAINING LENGTHS
- BEST FIT $\chi^2 = 1.34$: MINOR DATA INCOMPATIBILITIES (?)

PARAMETRIZATION INDEPENDENCE:

METHODOLOGY

- EFFECTIVELY 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")

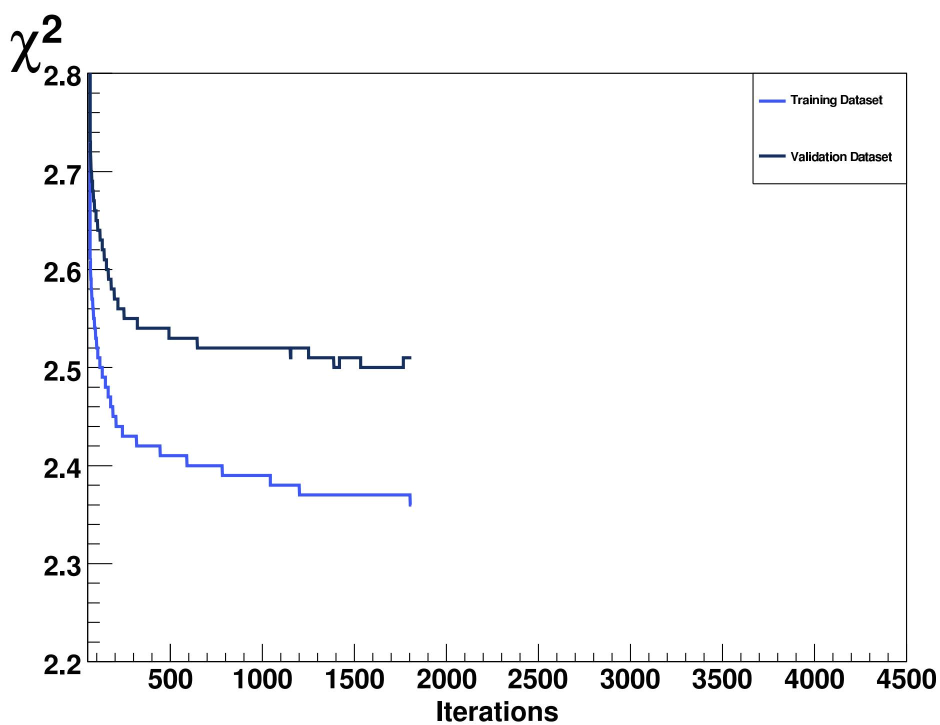
PARAMETRIZATION INDEPENDENCE:

REDUNDANCY AND OVERLEARNING

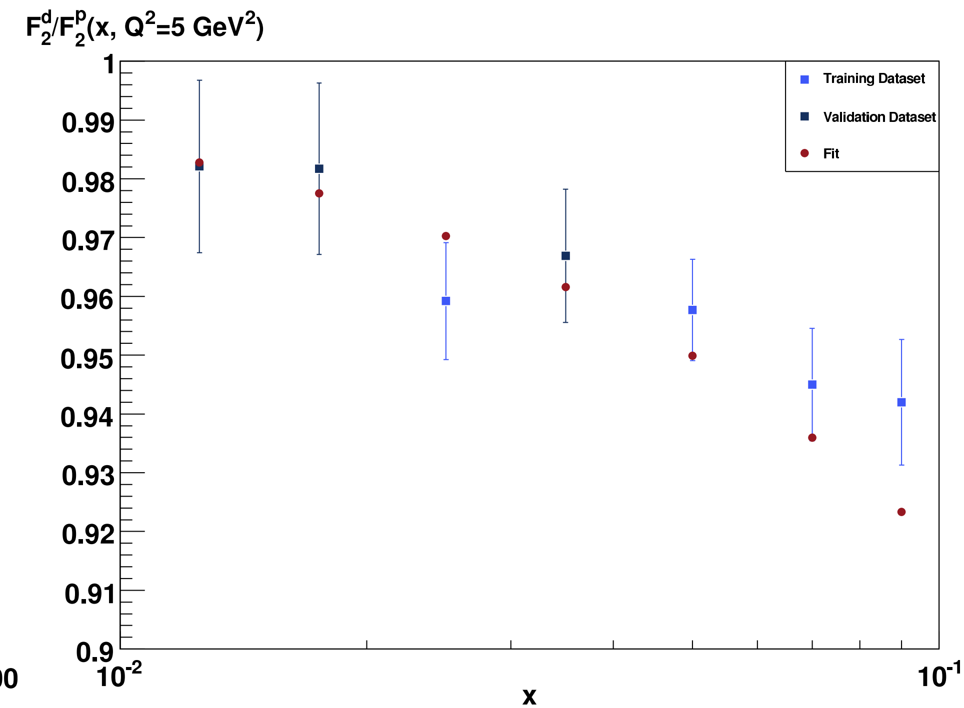
- OPTIMAL FIT OBTAINED WHEN QUALITY OF FIT TO VALIDATION (CONTROL) DATA STOPS IMPROVING
- POSSIBILITY OF OVERFITTING GUARANTESS THAT MINIMUM NOT DRIVEN BY PARAMETRIZATION

OPTIMAL FITTING

χ^2



FIT TO DATA



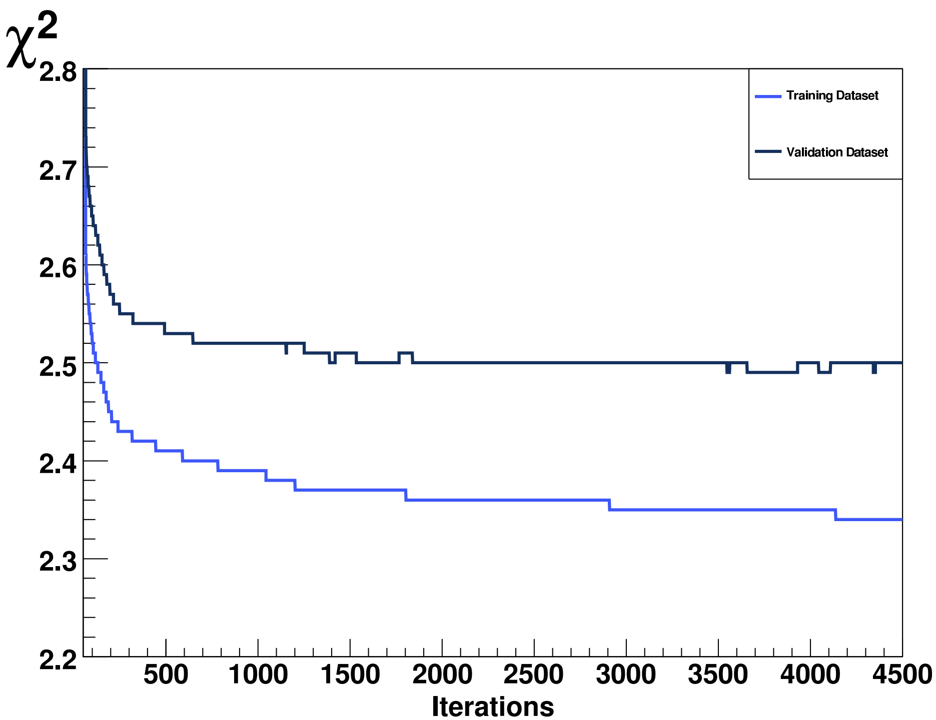
PARAMETRIZATION INDEPENDENCE:

REDUNDANCY AND OVERLEARNING

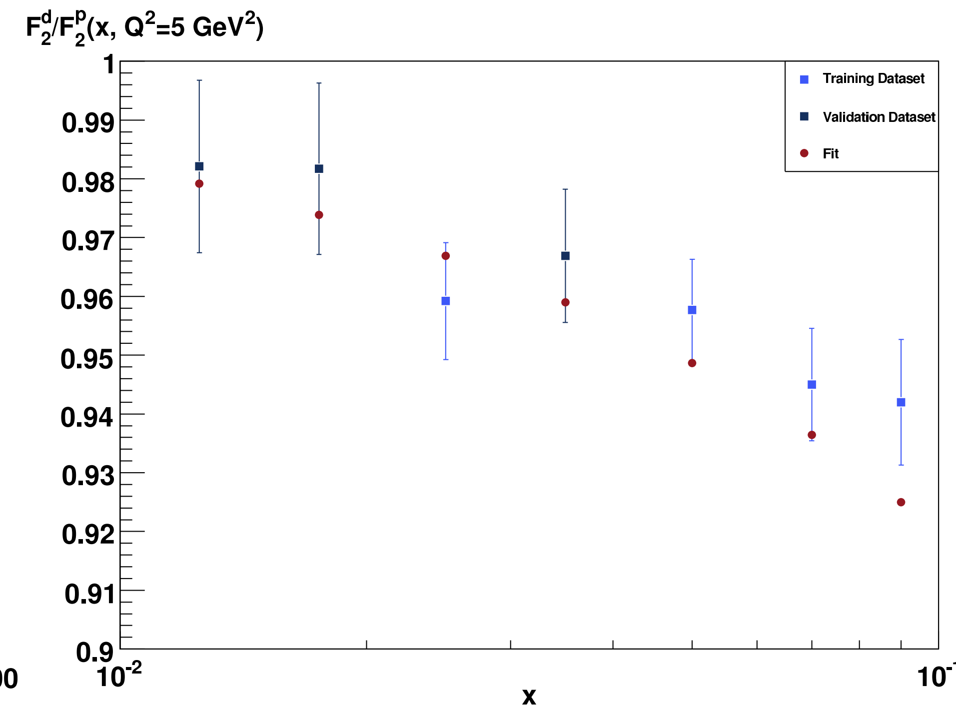
- OPTIMAL FIT OBTAINED WHEN QUALITY OF FIT TO VALIDATION (CONTROL) DATA STOPS IMPROVING
- POSSIBILITY OF OVERFITTING GUARANTESS THAT MINIMUM NOT DRIVEN BY PARAMETRIZATION

OVERFITTING

χ^2



FIT TO DATA



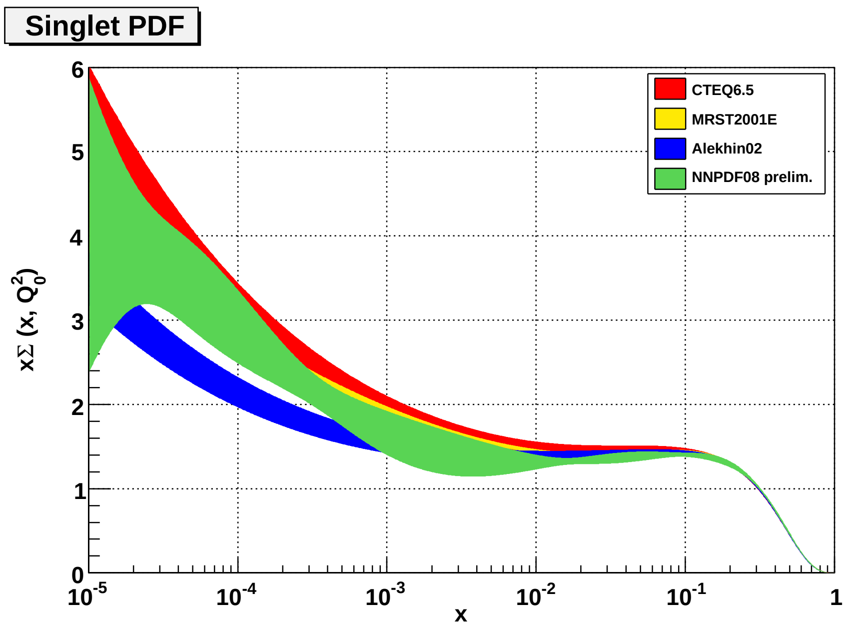
PARAMETRIZATION INDEPENDENCE:

LACK OF BIAS

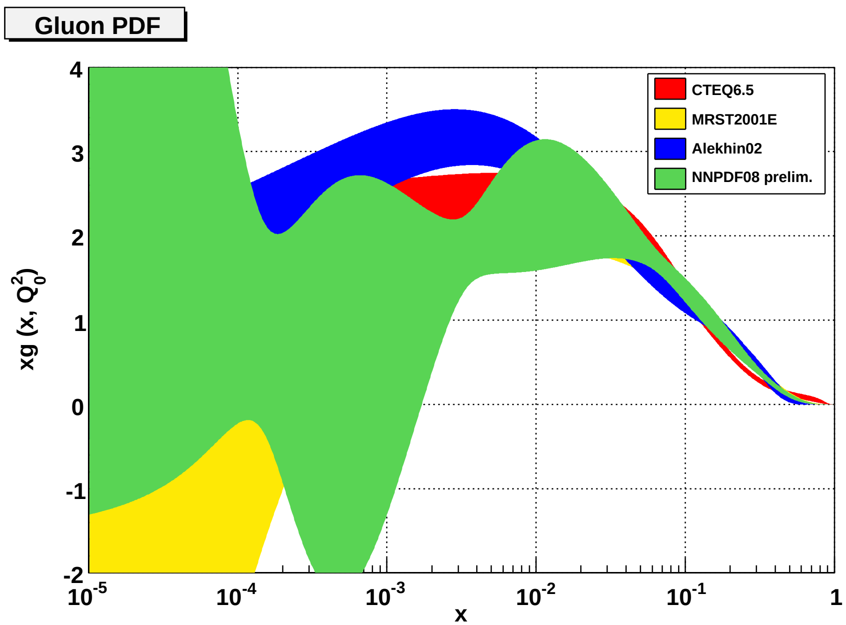
- IRREGULAR OR KNOTTY SHAPES ALLOWED IF DATA FLUCTUATE
- STATISTICS SHOW WHETHER THE EFFECT IS REAL

10 REPLICAS

SINGLET



GLUON



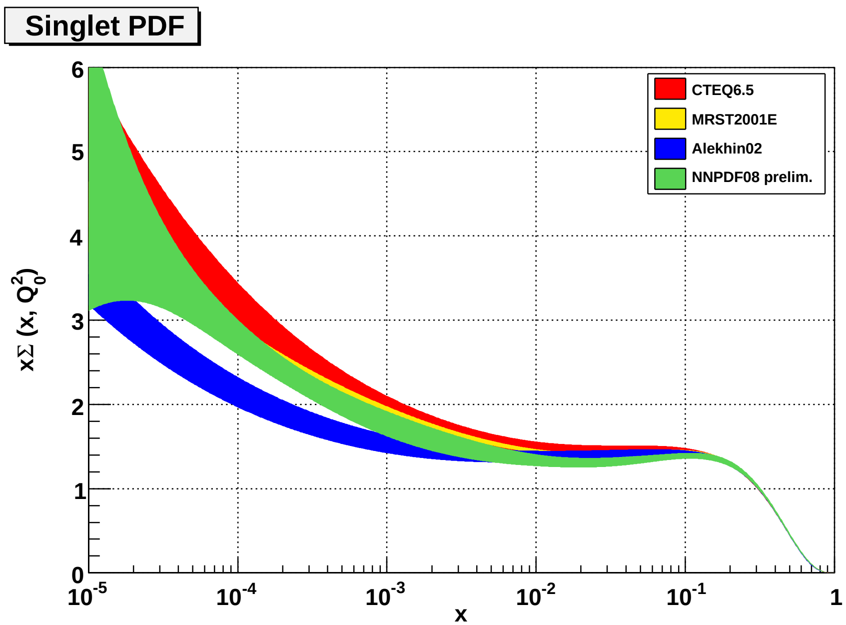
PARAMETRIZATION INDEPENDENCE:

LACK OF BIAS

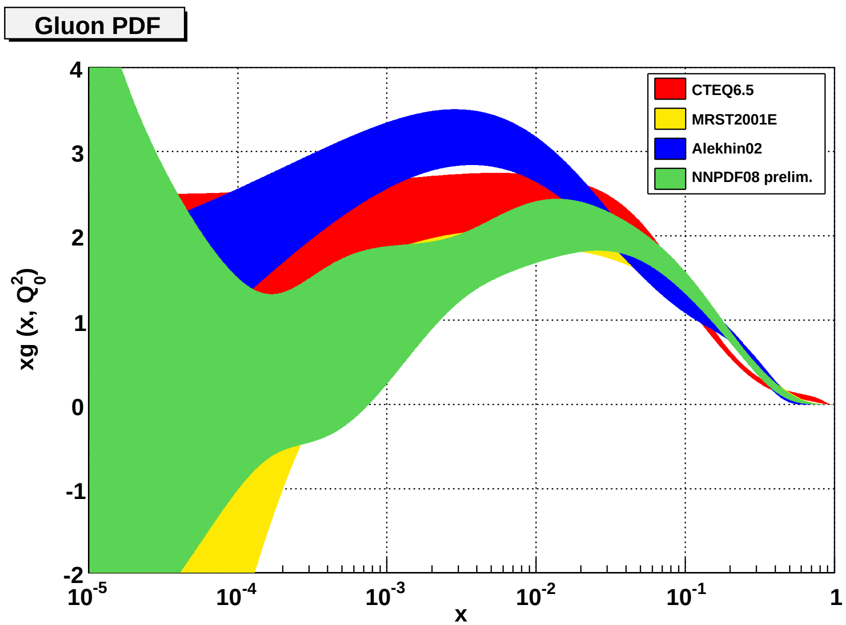
- IRREGULAR OR KNOTTY SHAPES ALLOWED IF DATA FLUCTUATE
- STATISTICS SHOW WHETHER THE EFFECT IS REAL

100 REPLICAS

SINGLET



GLUON



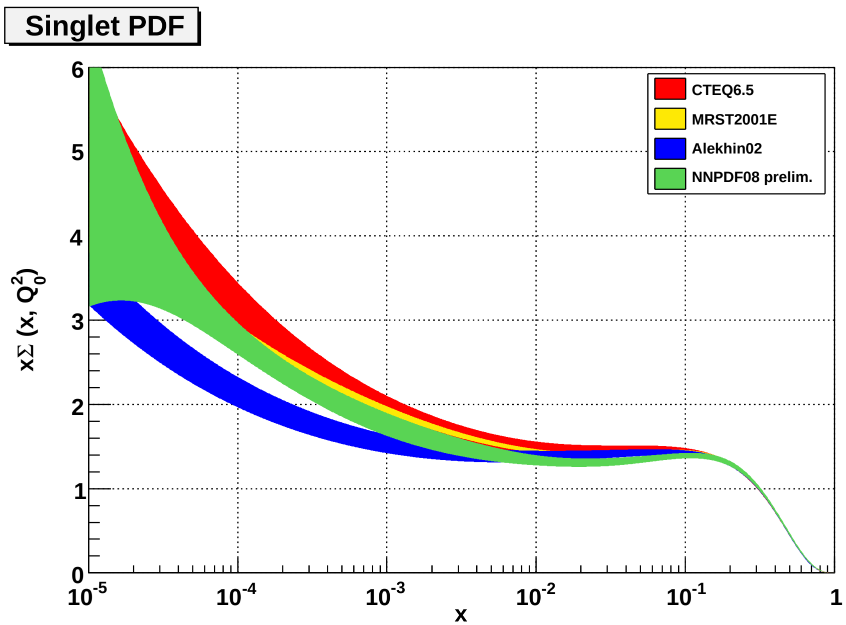
PARAMETRIZATION INDEPENDENCE:

LACK OF BIAS

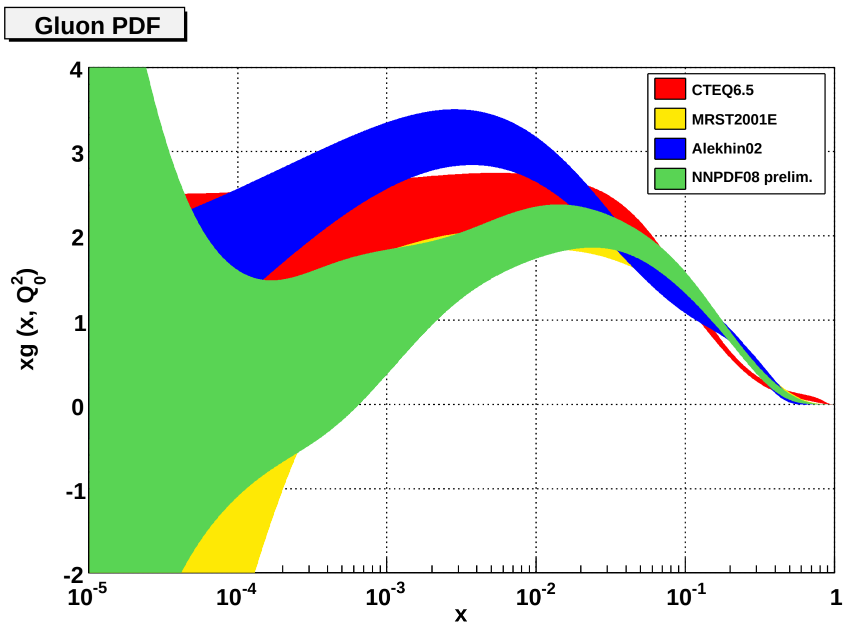
- IRREGULAR OR KNOTTY SHAPES ALLOWED IF DATA FLUCTUATE
- STATISTICS SHOW WHETHER THE EFFECT IS REAL

200 REPLICAS

SINGLET



GLUON



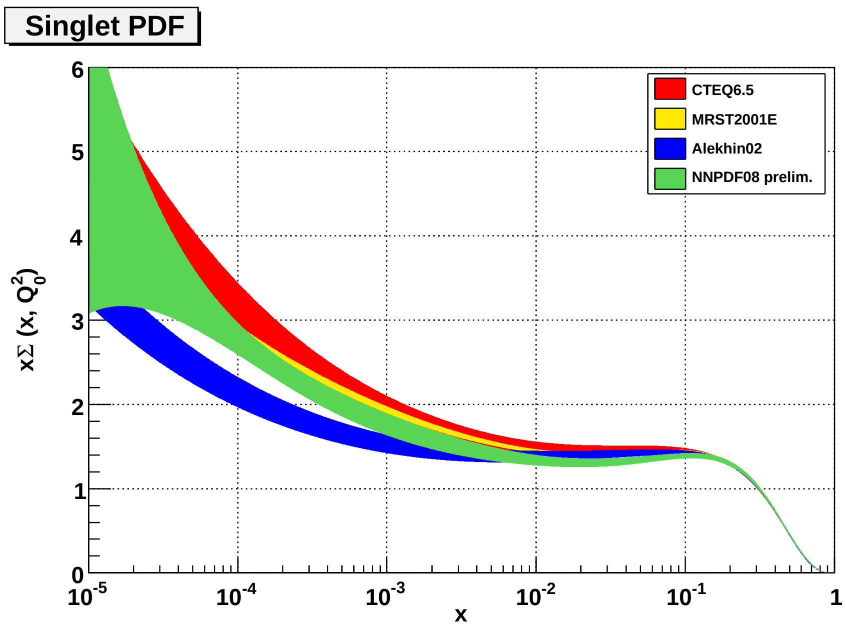
PARAMETRIZATION INDEPENDENCE:

LACK OF BIAS

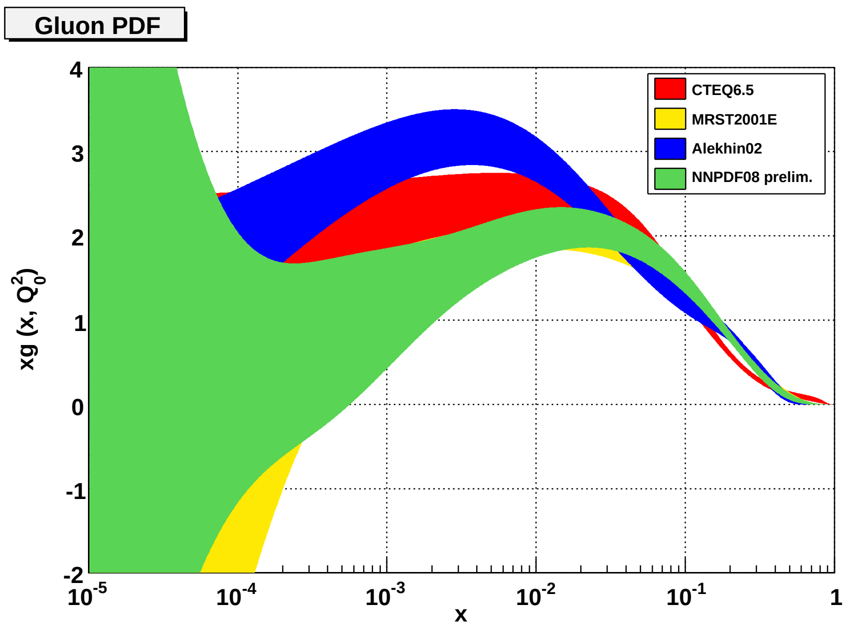
- IRREGULAR OR KNOTTY SHAPES ALLOWED IF DATA FLUCTUATE
- STATISTICS SHOW WHETHER THE EFFECT IS REAL

400 REPLICAS

SINGLET



GLUON



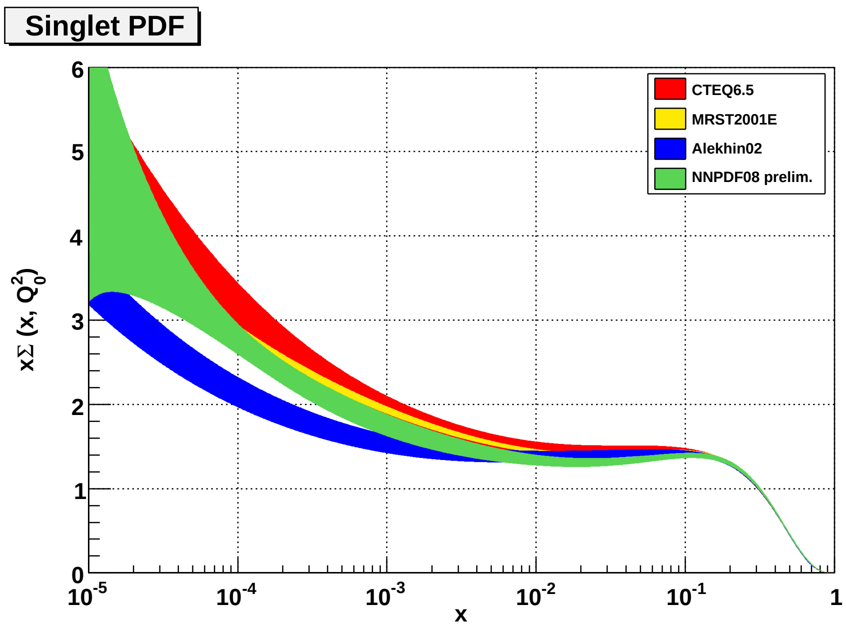
PARAMETRIZATION INDEPENDENCE:

LACK OF BIAS

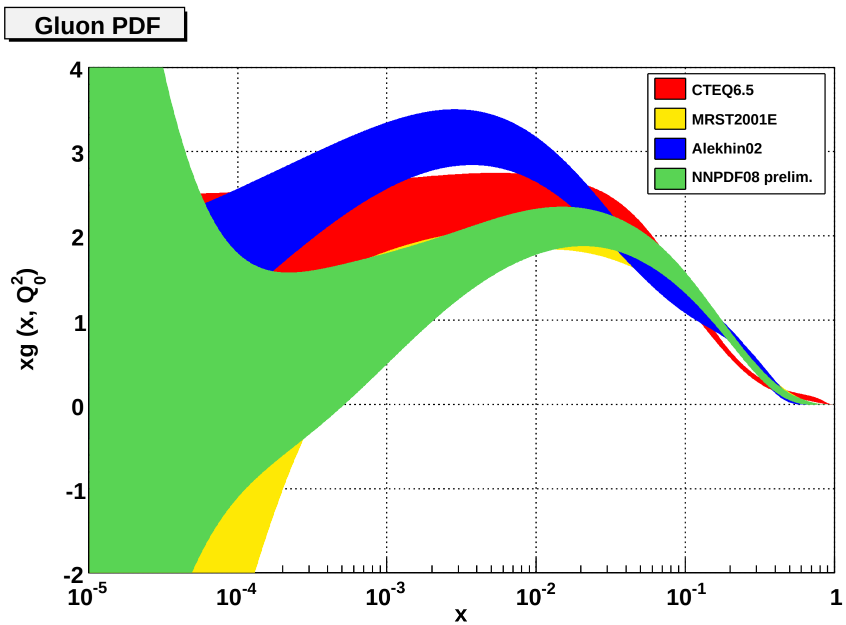
- IRREGULAR OR KNOTTY SHAPES ALLOWED IF DATA FLUCTUATE
- STATISTICS SHOW WHETHER THE EFFECT IS REAL

1000 REPLICAS

SINGLET



GLUON



PARAMETRIZATION INDEPENDENCE:

STATISTICAL STABILITY

COMPARE DISTANCE IN UNITS OF SIGMA OF RESULTS OBTAINED WITH DIFFERENT ASSUMPTIONS

- DISTANCE IN UNITS OF SIGMA

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

- NOTE σ $\Rightarrow$ ERROR ON AVERAGE
  = (ERROR ON q_i)/$\sqrt{N}$
  with 100 replicas, $d = 1$
  $\longrightarrow$ fits differ by 1/10 of nominal error

- TEST PREDICTIONS
  FOR CENTRAL VALUES & ERRORS

DISTANCE BETWEEN STANDARD & FIT WITH SMALLER NEURAL NETS

2-4-3-1 VS 2-5-3-1 ARCHITECTURE

(31 vs. 37 parms per net)

	DATA	EXTRAPOLATION
SINGLET	$0.005 \leq x \leq 0.1$	$10^{-4} \leq x \leq 10^{-3}$
$\langle d[q]\rangle$	0.96	1.32
$\langle d[\sigma]\rangle$	1.23	1.32
GLUON	$0.005 \leq x \leq 0.1$	$10^{-4} \leq x \leq 10^{-3}$
$\langle d[q]\rangle$	1.40	1.13
$\langle d[\sigma]\rangle$	1.17	1.06
VALENCE	$0.1 \leq x \leq 0.6$	$0.03 \leq x \leq 0.3$
$\langle d[q]\rangle$	1.40	0.93
$\langle d[\sigma]\rangle$	1.09	0.96
TRIPLET	$0.05 \leq x \leq 0.75$	$0.01 \leq x \leq 0.1$
$\langle d[q]\rangle$	1.05	1.09
$\langle d[\sigma]\rangle$	1.68	2.5

PARAMETRIZATION INDEPENDENCE:
THE "HERALHC BENCHMARK"

$Q^2 > 9\ \mathrm{GeV}^2;\ W^2 > 15\ \mathrm{GeV}^2$

REDUCED DATASET ⇒ WIDER ERROR BAND

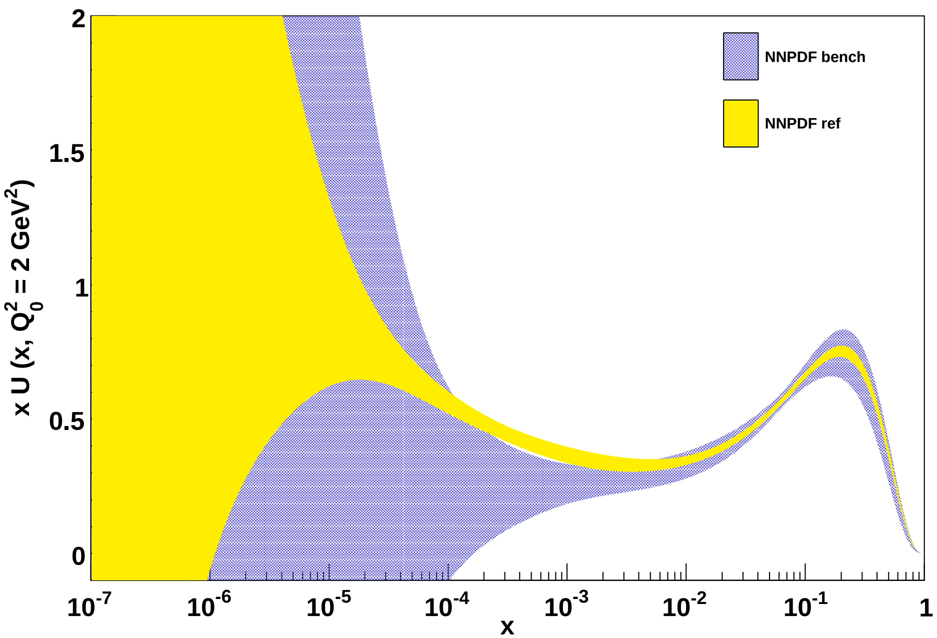
from 3161 to 773 datapoints

reduced info on small x sea (no low Q^2 data)

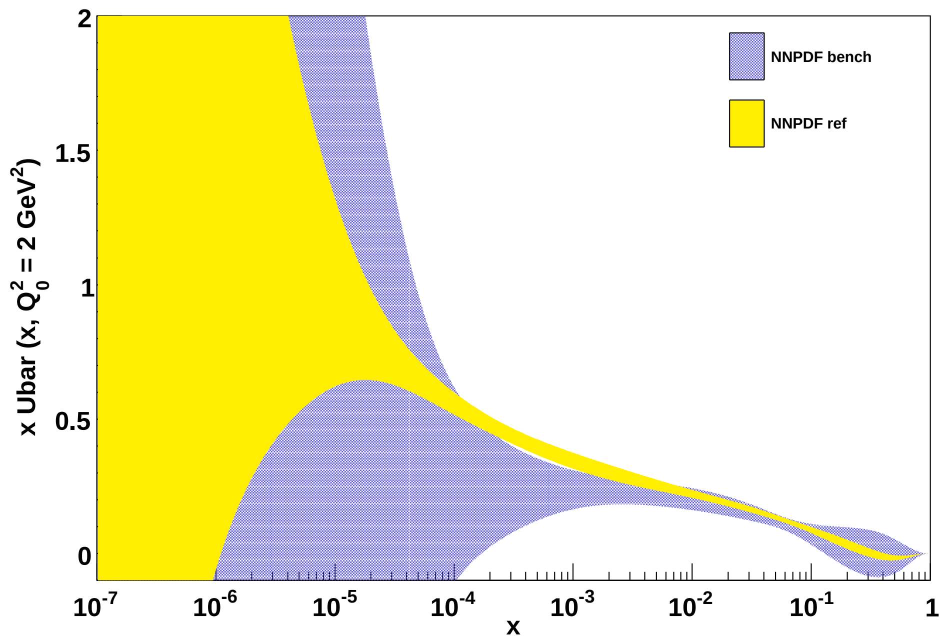
& large x valence (no neutrino data)

NAME	DATA POINTS	TARGET
NMC_PD	73	F_2^d / F_2^p
NMC	95	F_2^p
BCDMS	322	F_2^p
ZEUS97	206	F_2^p
H1LX97	77	F_2^p
TOTAL	773	

UP QUARK



UP ANTIQUARK



RESULTS COMPATIBLE TO WITHIN LESS THAN TWO SIGMA

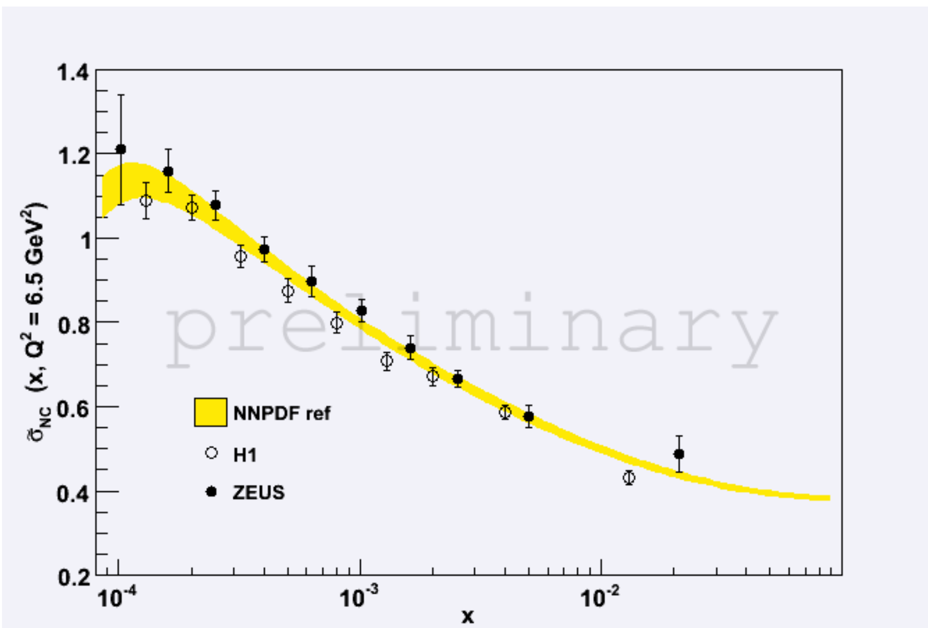
PARAMETRIZATION INDEPENDENCE:

THE "HERALHC BENCHMARK":INCOMPATIBLE DATA

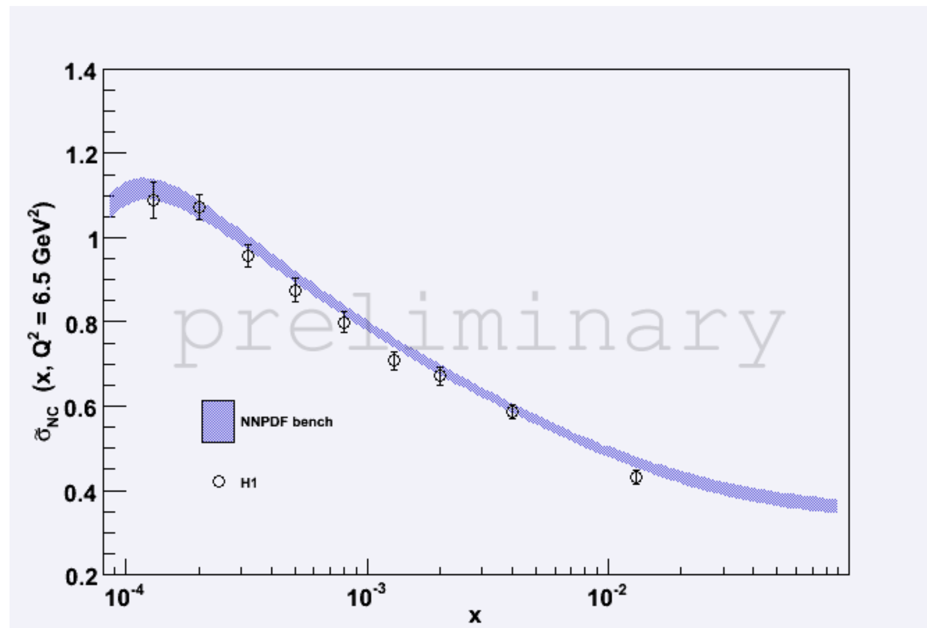
THE SMALL x CC REDUCED CROSS SECTION

results presented at the HERALHC workshop (NNPDF preliminary fit)

FULL DATASET



REDUCED "BENCHMARK" DATASET



NO ERROR REDUCTION WHEN DATA IN WIDER DATA SET ARE INCOMPATIBLE

DELIVERY:

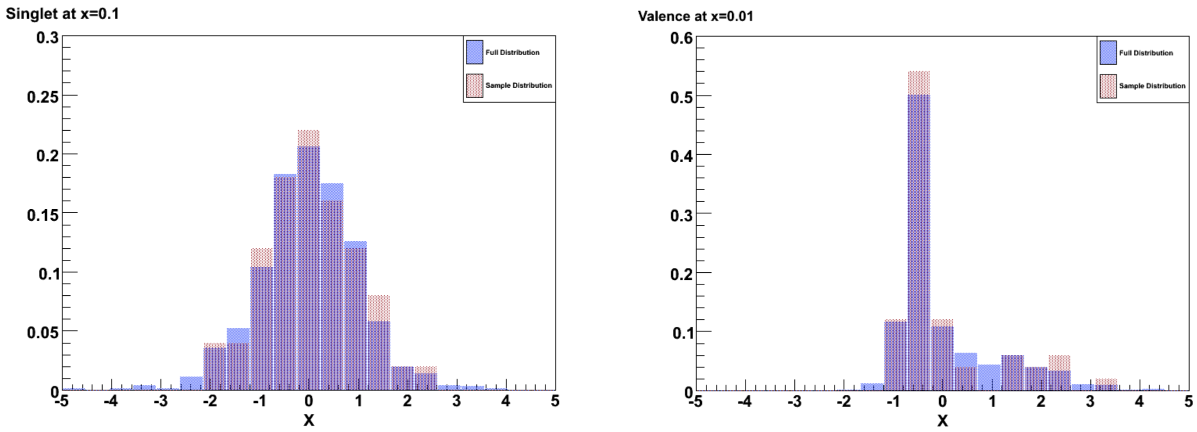
RESTRICTED SAMPLE OF REPLICAS

- WIDE SAMPLE OF PSEUDODATA ENDURES NO BIAS
- IMPRACTICAL TO AVERAGE OVER THOUSAND(S) OF REPLICAS
- SELECT SUBSET OF REPLICAS WITH APPROXIMATELY SAME STATISTICAL DISTRIBUTION AS FULL SET

- construct histogram for # of replicas n sigma away from mean
- compare result for subset & full
- minimize relative entropy of two histograms $S =$ $= \sum_{\text{i bins}} \left(p_i^{(1)} - p_i^{(2)}\right) \ln \frac{p_i^{(1)}}{p_i^{(2)}}$
- select with genetic algorithm subset which minimizes S

FULL (1000) VS. REDUCED (50) PROBABILITY HISTOGRAMS

SINGLET AT $x = 0.1$ VALENCE AT $x = 0.01$



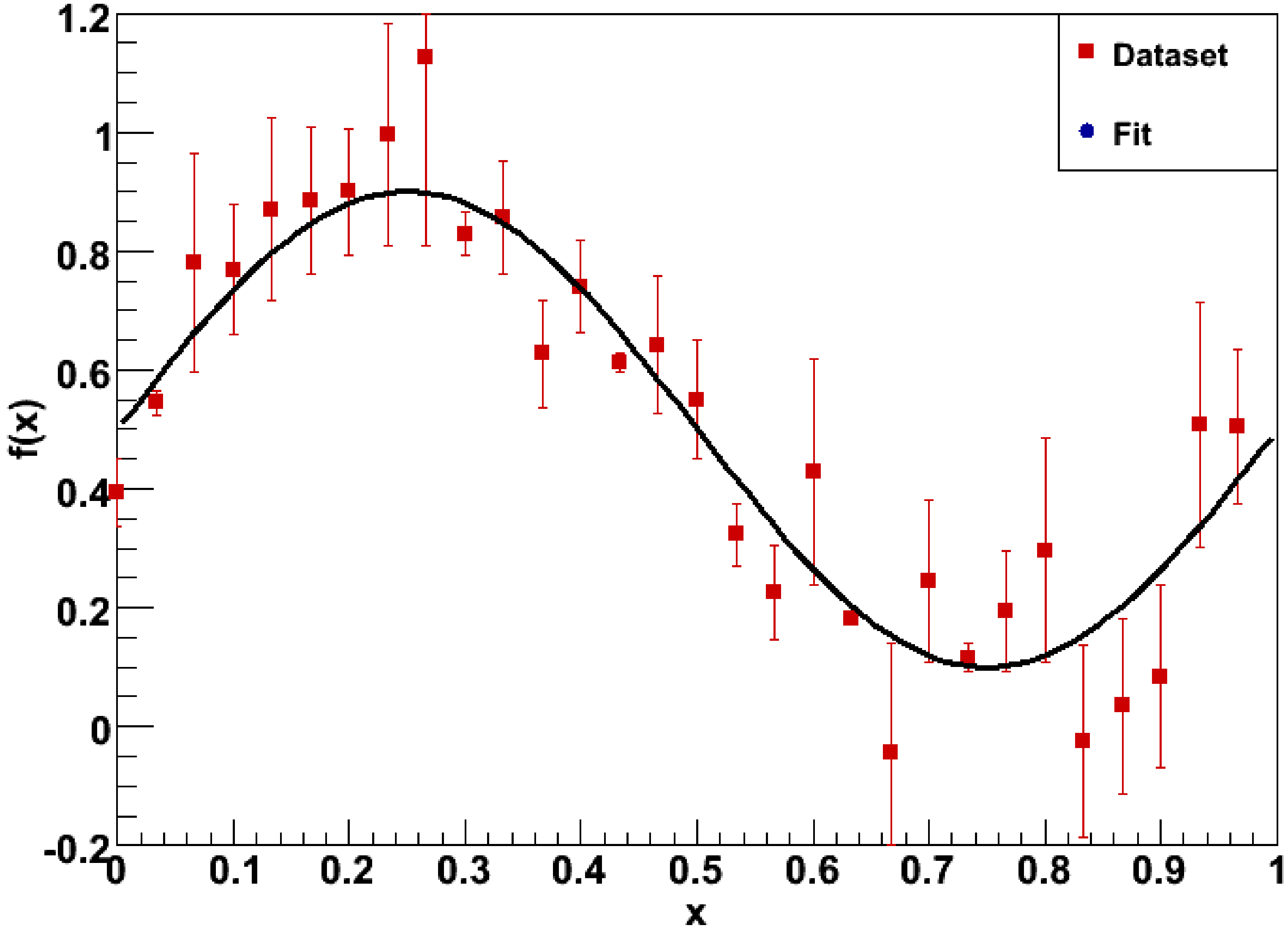
OUTLOOK

- FIT TO FULL DIS DATASET
  WITH UNBIASED ERROR ESTIMATE
- ALREADY INTERFACED TO LHAPDF
- FULL STATISTICAL FEATURES BASED ON SET OF
  1000 REPLICA PDFS,
  RESTRICTED SET OF 40 PDFS AVAILABLE SOON

EXTRAS

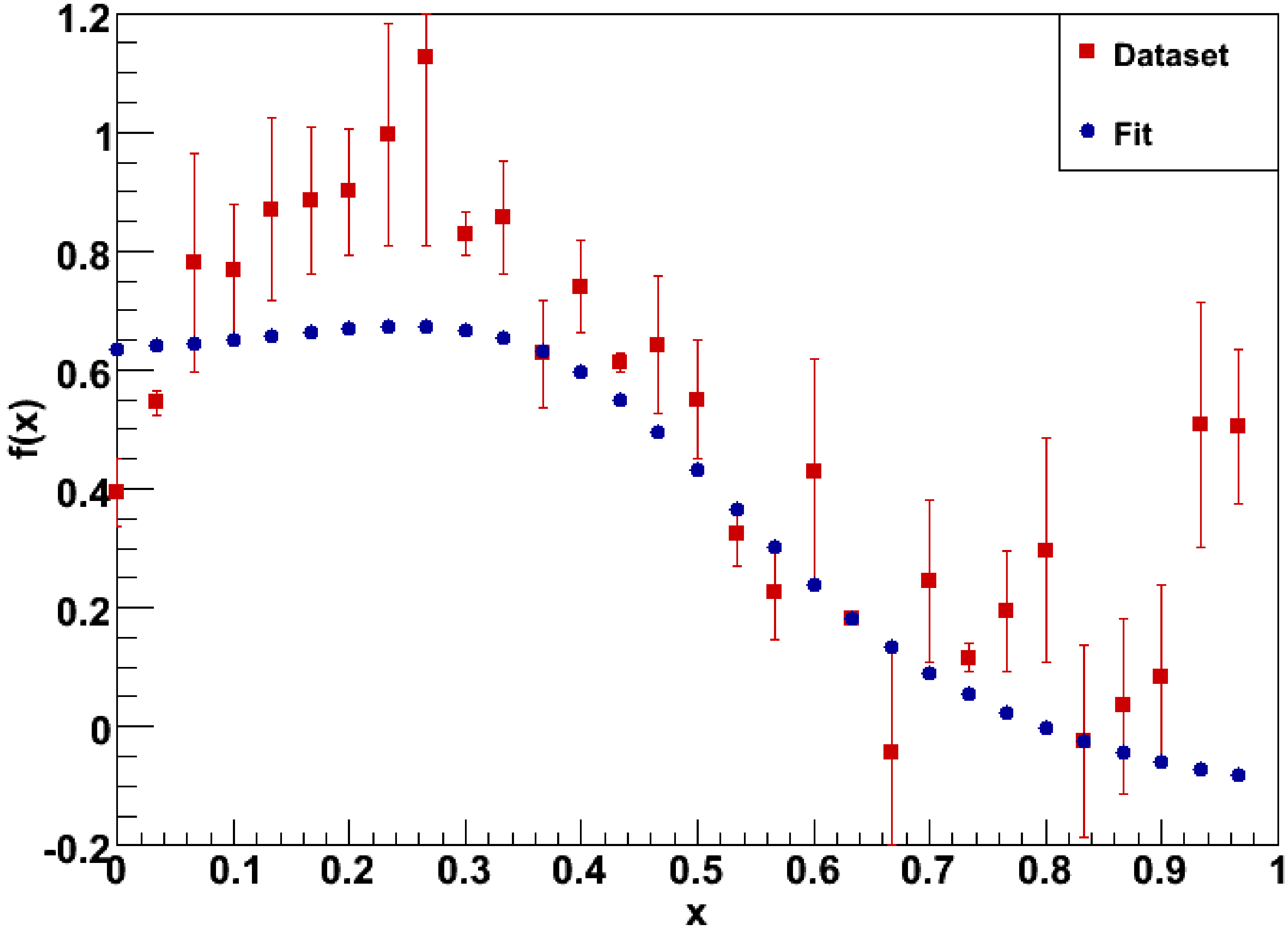
Optimal fitting vs. Overlearning: an example

the true function



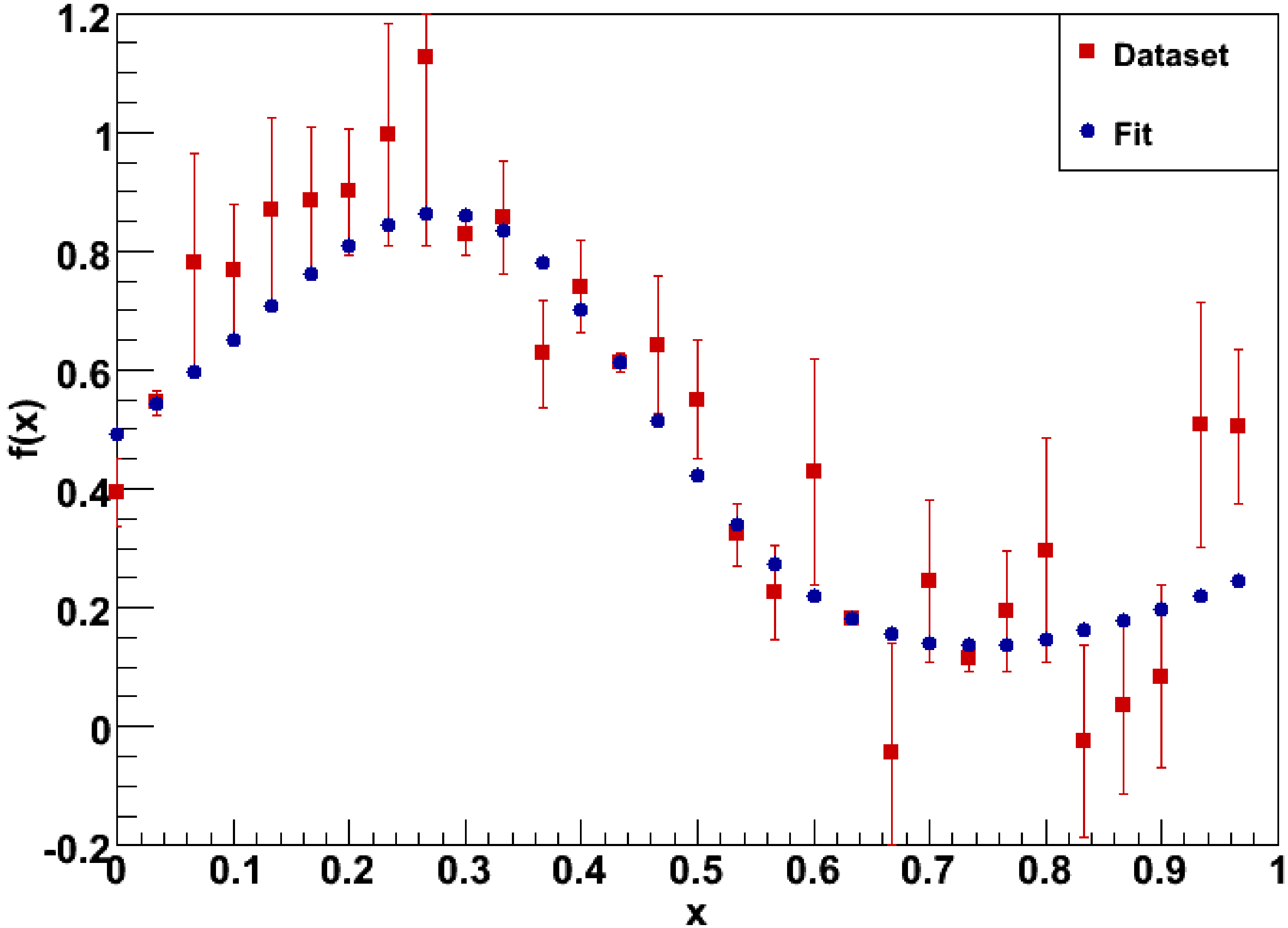
Optimal fitting vs. Overlearning: an example

UNDERLEARNING



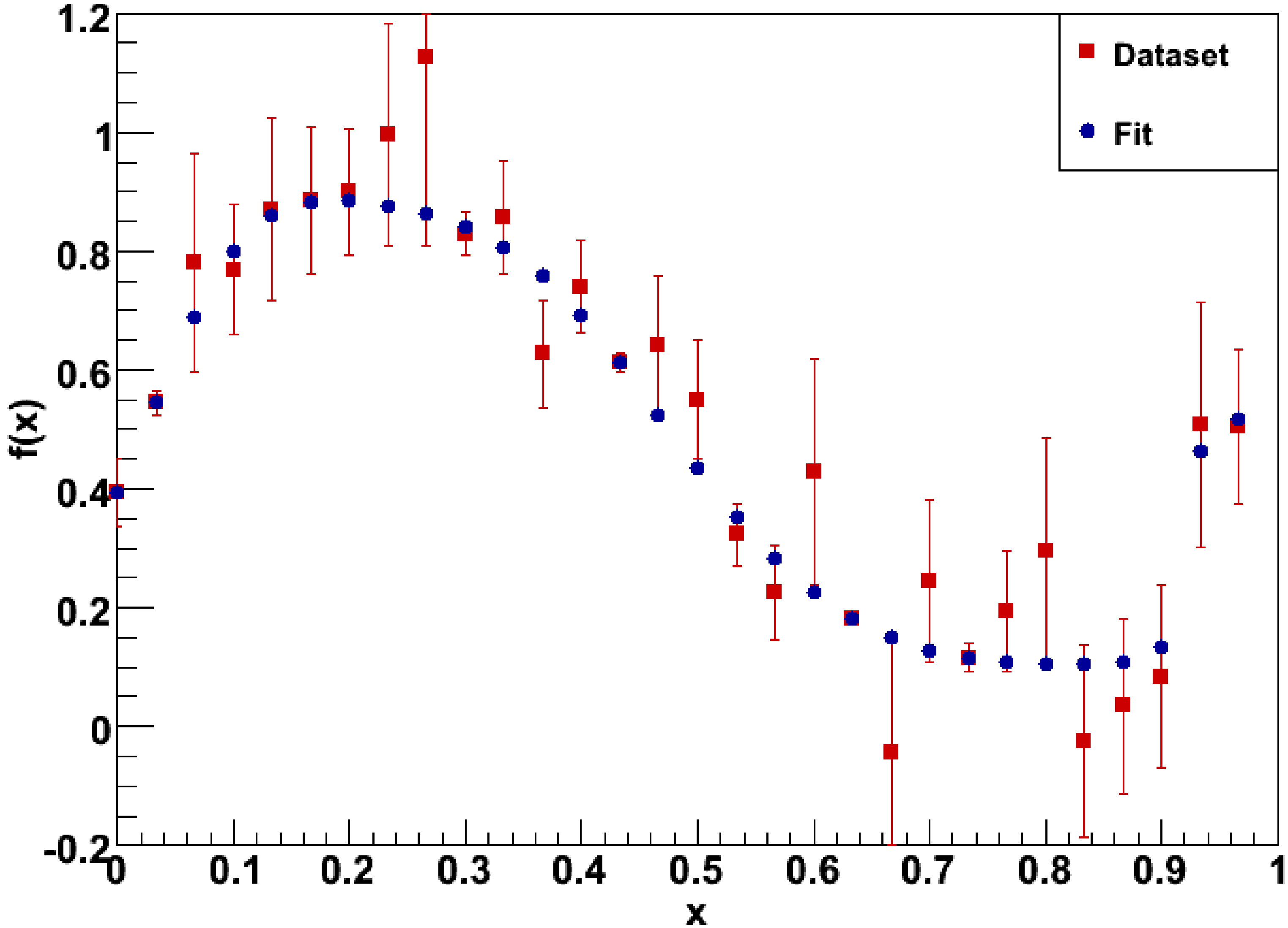
Optimal fitting vs. Overlearning: an example

optimal fit

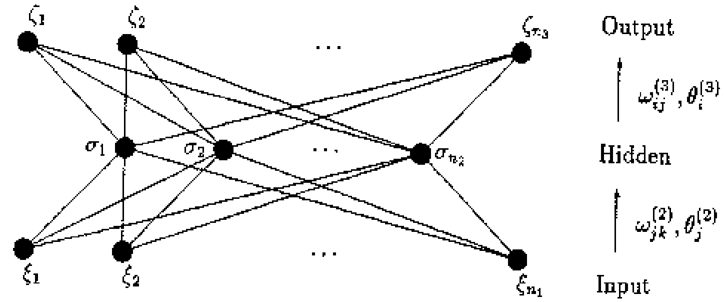


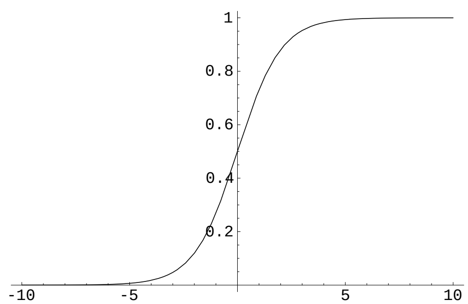
Optimal fitting vs. Overlearning: an example

overlearning



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

ANY FUNCTION CAN BE REPRESENTED BY A SUFFICIENTLY BIG NEURAL NETWORK

LESS PARAMETERS $\longrightarrow$ SMOOTHER FUNCTIONS

WHY NEURAL NETWORKS?

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?

WHY NEURAL NETWORKS?

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:

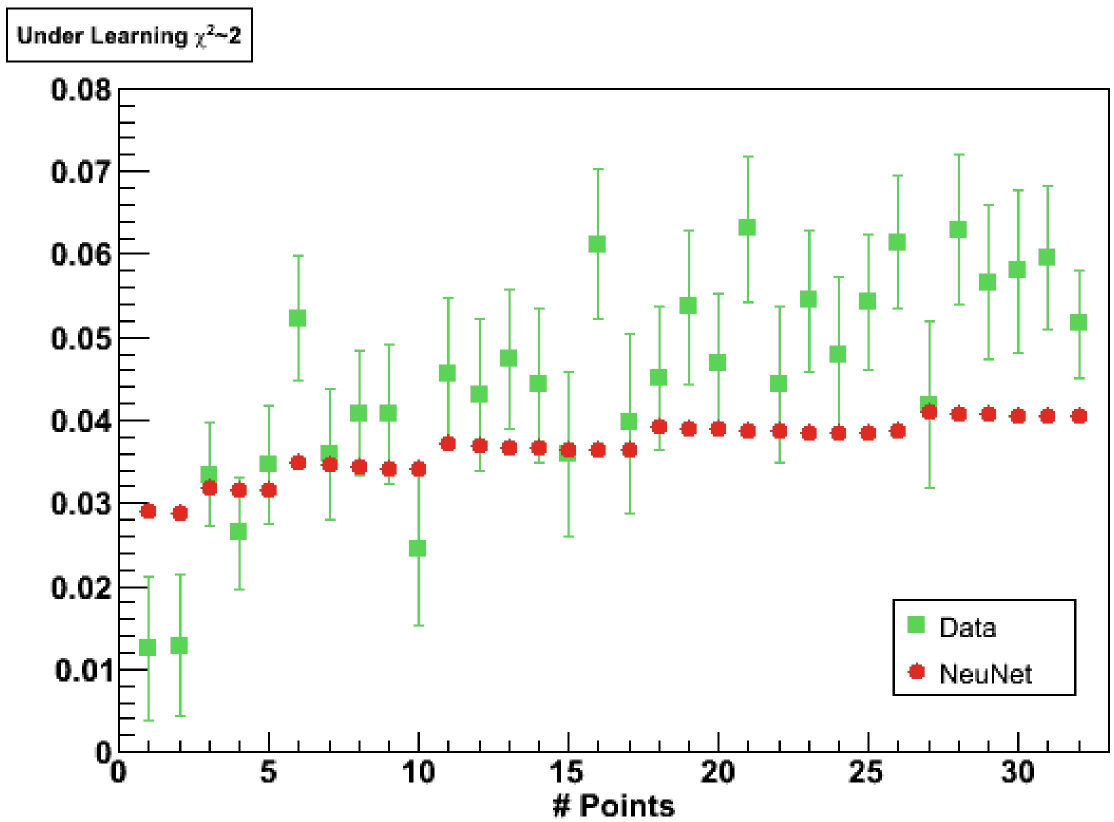
WHY NEURAL NETWORKS?

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



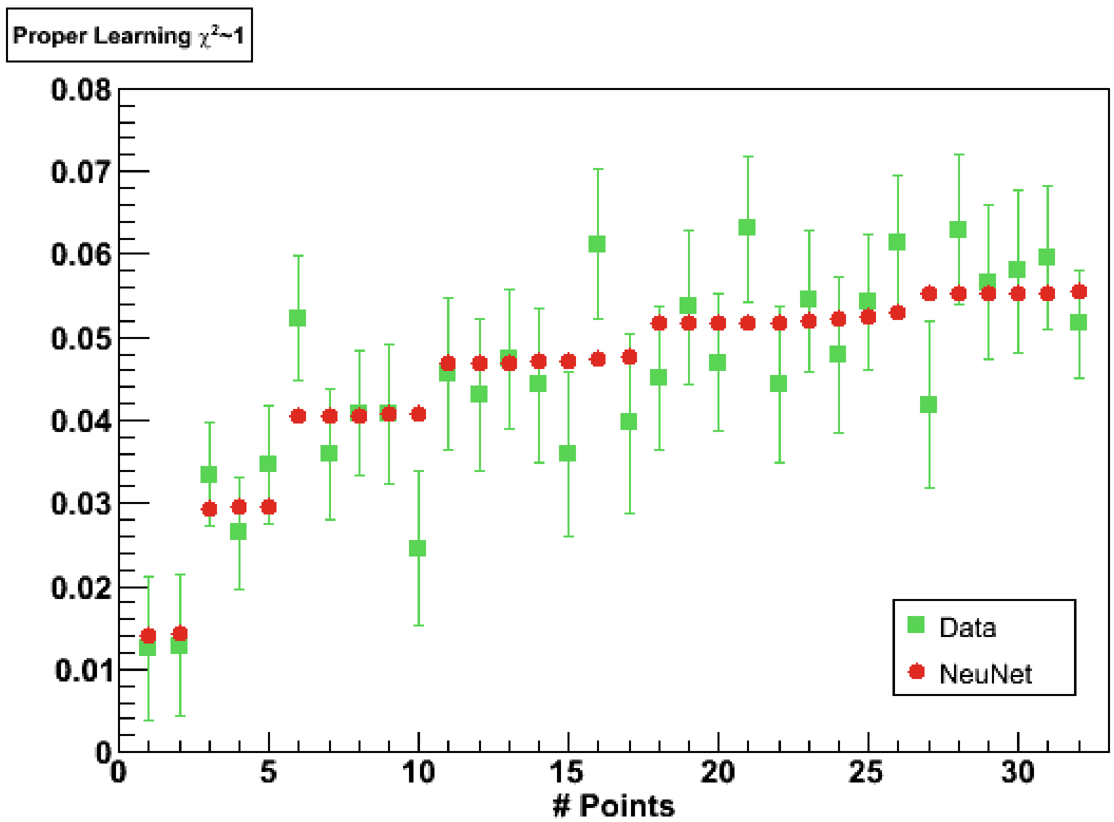
WHY NEURAL NETWORKS?

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



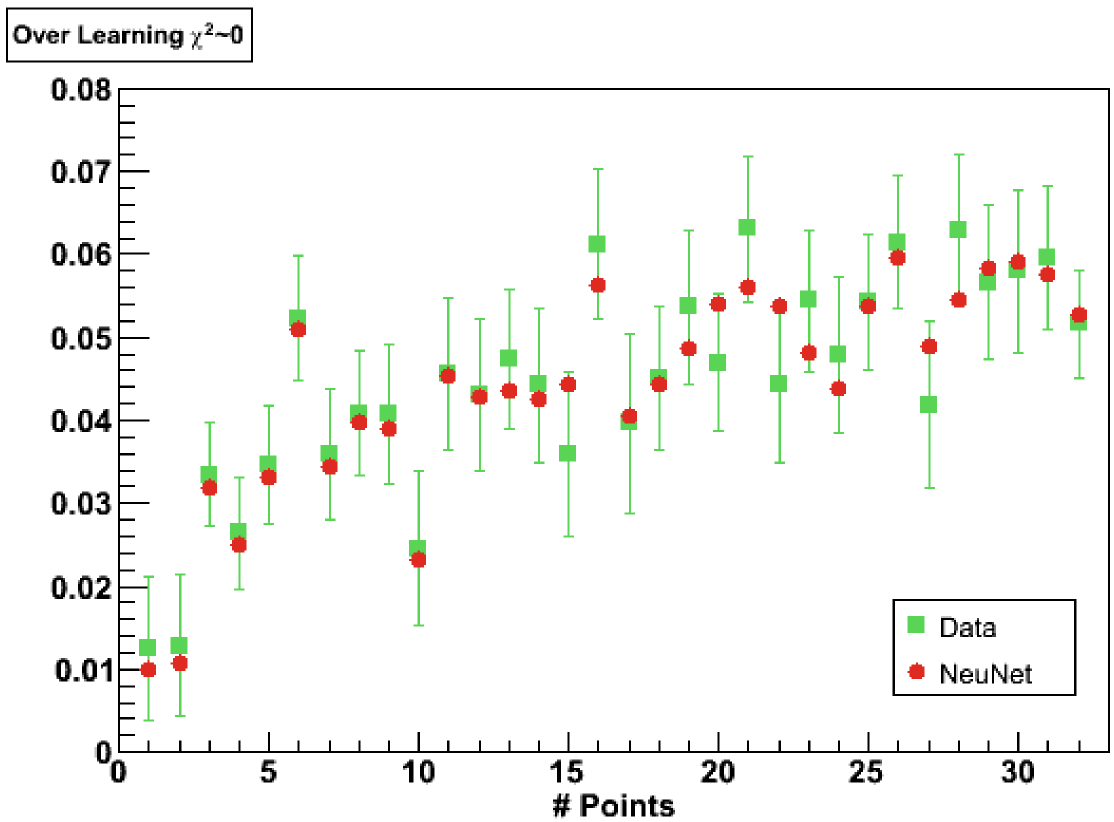
WHY NEURAL NETWORKS?

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



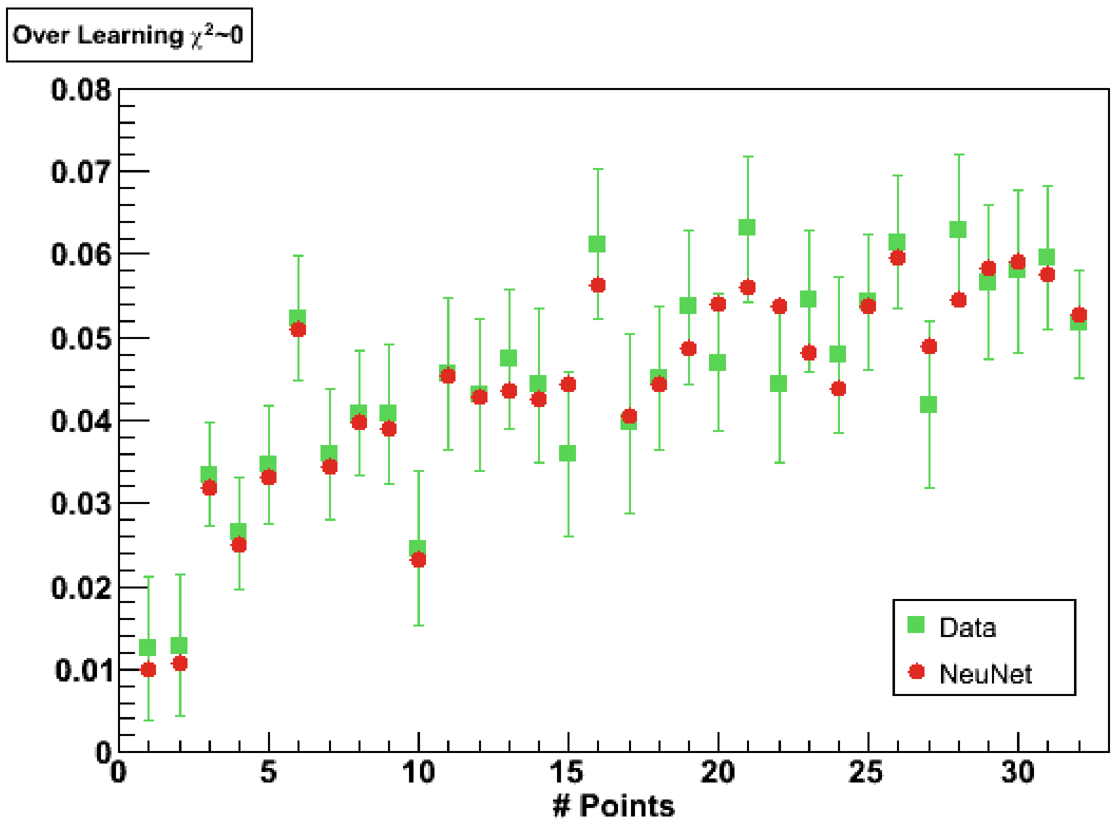
WHY NEURAL NETWORKS?

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!

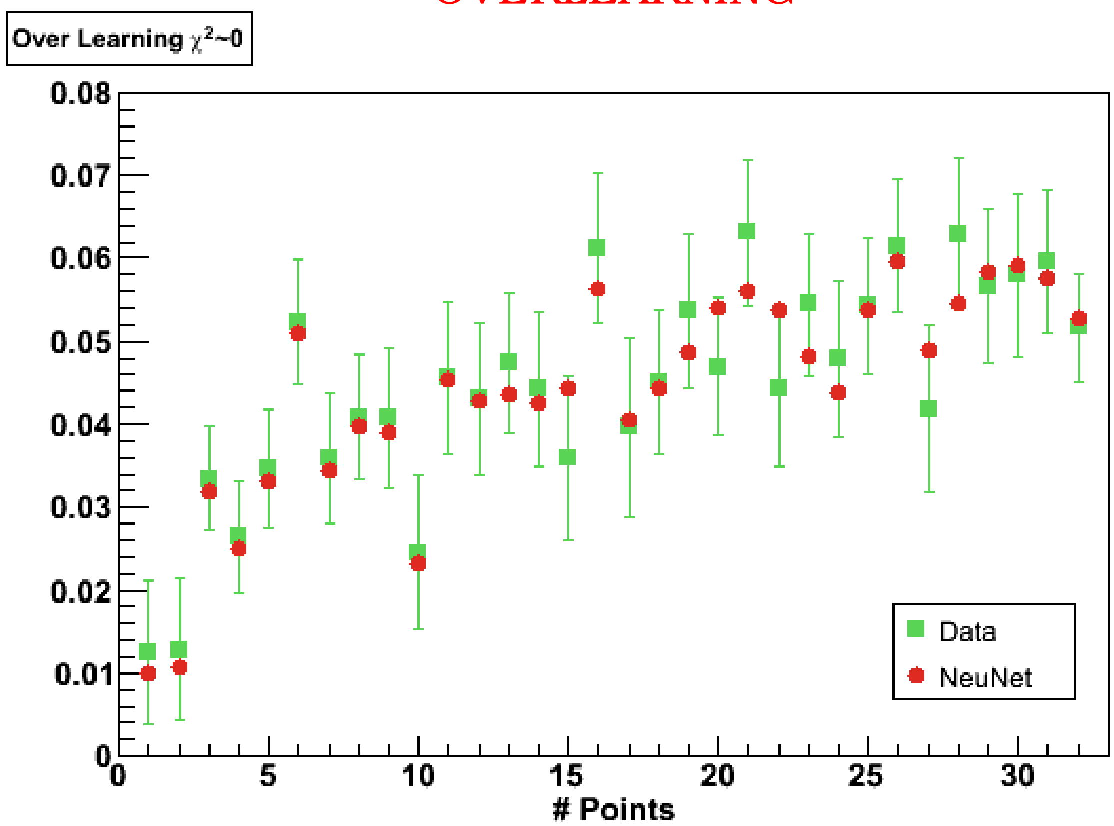
WHY NEURAL NETWORKS?

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

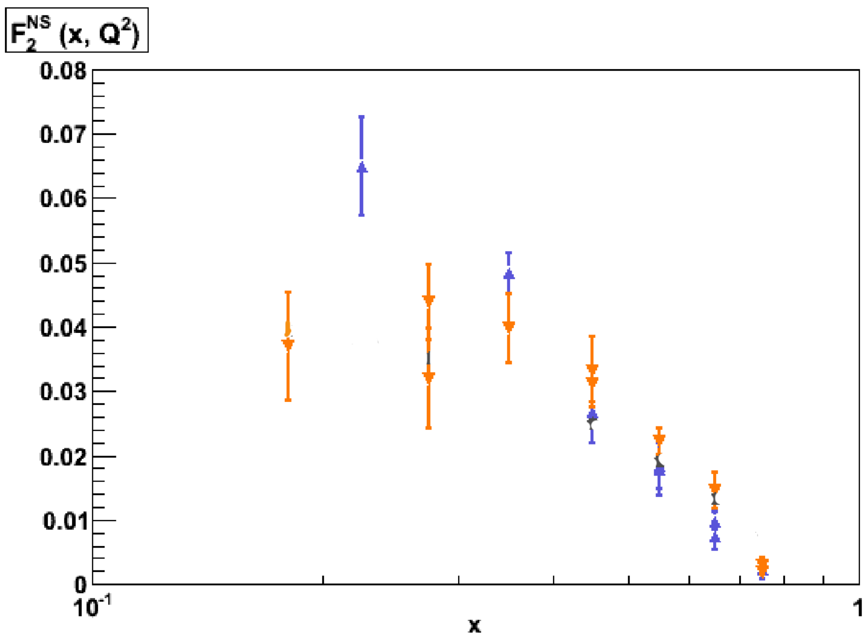


A: STOP THE FIT BEFORE OVERLEARNING SETS IN!
COULD BE DONE WITH STANDARD PARAMETRIZATIONS, BUT VERY INEFFICIENTLY

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

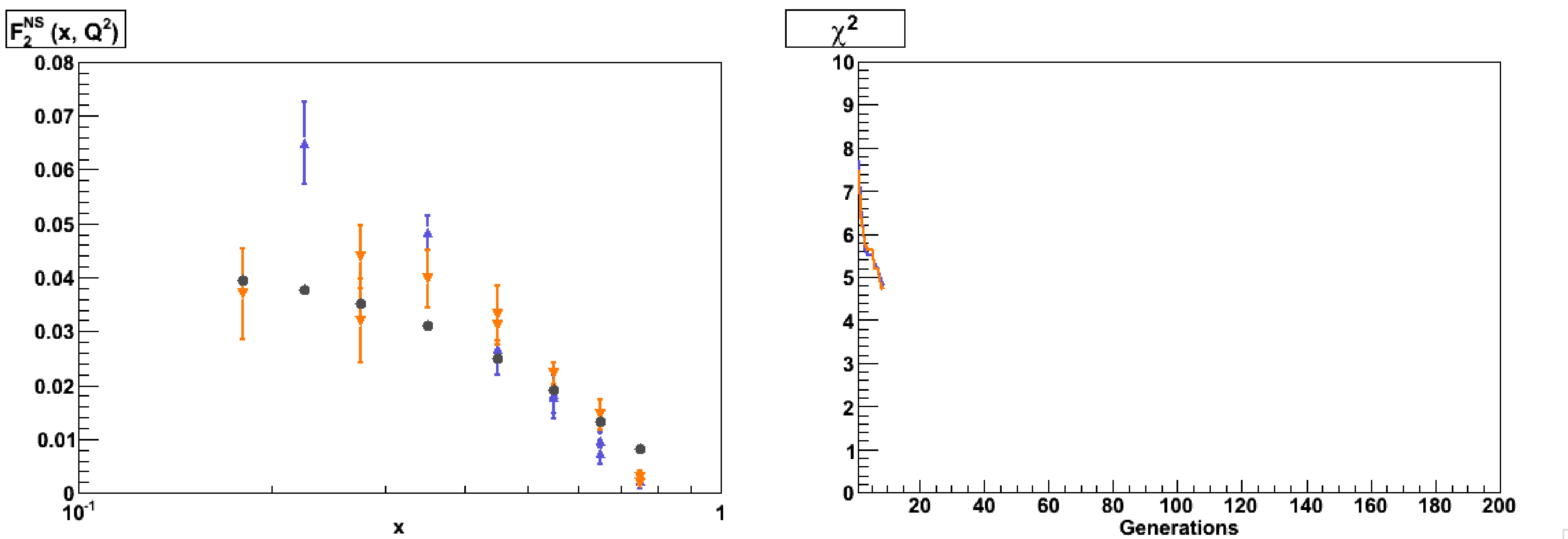


THE STOPPING CRITERION

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!

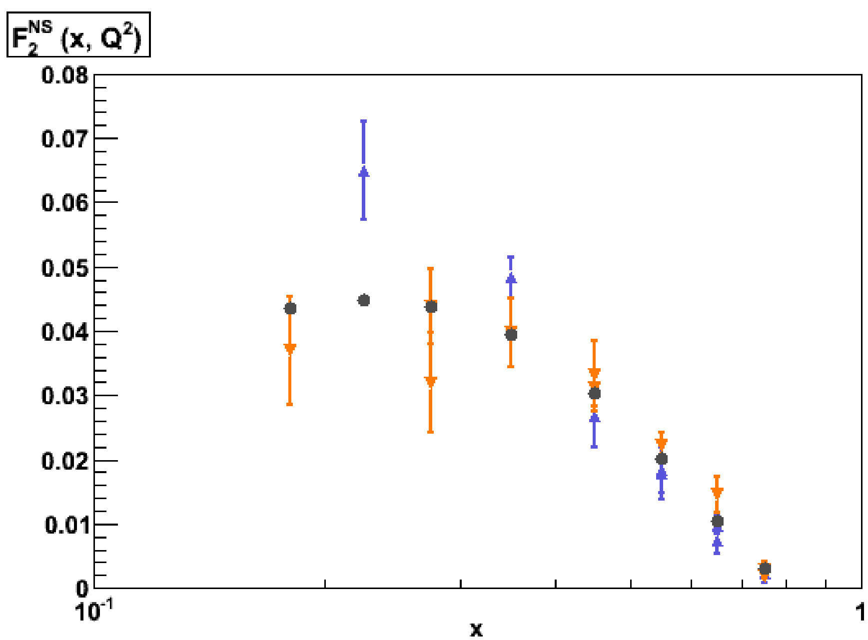


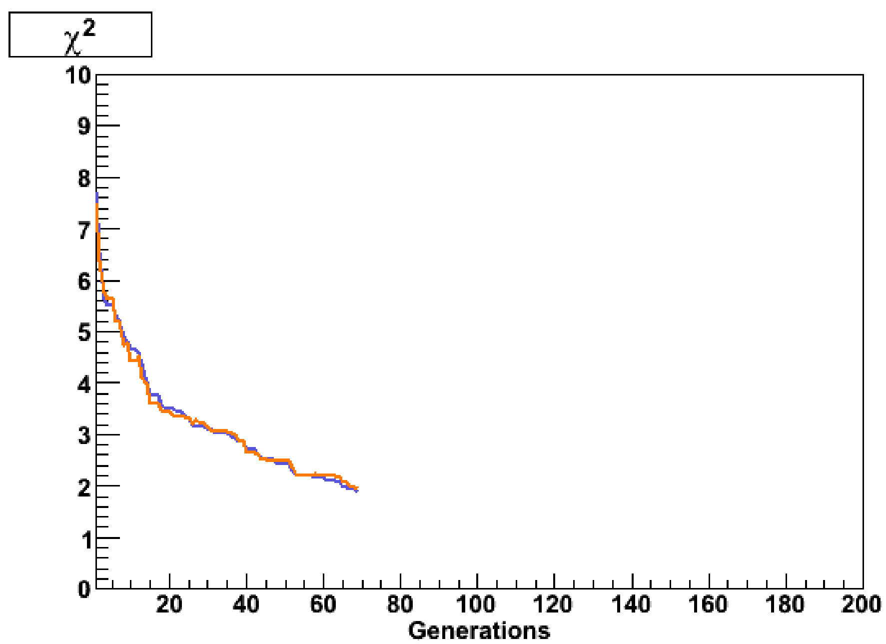
THE STOPPING CRITERION

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!



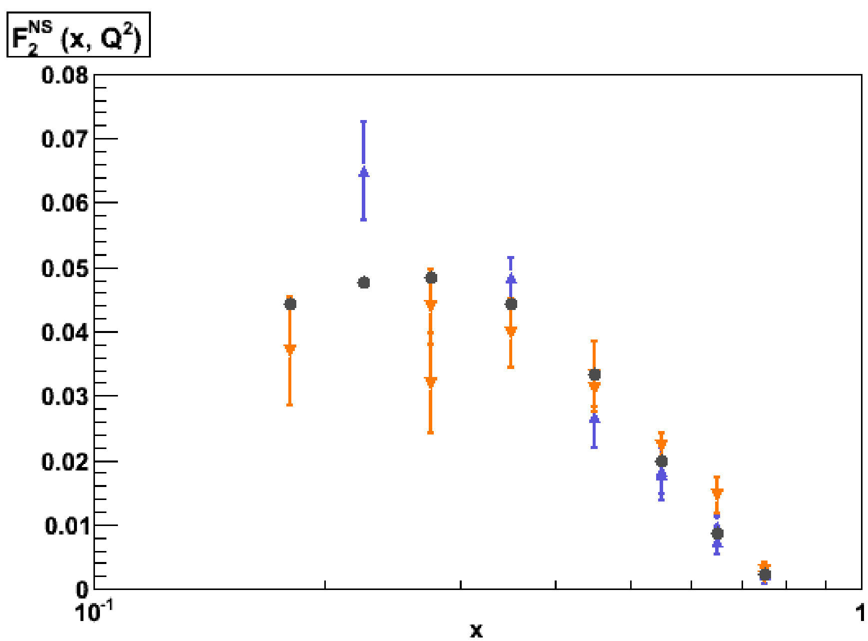


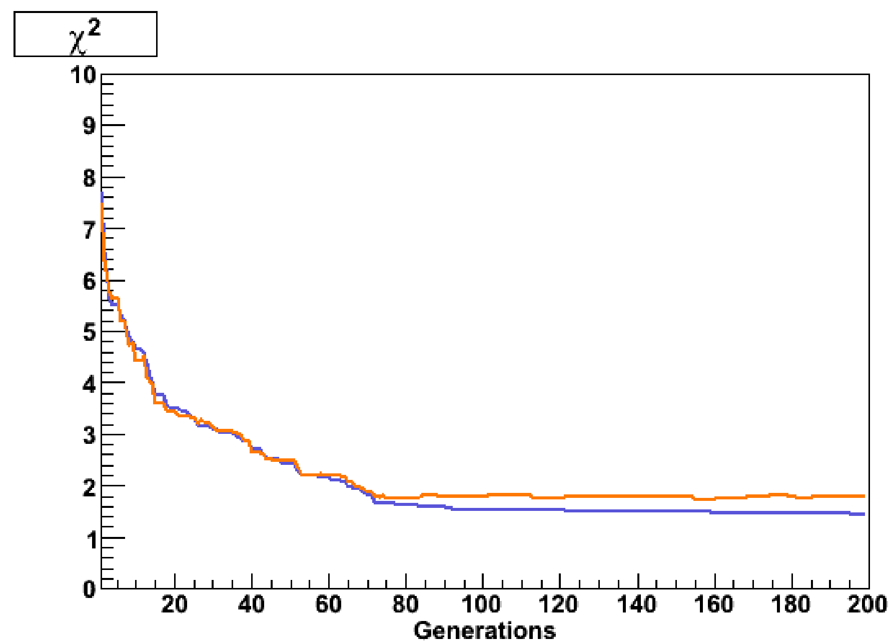
THE STOPPING CRITERION

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!





Monte Carlo data generation

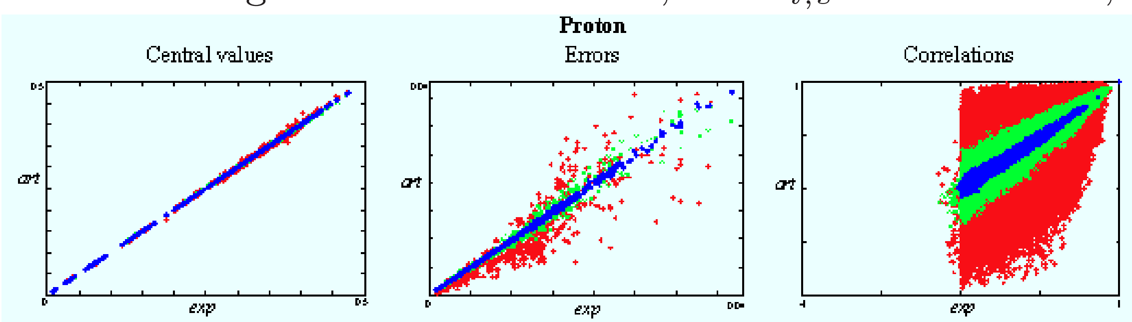
- BCDMS+ NMC proton & deuteron F_2 data (full correlated systematics available), taken at 4 beam energies
- on top of stat. errors, 4 systematics + 1 normalization (NMC) or 6 systematics + 1 absolute & 2 relative normalizations (BCDMS), with various forms of correlation (full, or for each target, or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



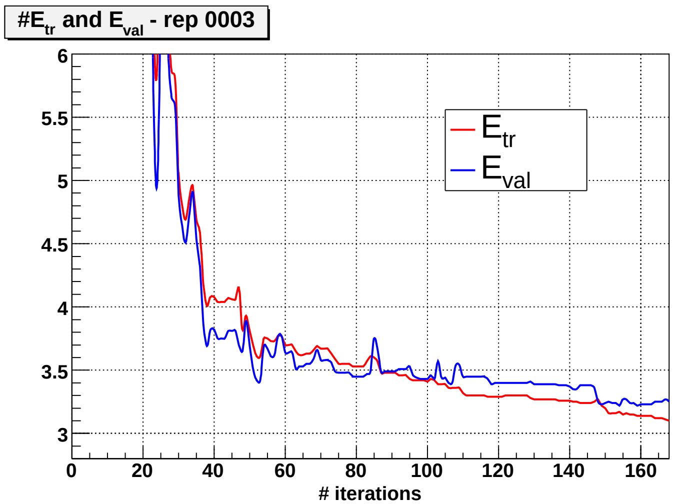
scatter plot art. vs. exp. for 10 (red) 100 (green) and 1000 (blue) replicas

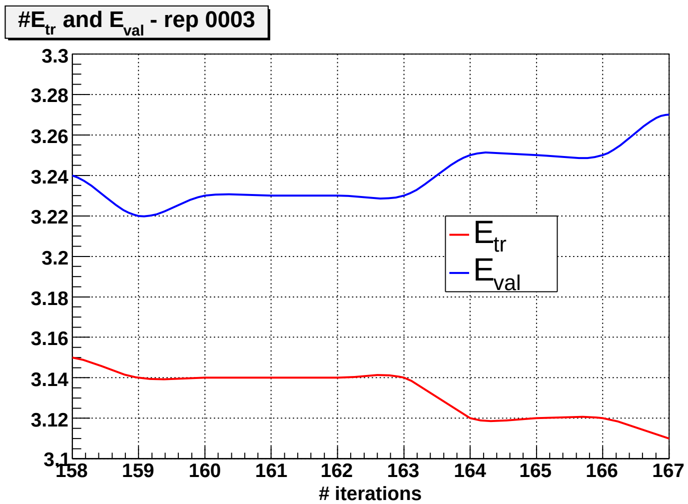
need 1000 replicas to reproduce correlations to percent accuracy

STOPPING I

- EACH NEURAL NET IS FITTED TO A PSEUDODATA REPLICA BY MINIMIZING THE χ^2 TO SUBSET OF DATA (TRAINING SET)
- FIT STOPS WHEN THE χ^2 OF THE REMAINING DATA STARTS TO GROW (VALIDATION SET)

STOPPING FOR THE χ^2 OF ONE REPLICA (FULL FIT)





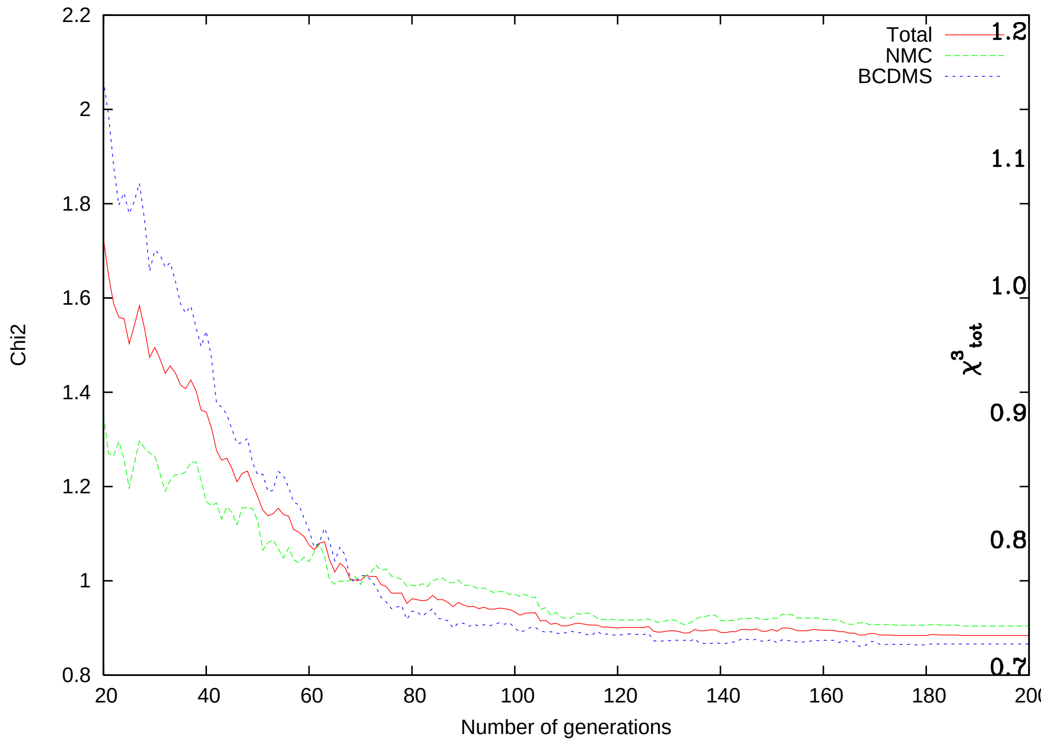
PERTURBATIVE EVOLUTION

- 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

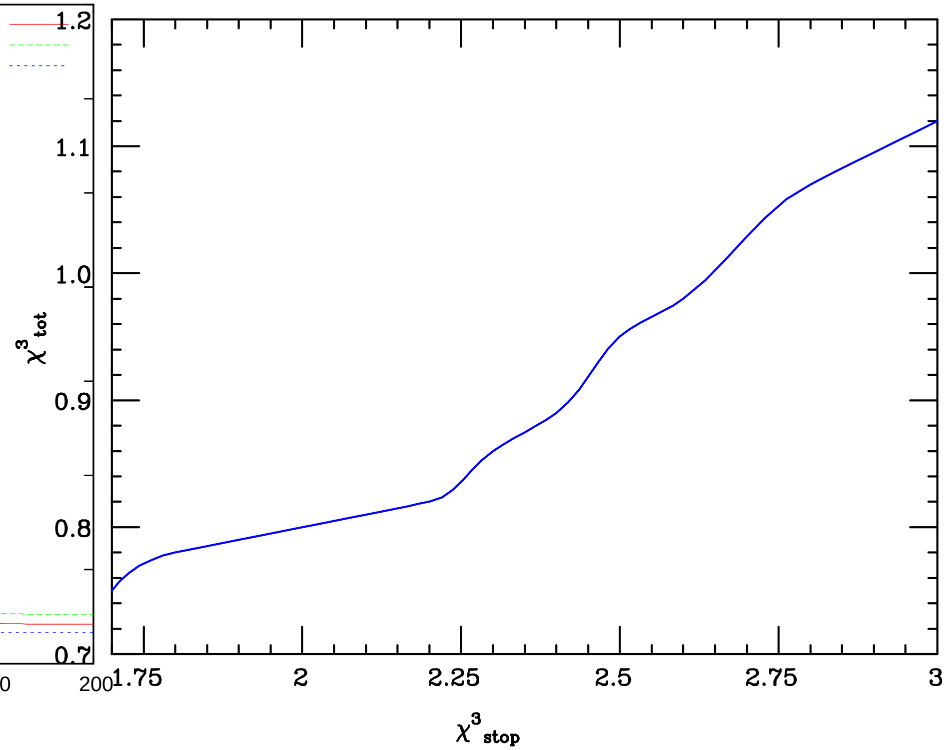
TRAINING...

- 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



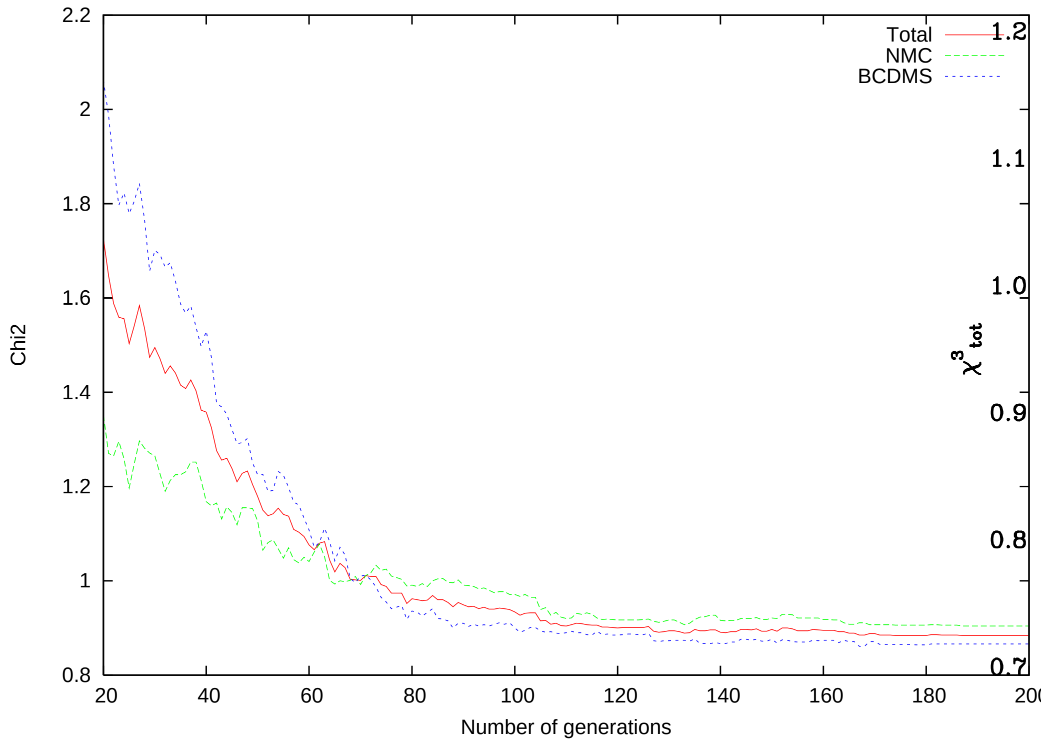
χ^2 OF BEST FIT VS. AVERAGE χ^2



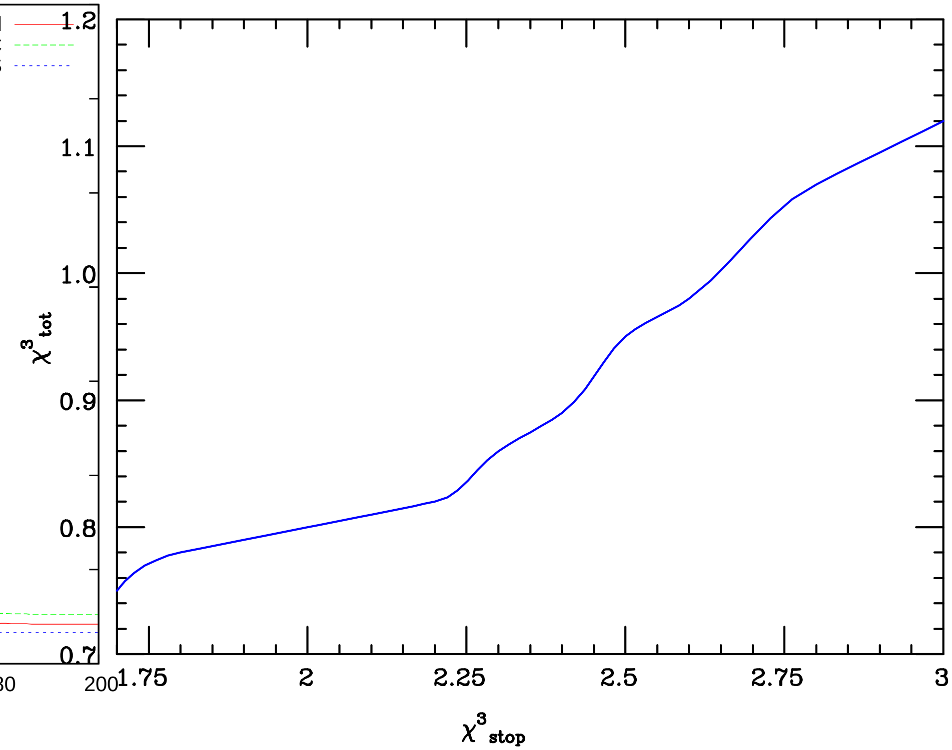
TRAINING...

- 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