MACHINE LEARNING AN UNKNOWN PHYSICAL LAW:

THE STRUCTURE OF THE PROTON

Stefano Forte

Università di Milano & INFN

UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI FISICA

CP[3] Lecture Series

Odense, December 2, 2019

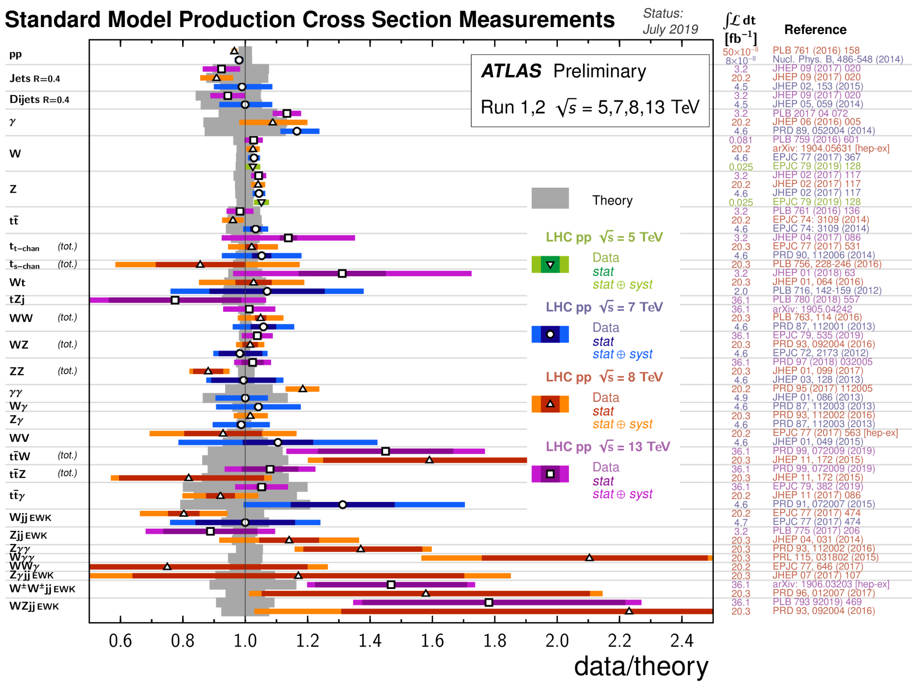
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 740006

PHYSICS AT THE LHC AS PRECSION PHYSICS

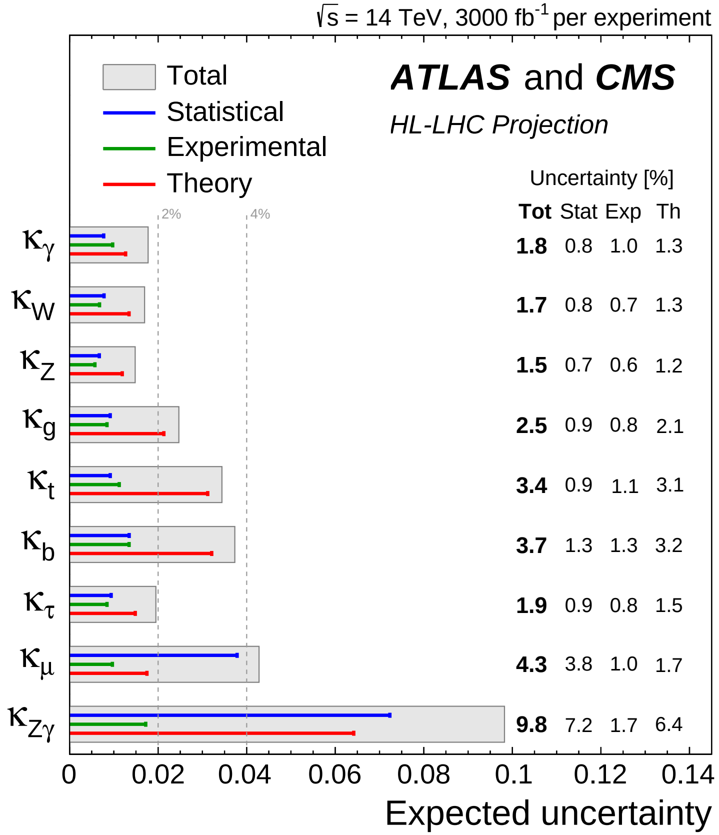
DEVIATIONS FROM SM

SM CROSS-SECTIONS TODAY:

TH. VS EXP



HL-LHC: 2024-2040



	Tot	Stat	Exp	Th
κ_γ	1.8	0.8	1.0	1.3
κ_W	1.7	0.8	0.7	1.3
κ_Z	1.5	0.7	0.6	1.2
κ_g	2.5	0.9	0.8	2.1
κ_t	3.4	0.9	1.1	3.1
κ_b	3.7	1.3	1.3	3.2
κ_τ	1.9	0.9	0.8	1.5
κ_μ	4.3	3.8	1.0	1.7
$\kappa_{Z\gamma}$	9.8	7.2	1.7	6.4

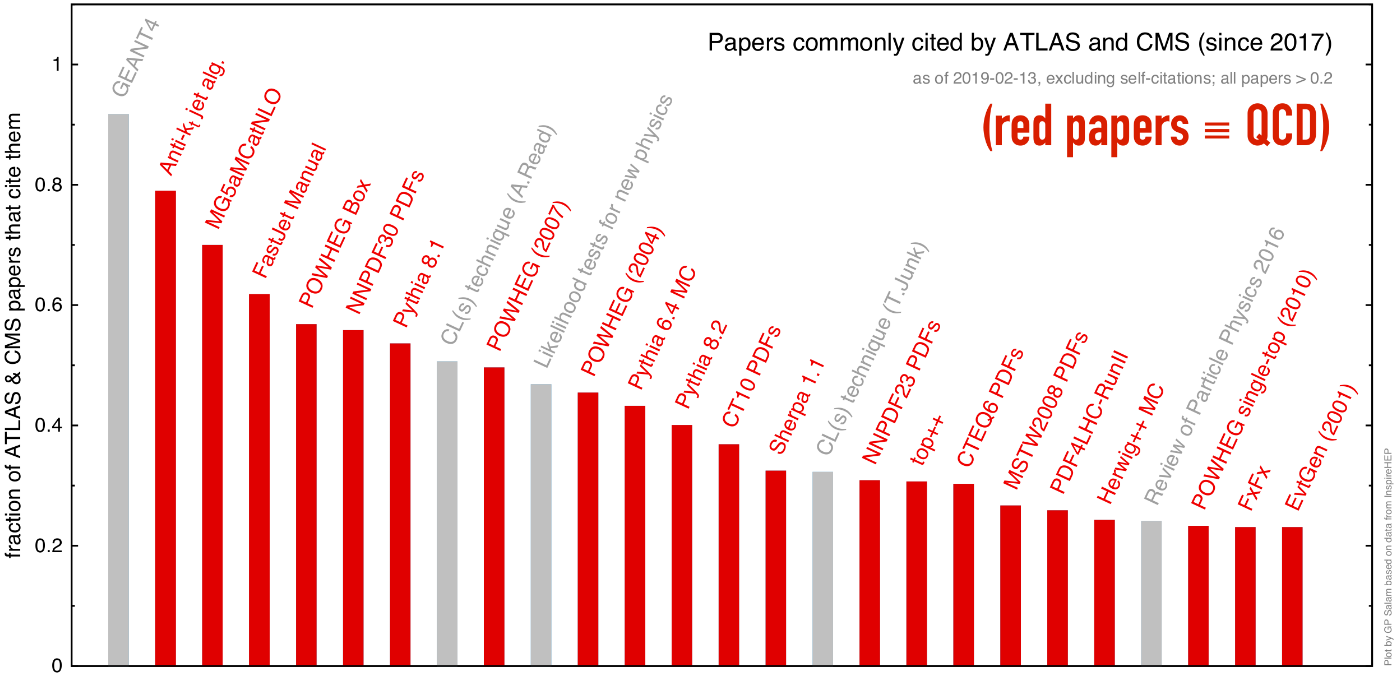
$$\kappa_j^2 = \sigma_j / \sigma^{\mathrm{SM}}$$

- SM TESTED AT THE PERCENT LEVEL
- SEEING DEVIATIONS REQUIRES SUB-PERCENT ACCURACY

UNCERTAINTIES AND QCD

- THE LHC IS A PROTON COLLIDER $\Rightarrow$ ANY INTERACTION CONTAINS A STRONG INTERACTION
- QCD IS THE MAIN THEORETICAL PROBLEM
- .

PAPERS MOST CITED BY ATLAS (BY FRACTION)

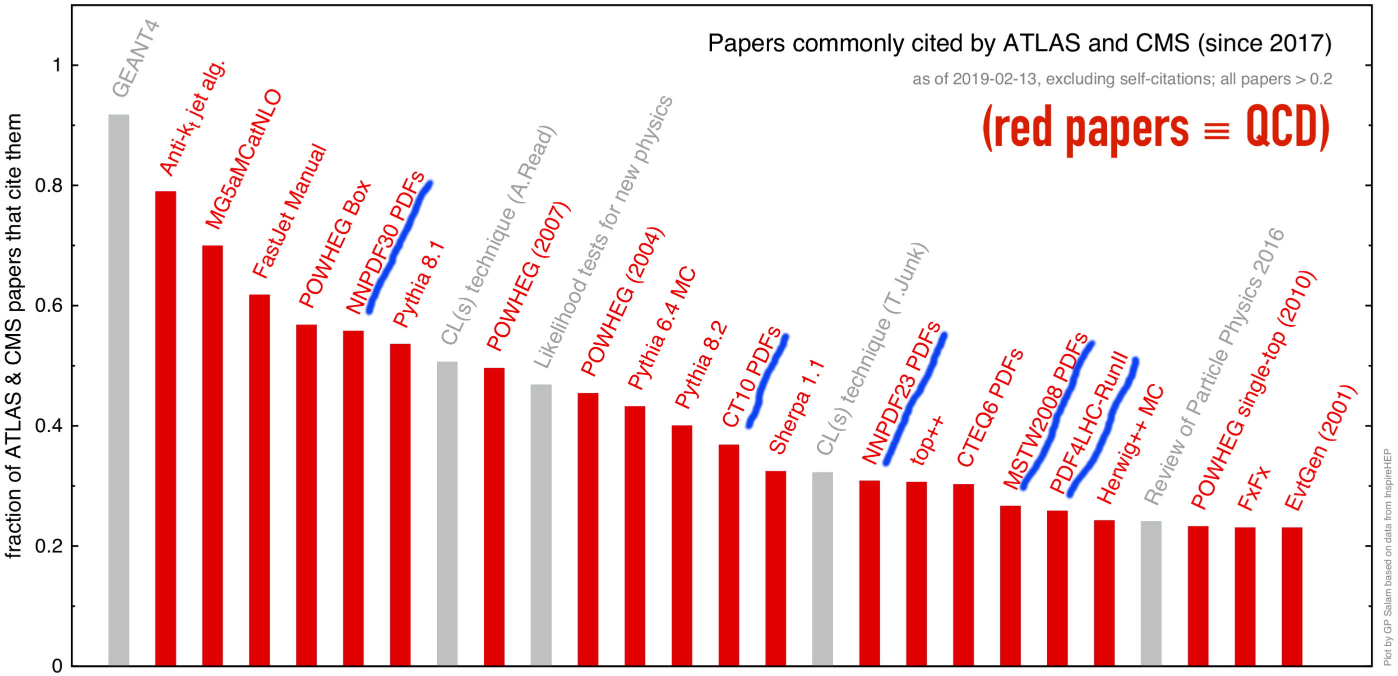


(G. Salam, 2019)

UNCERTAINTIES QCD, AND PDFS

- THE LHC IS A PROTON COLLIDER ⇒ ANY INTERACTION CONTAINS A STRONG INTERACTION
- QCD IS THE MAIN THEORETICAL PROBLEM
- PDFS ARE THE DOMINANT ISSUE

PAPERS MOST CITED BY ATLAS (BY FRACTION)

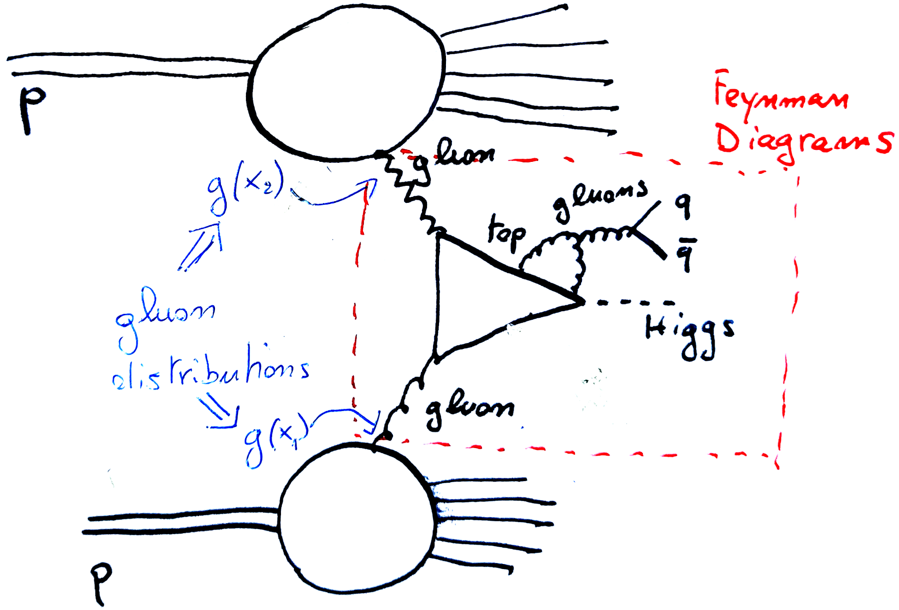


(G. Salam, 2019)

PDF papers underlined

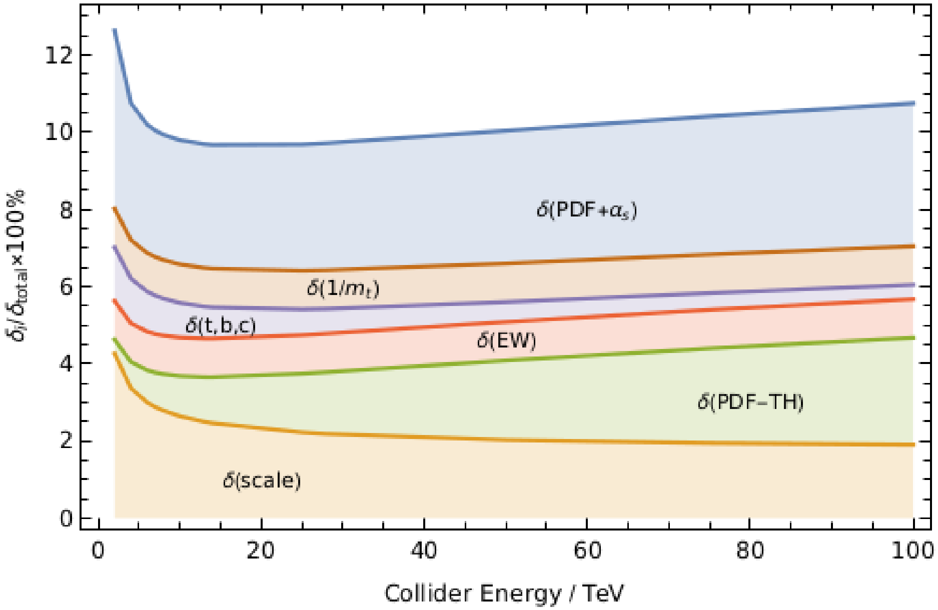
UNCERTAINTIES AND PDFs

QCD FACTORIZATION



UNCERTAINTIES:

HIGGS IN GLUON FUSION

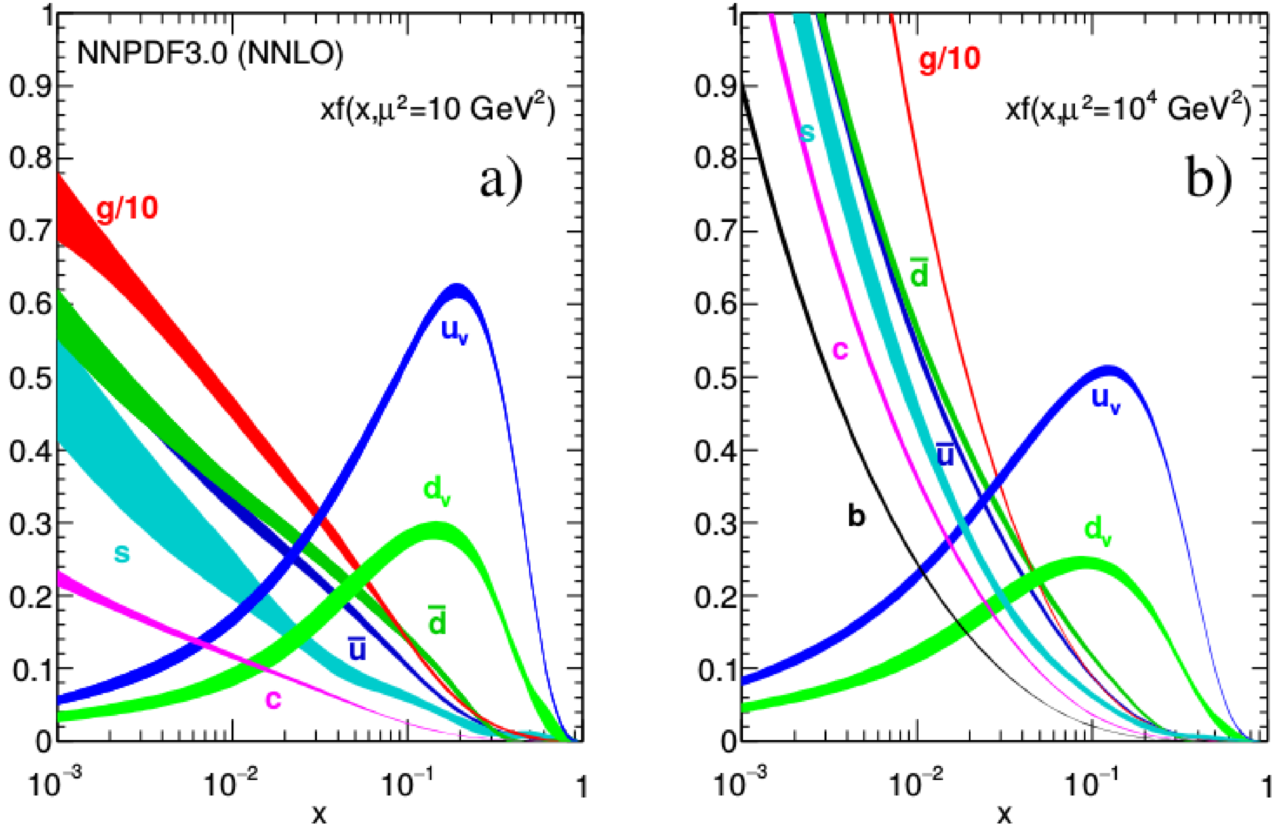


(HL-LHC Higgs WG report, 2019)

- PDF ESPRESS THE LIKELIHOOD OF A QUARK OR GLUONS (PARTONS) TO ENTER A COLLISION
- THEIR KNOWLEDGE IS A DOMINANT SOURCE OF UNCERTAINTY

A PORTRAIT OF THE PROTON

AS SEEN FROM A HIGGS BOSON



(PDG 2018)

- PARTON DISTRIBUTIONS: MOMENTUM FRACTION DISTRIBUTIONS FOR EACH TYPE OF QUARK, ANTIQUARK & THE GLUON
- EXTRACTED FROM DATA, COMPARING PDF-DEPENDENT PREDICTION & INVERTING

DISCOVERY AT A HADRON COLLIDER AND PDFS
THE DISCOVERY OF THE W (1984)

THEORETICAL PREDICTION

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

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

ALTARELLI, ELLIS, GRECO, MARTINELLI, 1984

EXPERIMENTAL DISCOVERY

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN-EP/85-108
11 July 1985

W PRODUCTION PROPERTIES AT THE CERN SPS COLLIDER

UA1 Collaboration, CERN, Geneva, Switzerland

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

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

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

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

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

DISCOVERY AT A HADRON COLLIDER AND PDFs
THE DISCOVERY OF THE W (1984)

PDFs IN 1984

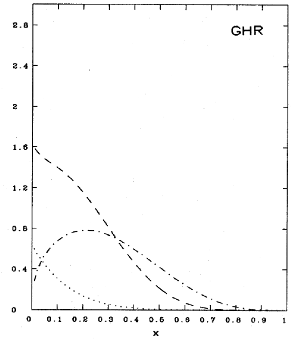


FIG. 25. Parton distributions of Glück, Hoffmann, and Reya (1982), at $Q^2=5$ GeV^2: valence quark distribution $x[u_v(x)+d_v(x)]$ (dotted-dashed line), $xG(x)$ (dashed line), and q_v (dotted line).

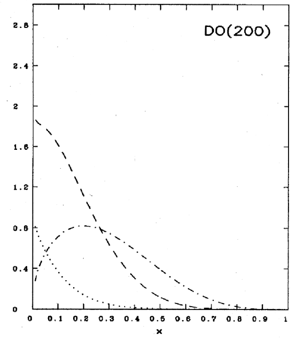


FIG. 27. "Soft-gluon" ($\Lambda=200$ MeV) parton distributions of Duke and Owens (1984) at $Q^2=5$ GeV^2: valence quark distribution $x[u_v(x)+d_v(x)]$ (dotted-dashed line), $xG(x)$ (dashed line), and $q_v(x)$ (dotted line).

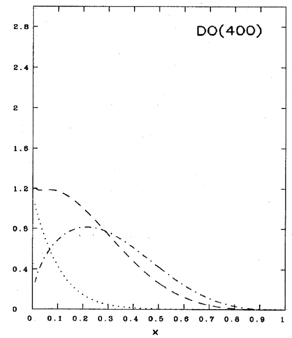


FIG. 26. "Hard-gluon" ($\Lambda=400$ MeV) parton distributions of Duke and Owens (1984) at $Q^2=5$ GeV^2: valence quark distribution $x[u_v(x)+d_v(x)]$ (dotted-dashed line), $xG(x)$ (dashed line), and $q_v(x)$ (dotted line).

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

GHR VS DUKE-OWENS

THEORETICAL PREDICTION

42 *G. Altarelli et al. / Vector boson production*

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

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

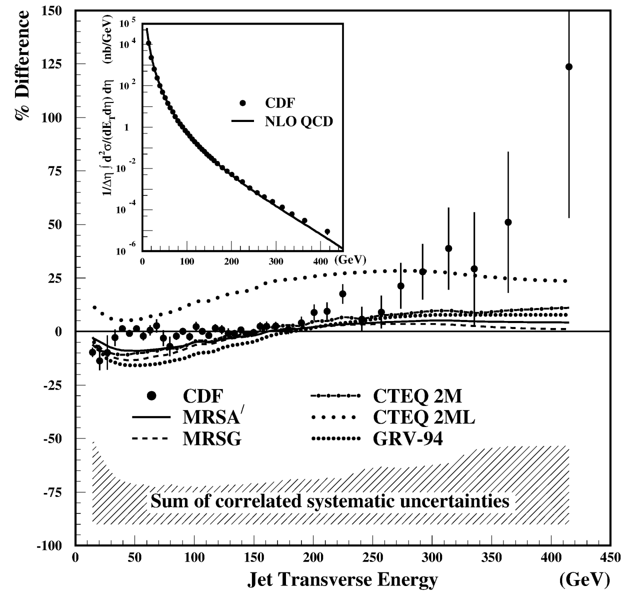
ALTARELLI, ELLIS, GRECO, MARTINELLI, 1984

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

DISCOVERY AT A HADRON COLLIDER AND PDFs

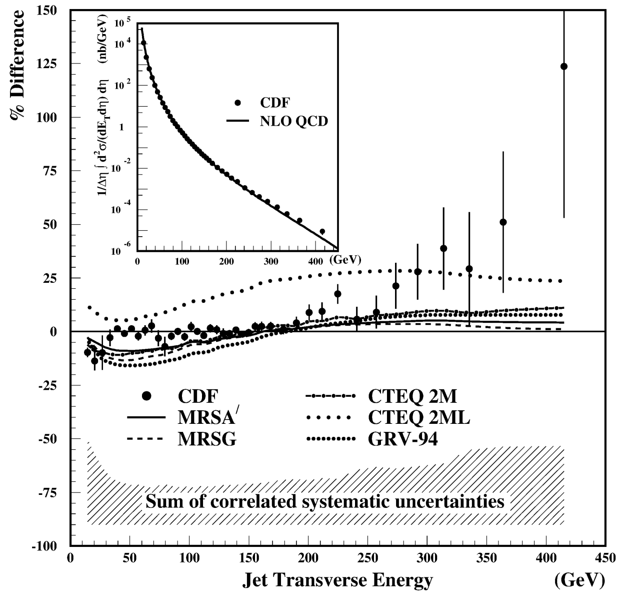
THE DISCOVERY OF QUARK COMPOSITENESS (1995)

- DISCREPANCY BETWEEN QCD CALCULATION AND CDF JET DATA (1995)
- EVIDENCE FOR QUARK COMPOSITENESS
- .



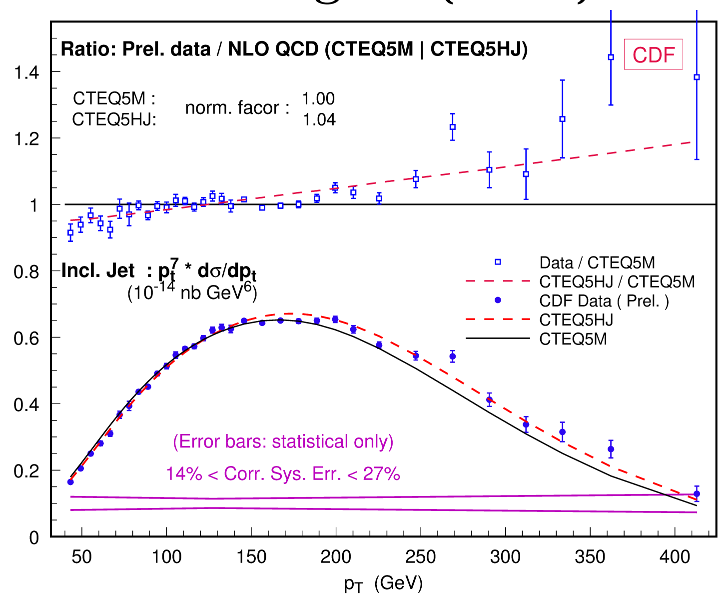
DISCOVERY AT A HADRON COLLIDER AND PDFs
A BETTER DETERMINATION OF THE GLUON PDF (1995)

- DISCREPANCY BETWEEN QCD CALCULATION AND CDF JET DATA (1995)
- ~~EVIDENCE FOR QUARK COMPOSITENESS~~
- NO INFO ON PARTON UNCERTAINTY $\Rightarrow$ RESULT STRONGLY DEPENDS ON GLUON AT $x \gtrsim 0.1$

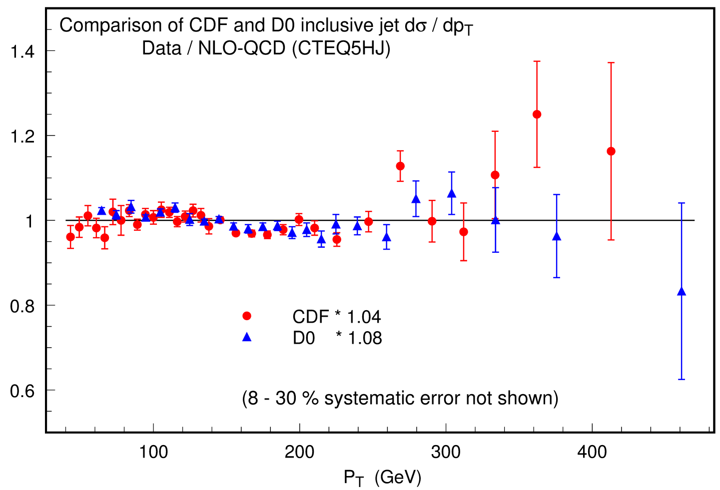


DISCREPANCY REMOVED IF JET DATA INCLUDED IN THE FIT

NEW CTEQ FIT (1996)



FINAL CTEQ FIT (1998)



WHAT'S THE PROBLEM ~ 2000

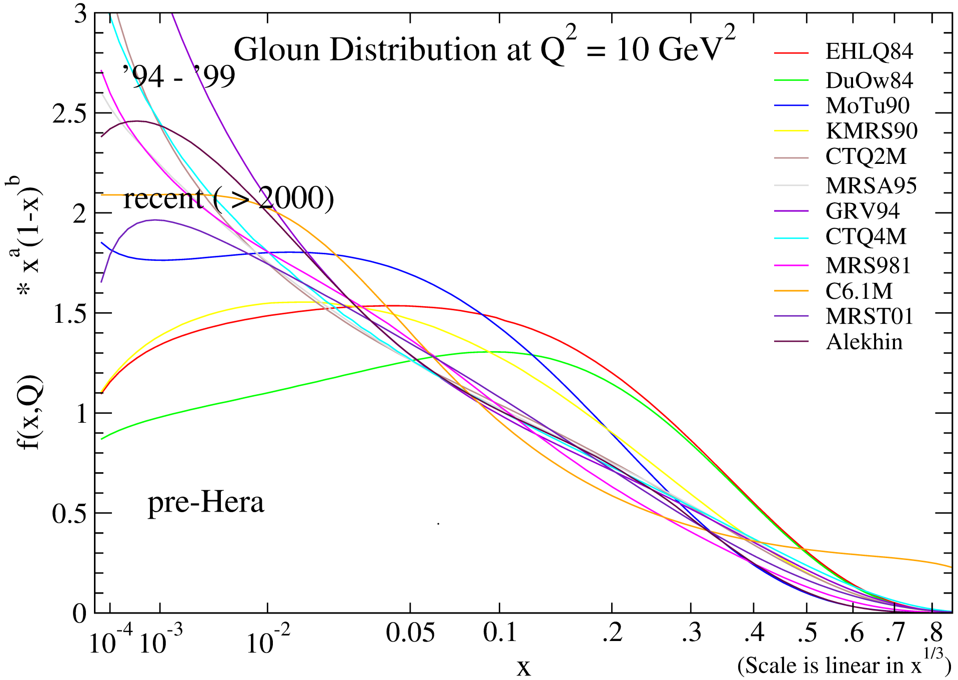
PDFs DETERMINED FITTING A MODEL-INSPIRED FUNCTIONAL FORM

gluon parametrization (MRST 2004)

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

- PROBLEM REDUCED TO FINITE-DIMENSIONAL
- WHO PICKS THE FUNCTIONAL FORM?

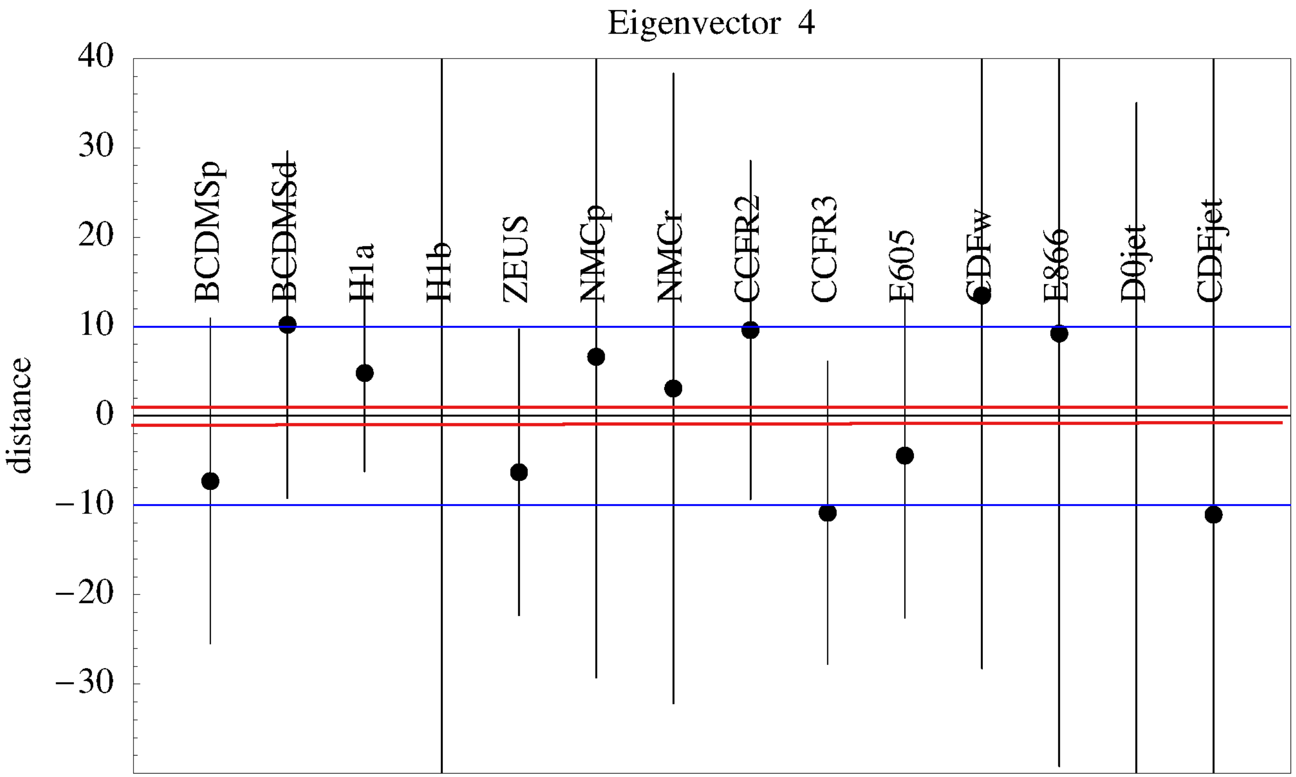
HISTORICAL COMPILATION OF GLUON PDFs



W.K.Tung, DIS 2004

FIRST PDFs WITH UNCERTAINTIES (2002)
"TOLERANCE"

one sigma & ten sigma intervals for typical
covariance matrix eigenvalue
vs best value and uncertainty from individual experiments

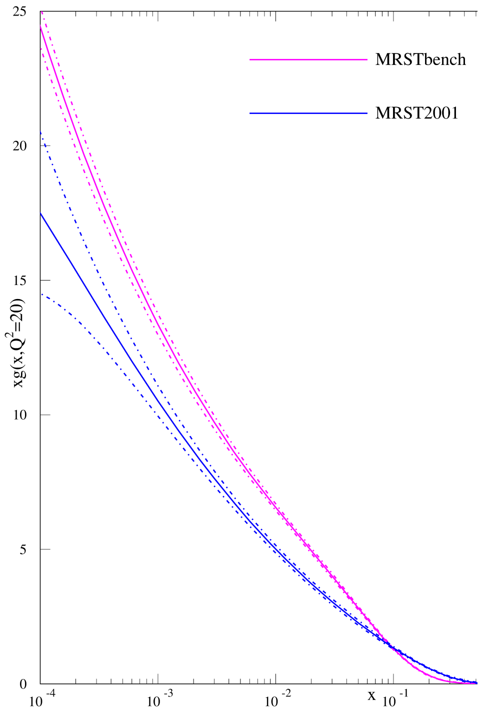


- SPREAD OF BEST-FIT FROM DIFFERENT DATA HUGE W.R. TO TEXTBOOK UNCERTAINTIES
- PDF UNCERTAINTIES RESCALED BY "TOLERANCE" $T \sim 10$

THE HERA-LHC BENCHMARK (2005)

- RESTRICTED AND VERY CONSISTENT DATASET USED
- RESULTS COMPARED TO THEN-BEST RESULT FROM FULL DATASET

BENCHMARK VS DEFAULT GLUON



"...the partons extracted using a very limited data set are completely incompatible, even allowing for the uncertainties, with those obtained from a global fit with an identical treatment of errors...The comparison illustrates the problems in determining the true uncertainty on parton distributions." (R.Thorne, HERALHC, 2005)

A NEW APPROACH: USING AI TOOLS

PUBLISHED BY INSTITUTE OF PHYSICS PUBLISHING FOR SISSA/ISAS

RECEIVED: *April 24, 2002*
REVISED: *May 30, 2002*
ACCEPTED: *May 31, 2002*

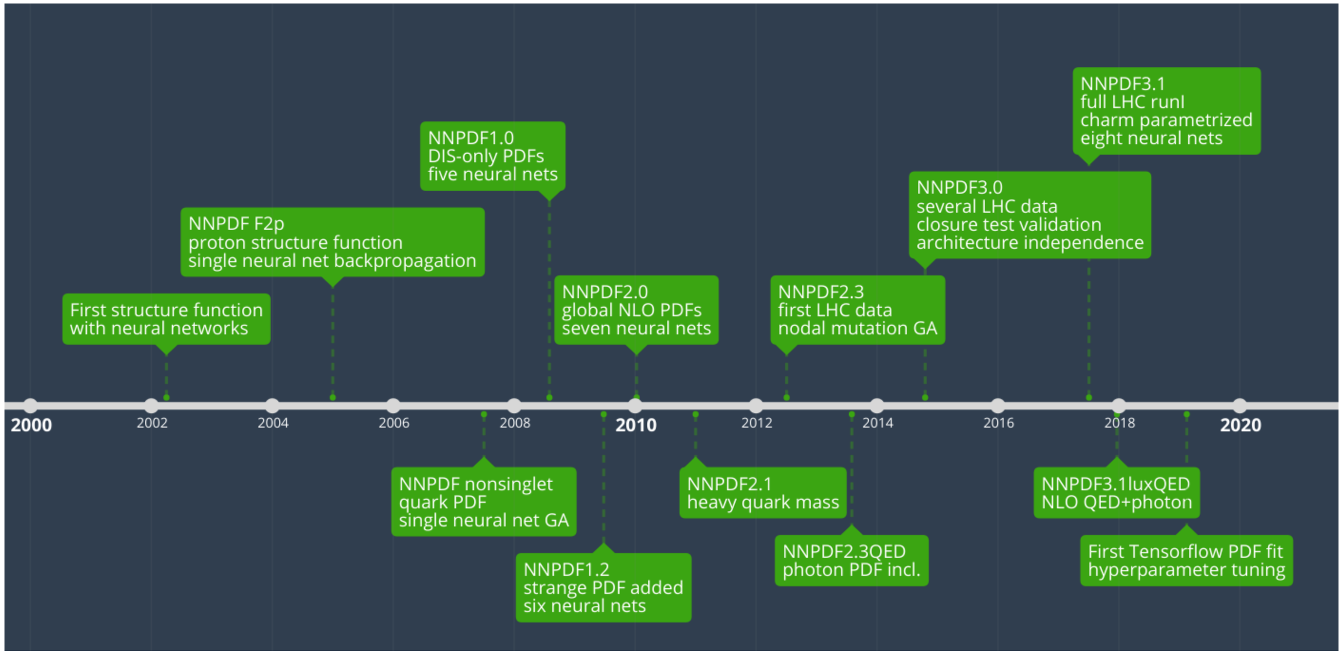
Neural network parametrization of deep-inelastic structure functions

Stefano Forte,[a] Lluís Garrido,[b] José I. Latorre[b] and Andrea Piccione[c]

[a]*INFN, Sezione di Roma Tre*
Via della Vasca Navale 84, I-00146 Rome, Italy
[b]*Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona,*
Diagonal 647, E-08028 Barcelona, Spain
[c]*INFN sezione di Genova and*
Dipartimento di Fisica, Università di Genova,
via Dodecaneso 33, I-16146 Genova, Italy

FROM THE PROOF OF CONCEPT...

...TO THE NNPDF TIMELINE



THE NNPDF APPROACH

COMBINING DATA BY MONTE CARLO

TWO MEASUREMENTS: $\mu_1 \pm \sigma_1$; $\quad \mu_2 \pm \sigma_2$

ML COMBINATION: $\bar\mu \pm \bar\sigma$; $\quad \bar\mu = \frac{\frac{\mu_1}{\sigma_1^2}+\frac{\mu_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}}$; $\bar\sigma^2 = \frac{1}{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}}$

MONTE CARLO REPRESENTATION

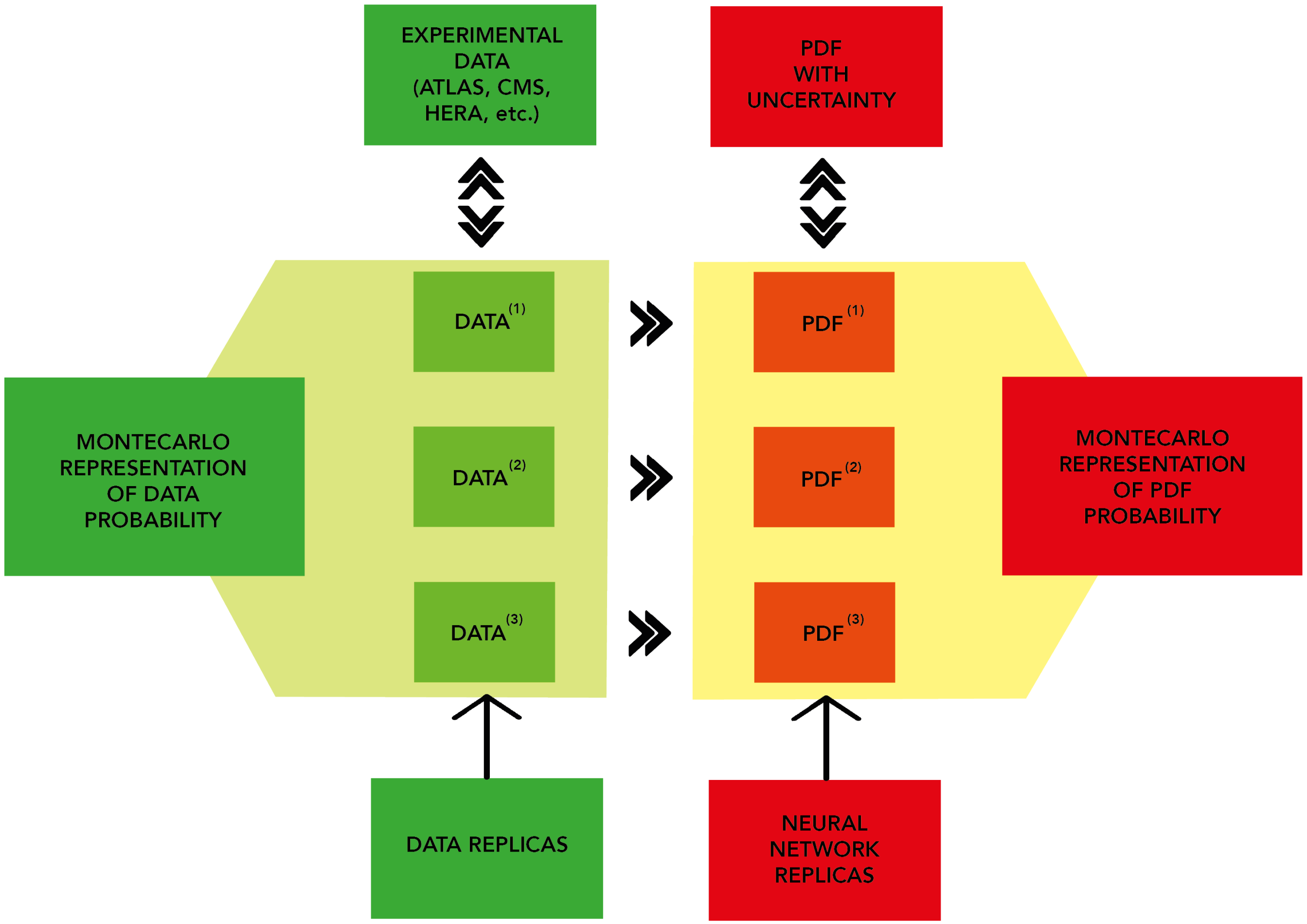
$$
\begin{array}{cccc}
\mu_1 \pm \sigma_1 & \mu_2 \pm \sigma_2 & \longrightarrow & \bar\mu + \bar\sigma \\
\updownarrow & \updownarrow & & \updownarrow \\
\mu_1^{(1)} & \mu_2^{(1)} & \longrightarrow & \bar\mu^{(1)} \\
\mu_1^{(2)} & \mu_2^{(2)} & \longrightarrow & \bar\mu^{(2)} \\
\vdots & \vdots & \vdots & \\
\mu_1^{(n)} & \mu_2^{(n)} & \longrightarrow & \bar\mu^{(n)}
\end{array}
$$

$\mu^{(i)} \Leftrightarrow$ REPLICA SAMPLE $\Leftrightarrow$ REPRESENTATION OF PROBABILITY DISTRIBUTION

THE NNPDF APPROACH

THE FUNCTIONAL MONTE CARLO

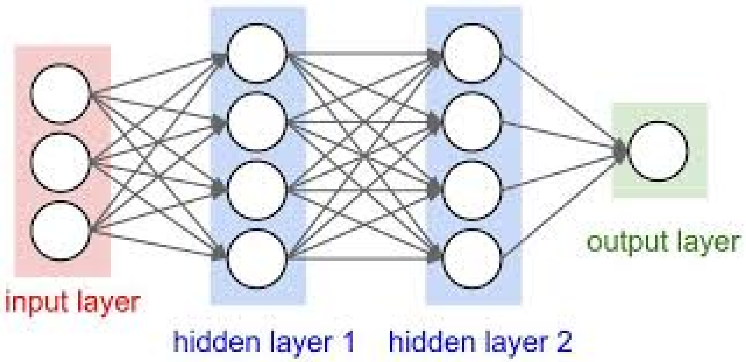
REPLICA SAMPLE OF FUNCTIONS $\Leftrightarrow$ PROBABILITY DENSITY IN FUNCTION SPACE



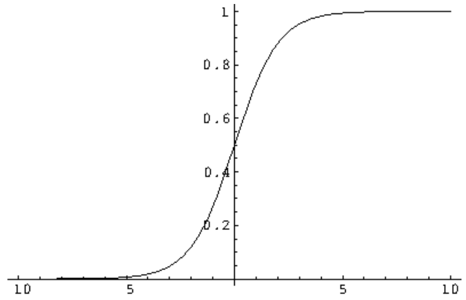
FINAL PDF SET: $f_i^{(a)}(x,\mu)$; i =up, antiup, down, antidown, strange, antistrange, charm, gluon; $j = 1, 2, \ldots N_{\rm rep}$

UNBIASED INTERPOLANTS: NEURAL NETWORKS

ARCHITECTURE



ACTIVATION FUNCTION



PARAMETERS

- WEIGHTS ω_{ij}
- THRESHOLDS θ_i

$$F_{\rm out}^{(i)}(\vec{x}_{\rm in}) = F\left(\sum_j \omega_{ij} x_{\rm in}^j - \theta_i\right)$$

SIMPLEST EXAMPLE

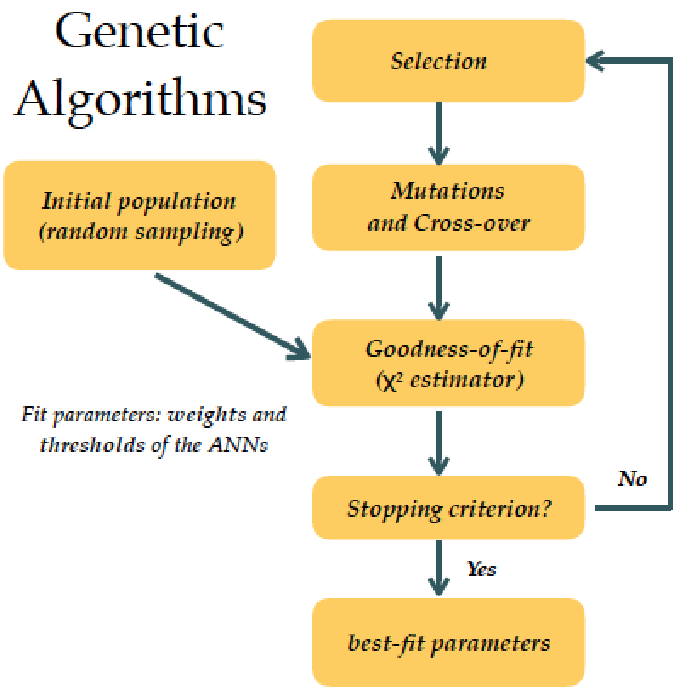
1-2-1

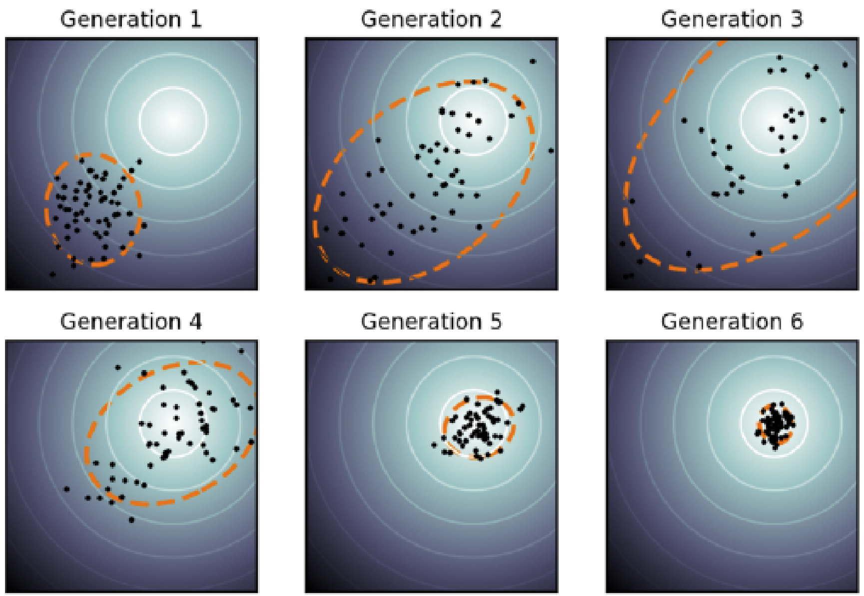
$$f(x) = \frac{1}{1+e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)} - x\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)} - x\omega_{21}^{(1)}}}}}$$

NNPDF: $2-5-3-1$ NN FOR EACH PDF: $37 \times 8 = 296$ PARAMETERS

GENETIC ALGORITHMS

- BASIC IDEA: RANDOM MUTATION OF THE NN PARAMETER
- SELECTION OF THE FITTEST

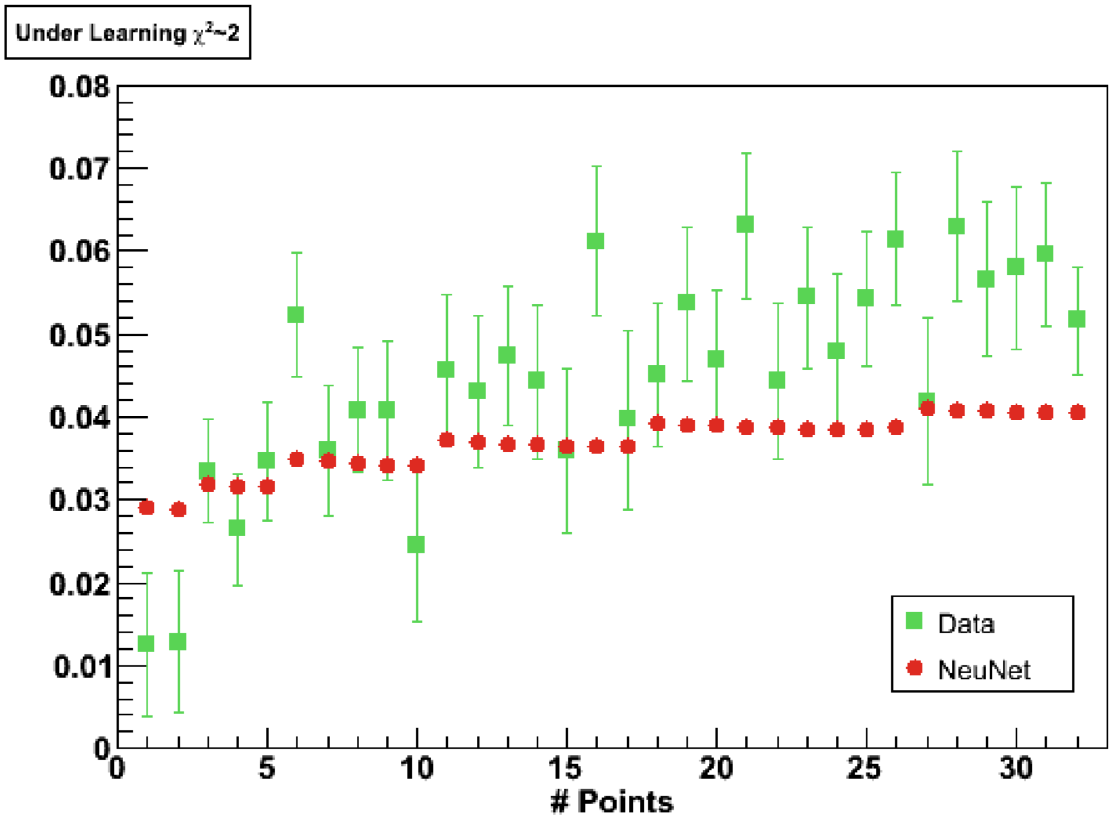




NEURAL LEARNING

- COMPLEXITY INCREASES AS THE FITTING PROCEEDS
- UNTIL LEARNING NOISE
- WHEN SHOULD ONE STOP?

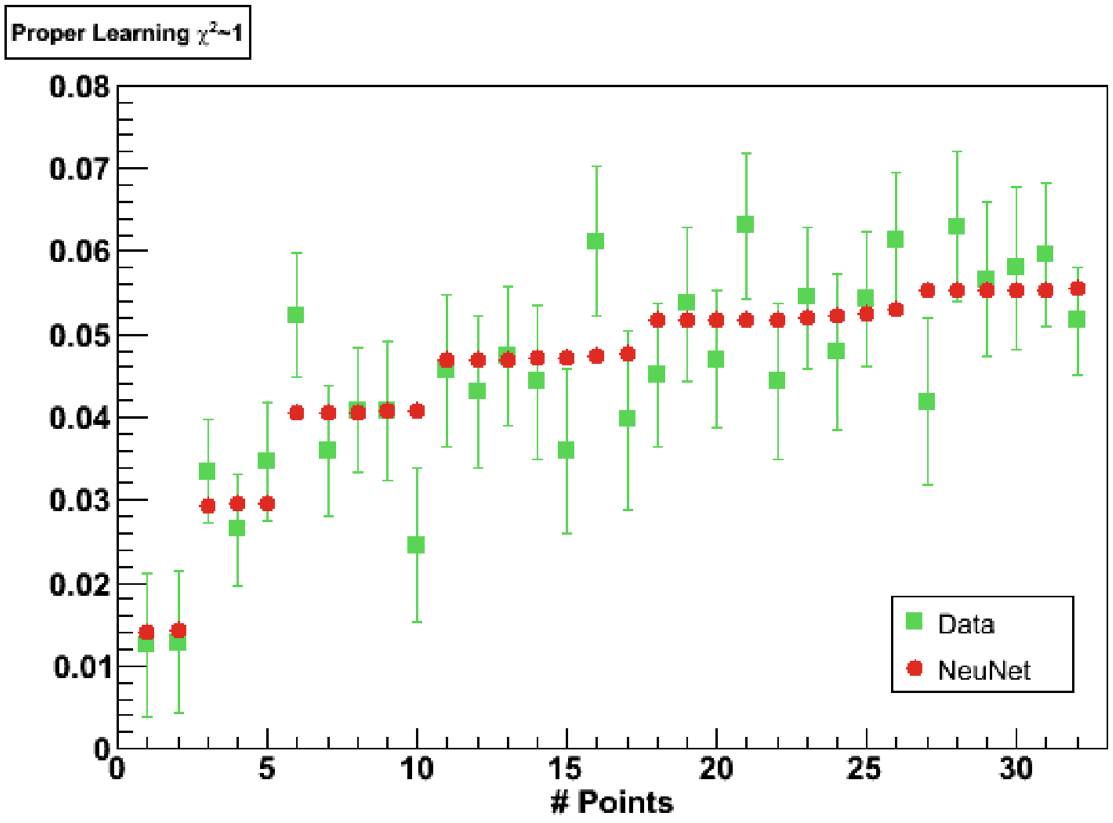
UNDERLEARNING



NEURAL LEARNING

- COMPLEXITY INCREASES AS THE FITTING PROCEEDS
- UNTIL LEARNING NOISE
- WHEN SHOULD ONE STOP?

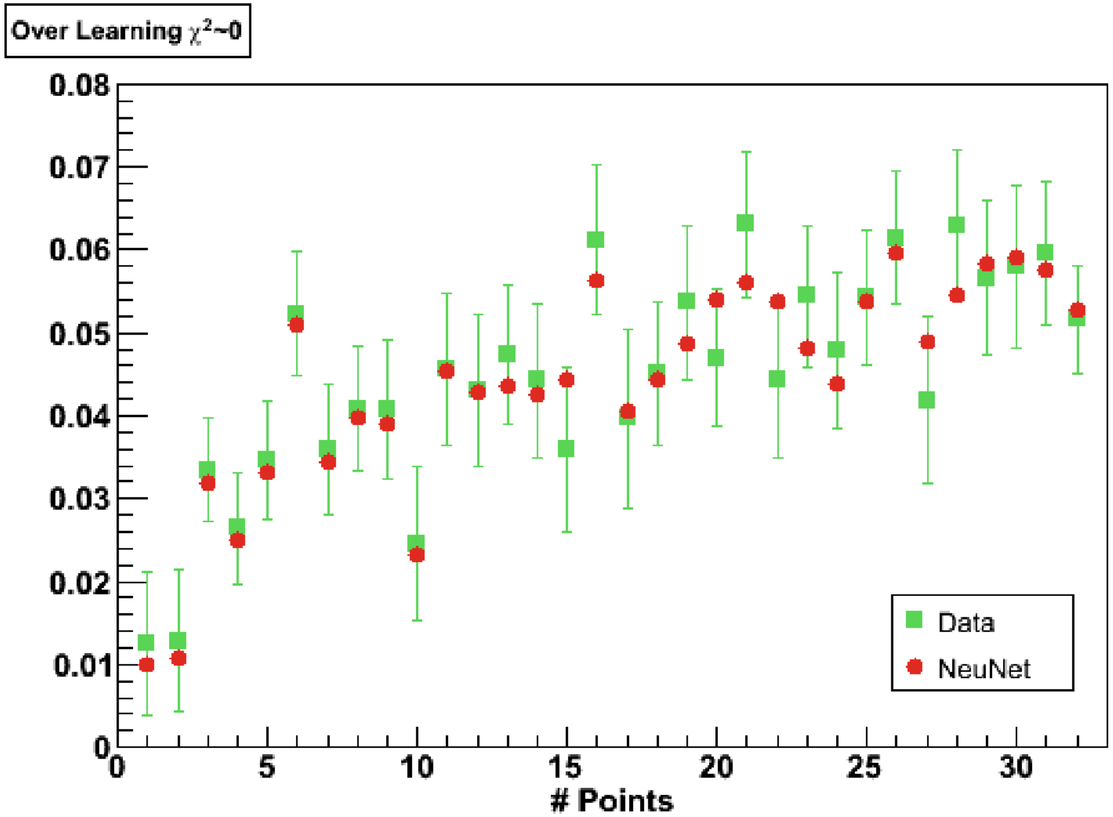
PROPER LEARNING



NEURAL LEARNING

- COMPLEXITY INCREASES AS THE FITTING PROCEEDS
- UNTIL LEARNING NOISE
- WHEN SHOULD ONE STOP?

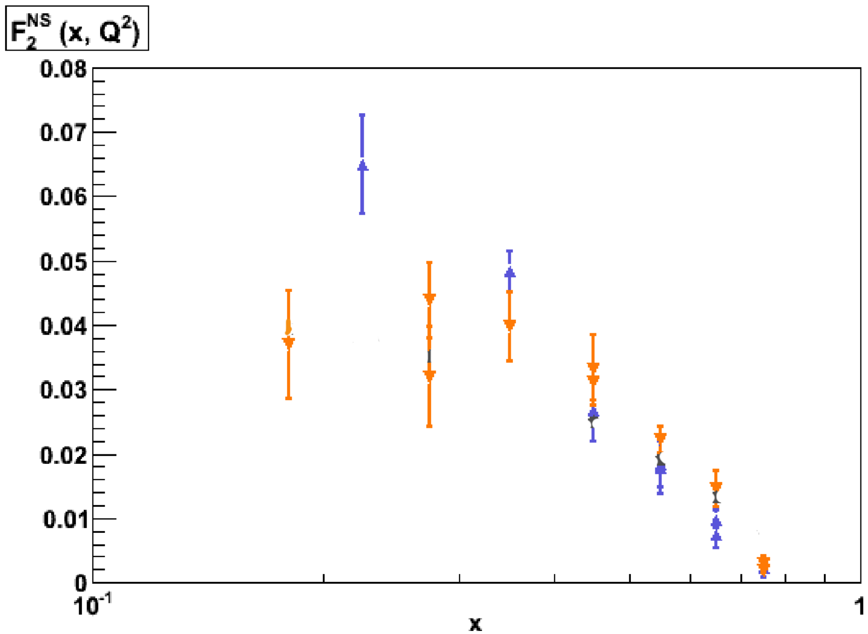
OVERLEARNING



OPTIMAL FIT: CROSS-VALIDATION

GENETIC MINIMIZATION:
AT EACH GENERATION, 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

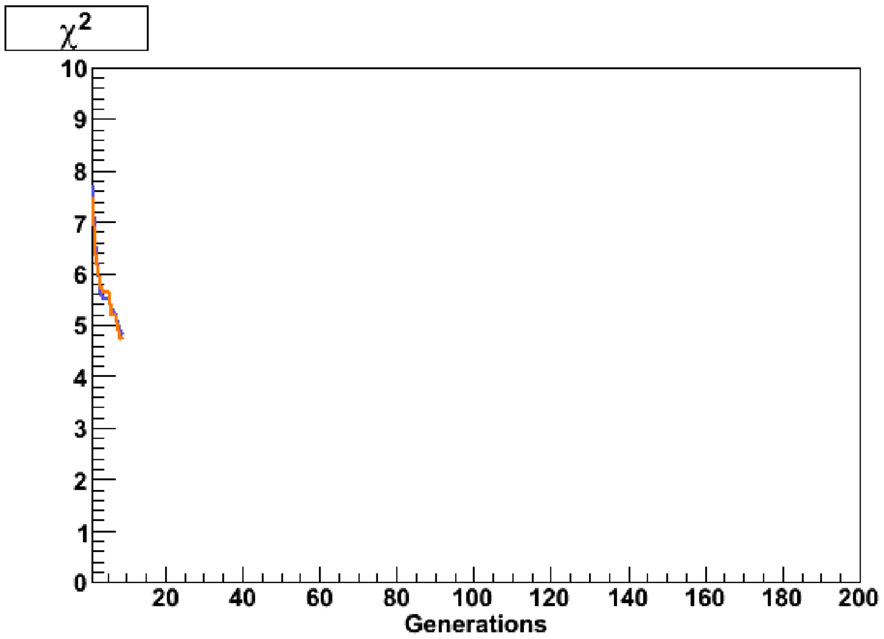


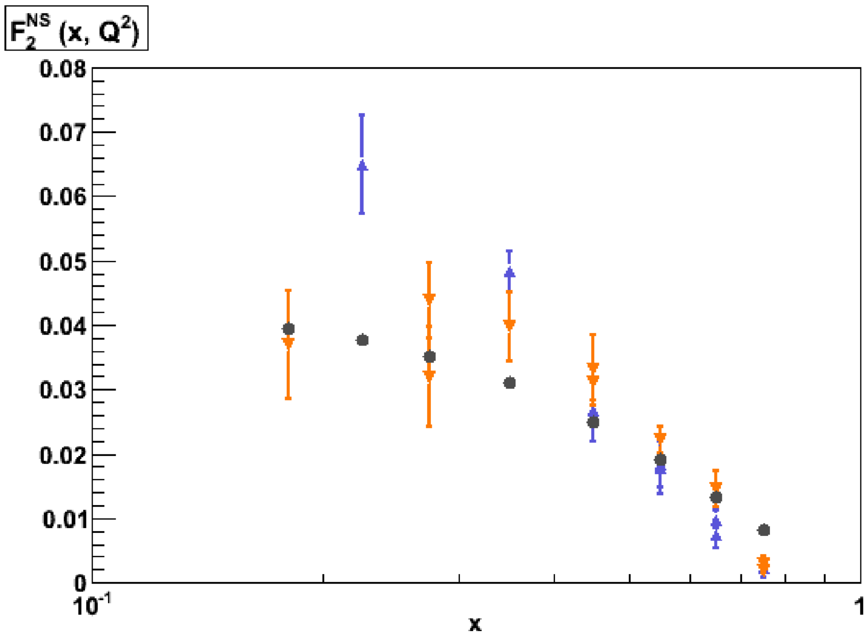
OPTIMAL FIT: CROSS-VALIDATION

GENETIC MINIMIZATION:
AT EACH GENERATION, 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!



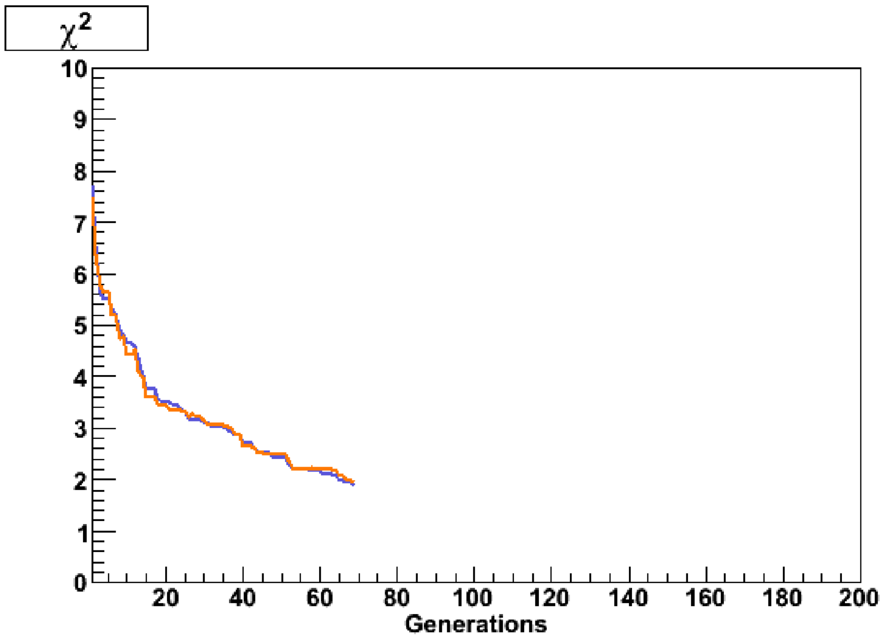


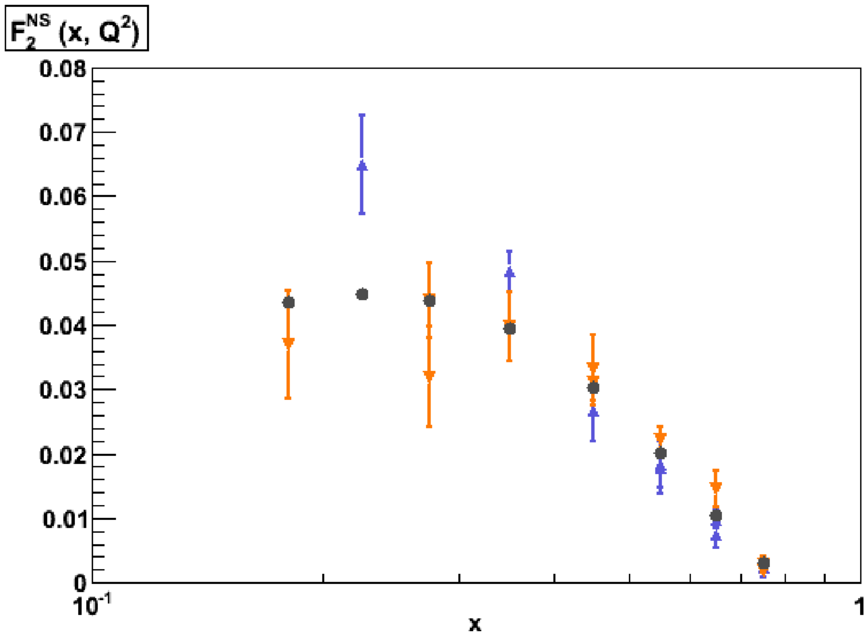
OPTIMAL FIT: CROSS-VALIDATION

GENETIC MINIMIZATION:
AT EACH GENERATION, 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!



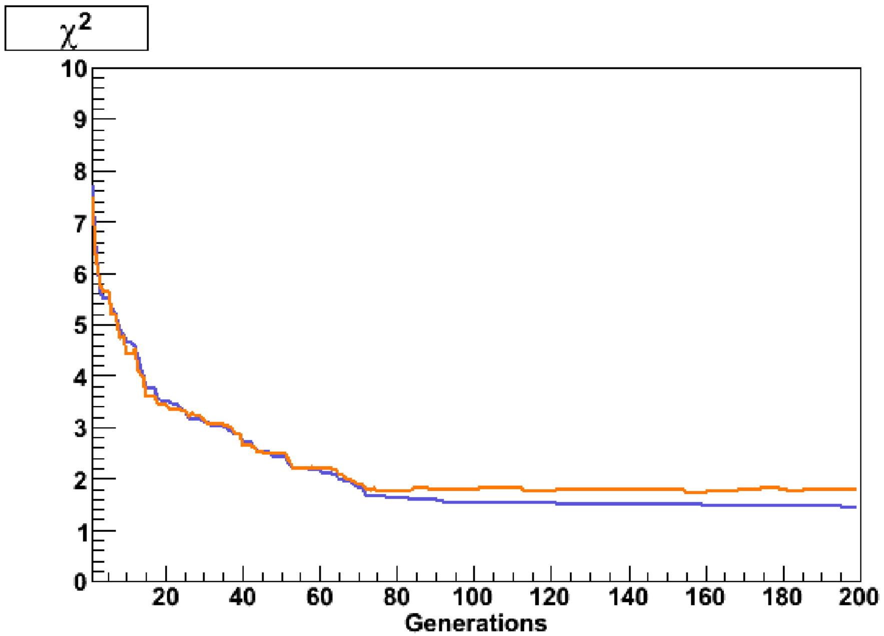


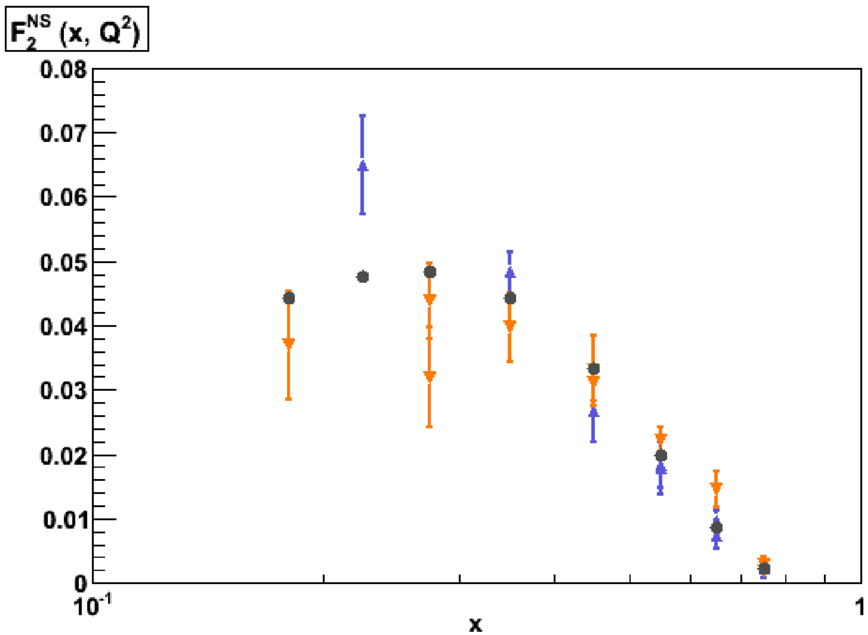
OPTIMAL FIT: CROSS-VALIDATION

GENETIC MINIMIZATION:
AT EACH GENERATION, 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

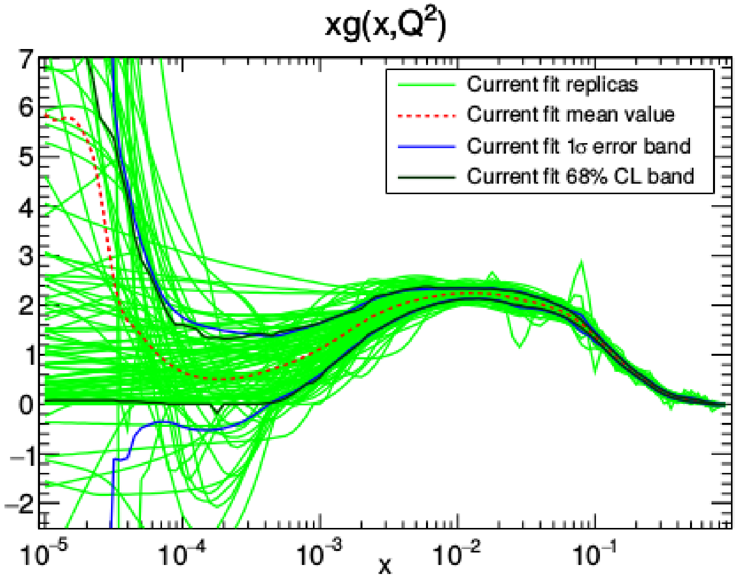
TOO LATE!



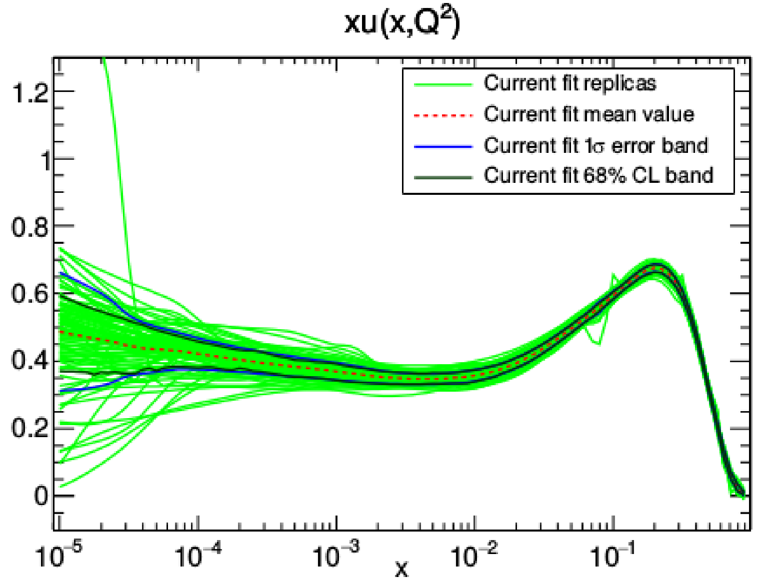


CURRENT STATUS
NNPDF3.1 (2017)

GLUON



UP



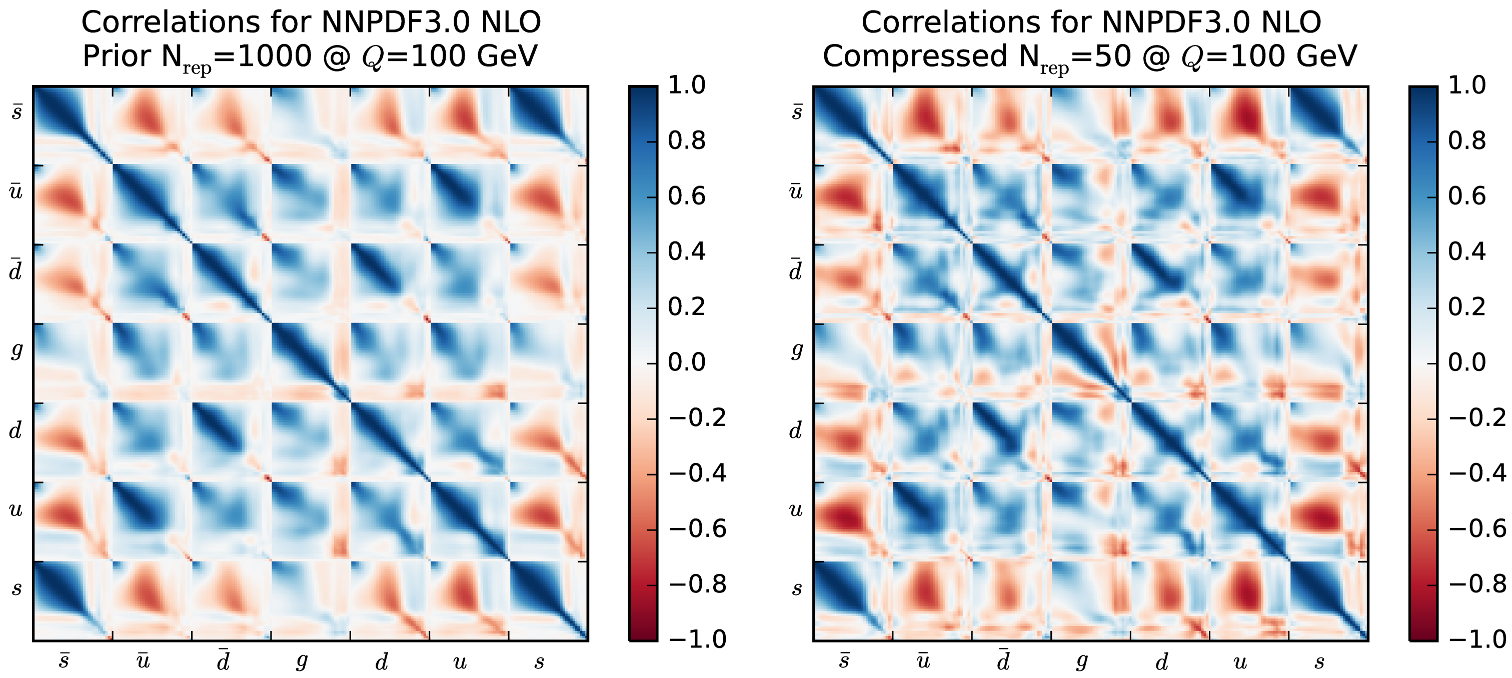
- PROBABILITY DISTRIBUTION OF PDFs ↔ ENSEMBLE OF REPLICAS
- EXPECTED CENTRAL VALUE ↔ MEAN
- UNCERTAINTY ↔ STANDARD DEVIATION
- 68% C.L. ALSO SHOWN

UNSUPERVISED LEARNING

OPTIMIZATION

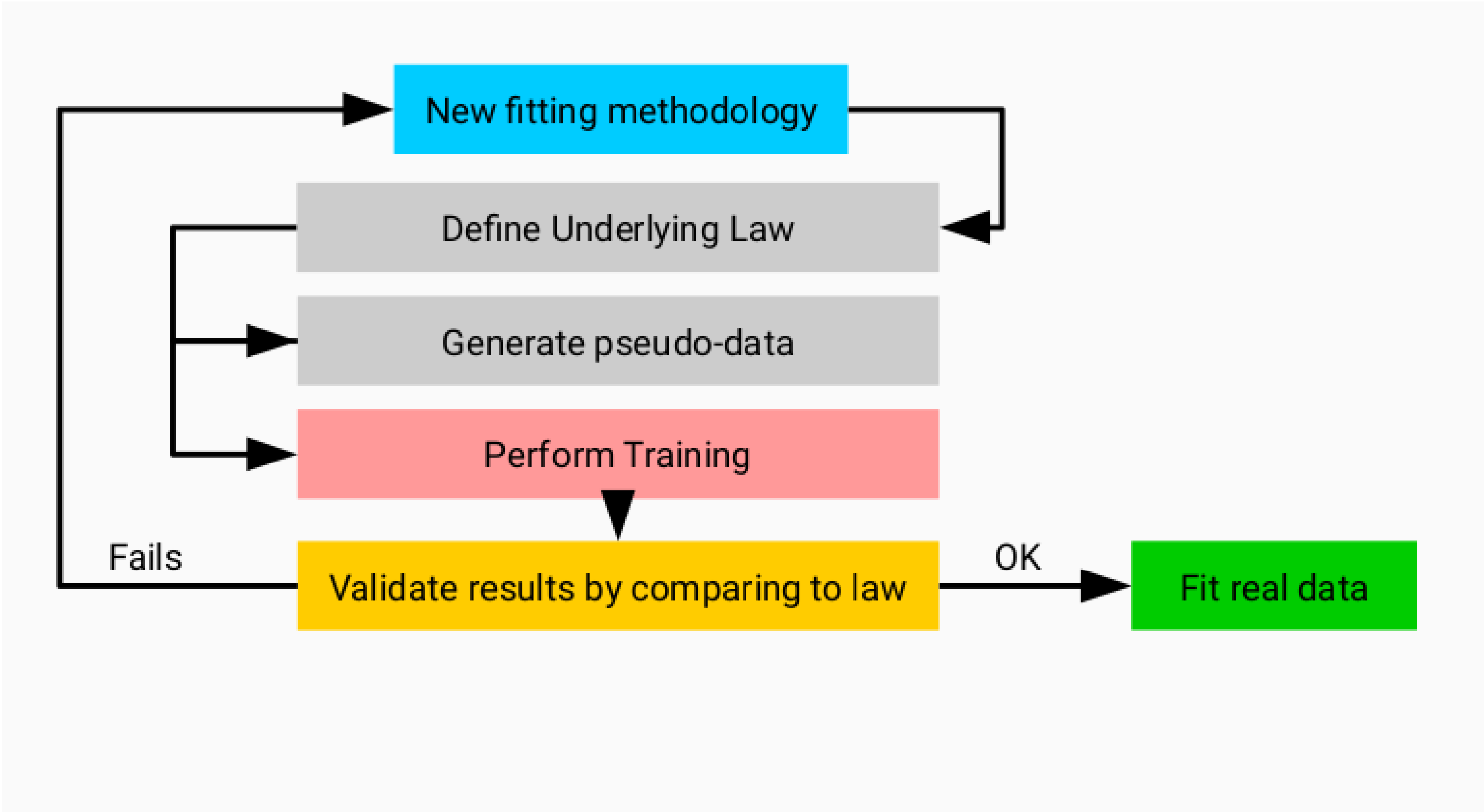
- HOW TO MAXIMIZE ACCURACY?
- LARGE (PRIOR) REPLICA SET
- GENETIC SELECTION $\Rightarrow$ OPTIMIZATION OF STATISTICAL INDICATORS (KULLBACK-LEIBLER DIVERGENCE)
- 50 OPTIMIZES REPLICAS $\Leftrightarrow$ 1000 STARTING REPLICAS

CORRELATION MATRIX



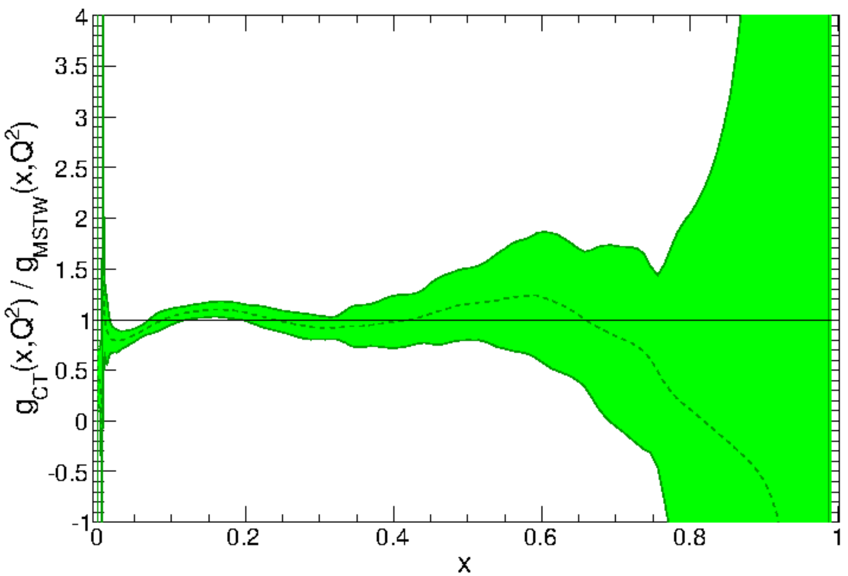
HOW DO WE KNOW THAT WE GOT THE RIGHT ANSWER

CLOSURE TEST

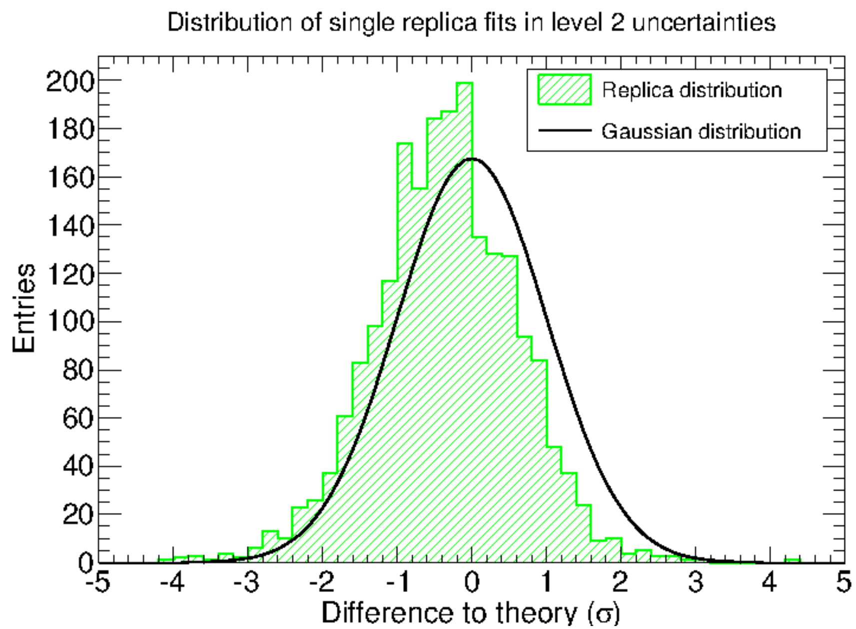


FIRST CLOSURE TEST (NNPDF3.0; 2014)

THE GLUON: RESULT/"TRUTH"



NORMALIZED DISTRIBUTION OF DEVIATIONS

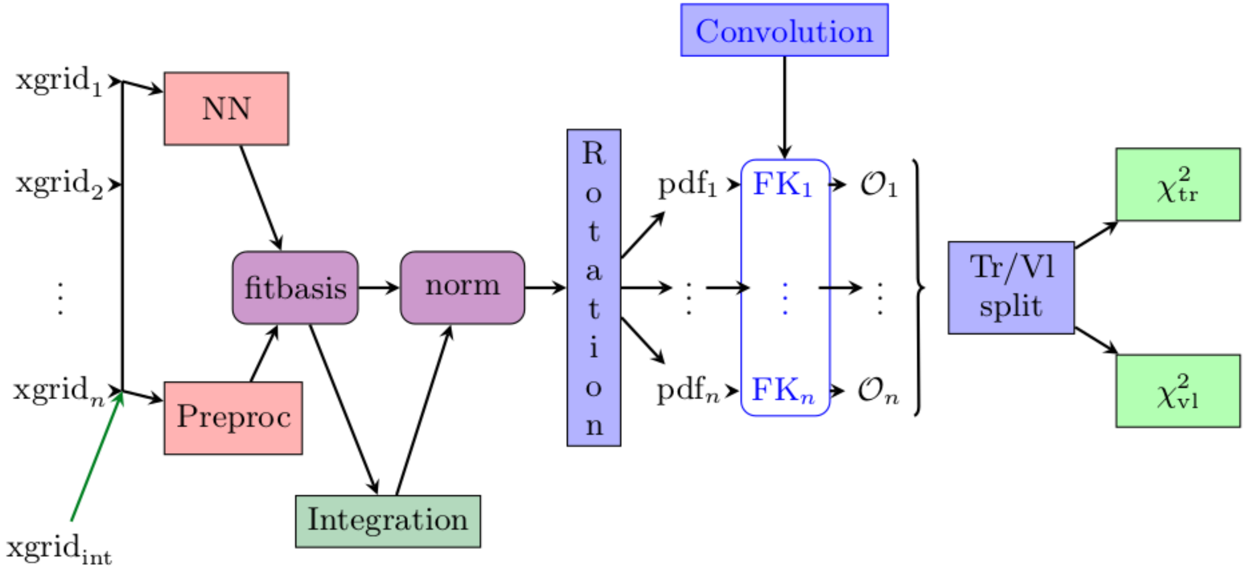


- THE METHODOLOGY IS FAITHFUL
- BUT IS IT OPTIMAL?

FITTING THE METHODOLOGY
THE N3FIT PROJECT

HOW DO WE KNOW THAT THE METHODOLOGY IS THE BEST?
"ACCUMULATED WISDOM" INEFFICIENT AND SLOW

CHANGE OF PHILOSOPHY $\Rightarrow$ DETERMINISTIC MINIMIZATION (GRADIENT DESCENT)
GO FOR THE ABSOLUTE MINIMUM, AND (HYPER)OPTIMIZE

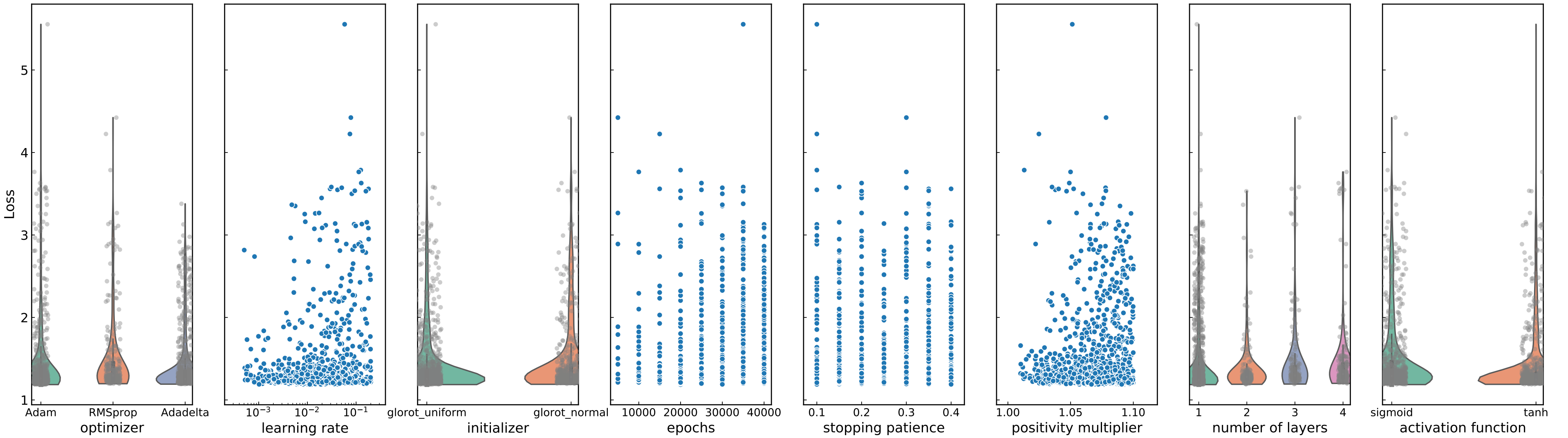


(Carrazza, Cruz-Martinez, 2019)

- PYTHON-BASED KERAS + TENSORFLOW FRAMEWORK
- EACH BLOCK INDEPENDENT LAYER
- CAN VARY ALL ASPECT OF METHODOLOGY

FITTING THE METHODOLOGY

HYPEROPTIMIZATION SCANS



HYPEROPT PARAMETERS

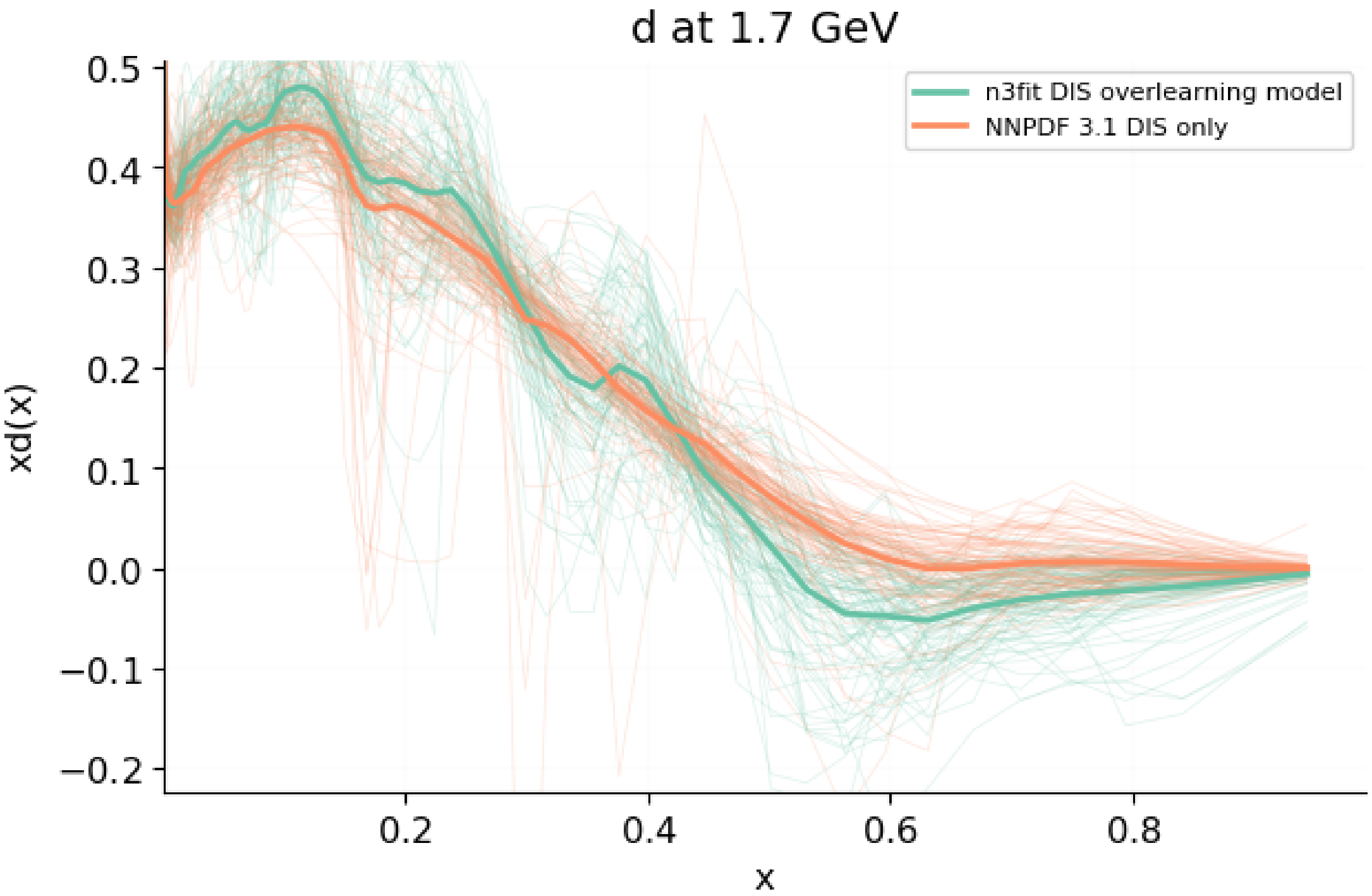
Neural Network	Fit options
Number of layers (*)	Optimizer (*)
Size of each layer	Initial learning rate (*)
Dropout	Maximum number of epochs (*)
Activation functions (*)	Stopping Patience (*)
Initialization functions (*)	Positivity multiplier (*)

- SCAN PARAMETER SPACE
- OPTIMIZE FIGURE OF MERIT
- BAYESIAN UPDATING

FITTING THE METHODOLOGY

THE OVERFITTING PROBLEM

DOWN QUARK: HYPEROPTIMIZED VS. STANDARD



- OVERFITTING $\Rightarrow \chi^2_{\mathrm{train}} << \chi^2_{\mathrm{valid}}$!! & WIGGLY PDFS
- CORRELATIONS BETWEEN DATA IN A SET

WHAT HAPPENED?

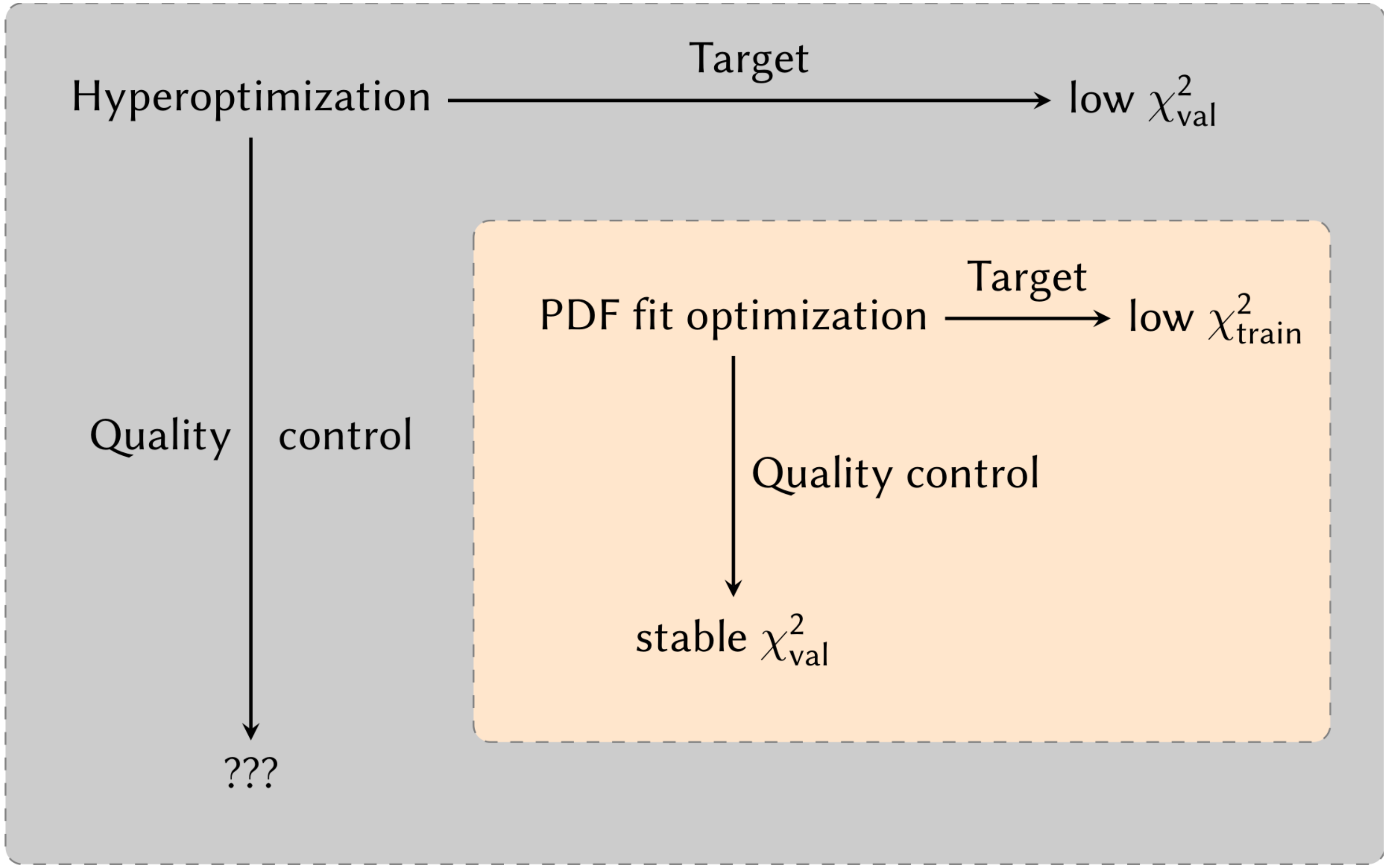
OPTIMIZATION

$$\begin{tikzcd} \text{PDF fit optimization} \arrow[r, "\text{Target}"] \arrow[d, "\text{Quality control}"] & \text{low } \chi^2_{\text{train}} \\ \text{stable } \chi^2_{\text{val}} & \end{tikzcd}$$

CROSS-VALIDATION SELECTS THE OPTIMAL MINIMUM

WHAT HAPPENED?

HYPEROPTIMIZATION

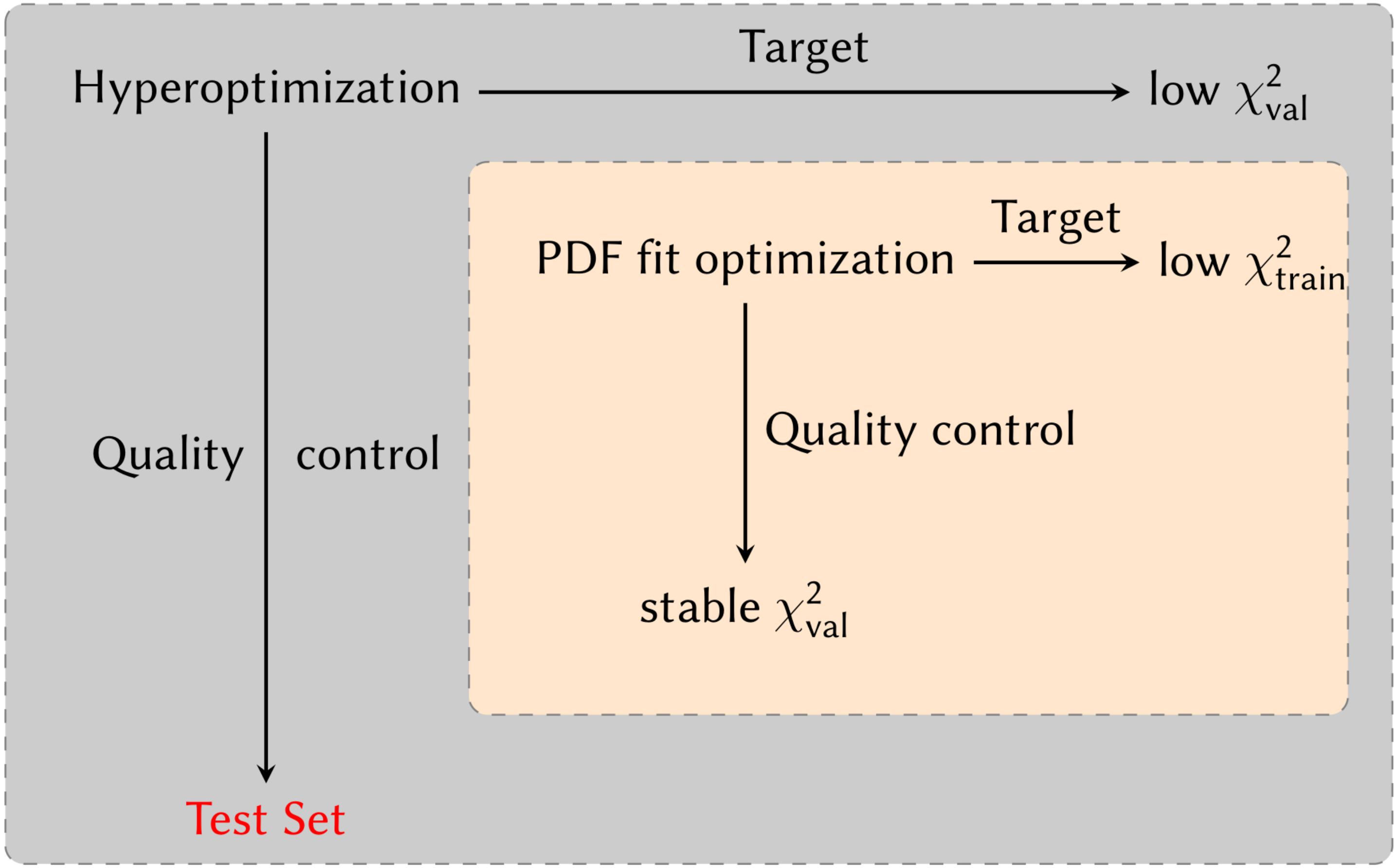


WE ARE MISSING A SELECTION CRITERION

MACHINE LEARNING

THE SOLUTION

TUNED HYPEROPTIMIZATION



COMPARE TO A A TEST SET (NEW SET OF DATA PREVIOUSLY NOT USED AT AL)

TESTS GENERALIZATION POWER

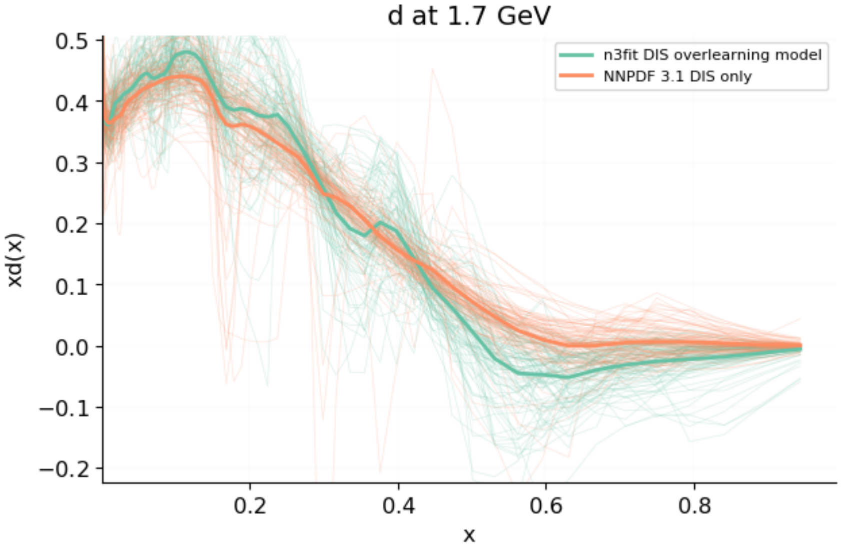
THE TEST SET METHOD

- COMPLETELY UNCORRELATED TEST SET
- OPTIMIZE ON WEIGHTED AVERAGE OF VALIDATION AND TEST
  $\Rightarrow$ NO OVERLEARNING

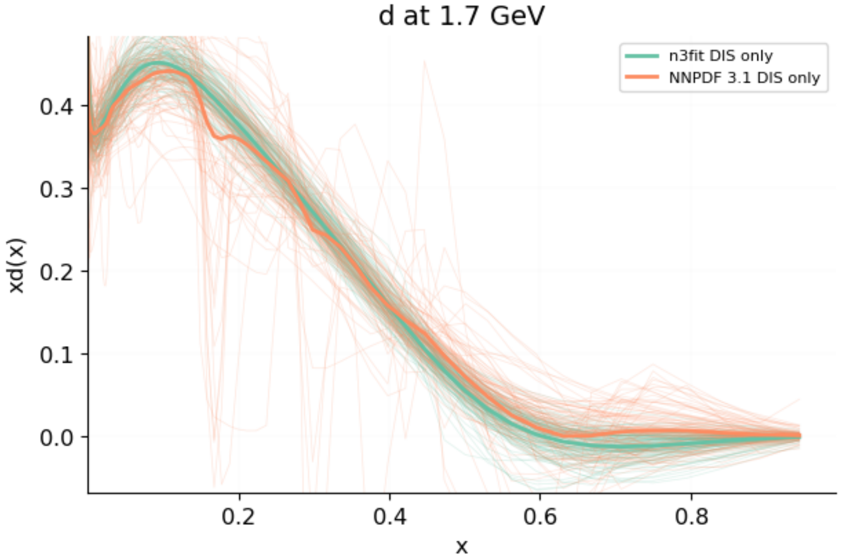
OPTIMIZED PDFS

DOWN QUARK

N3 OVERFIT vs NNPDF3.1
d QUARK



N3FIT vs NNPDF3.1

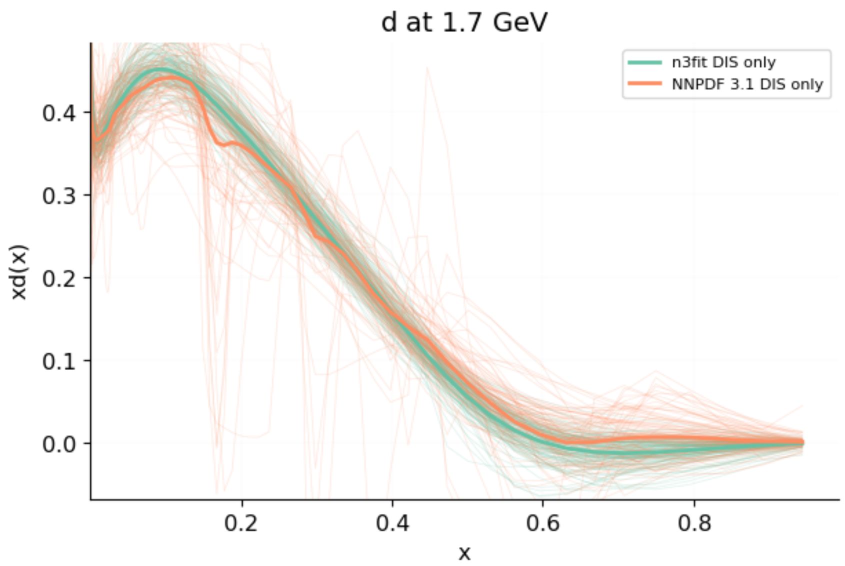


THE TEST SET METHOD

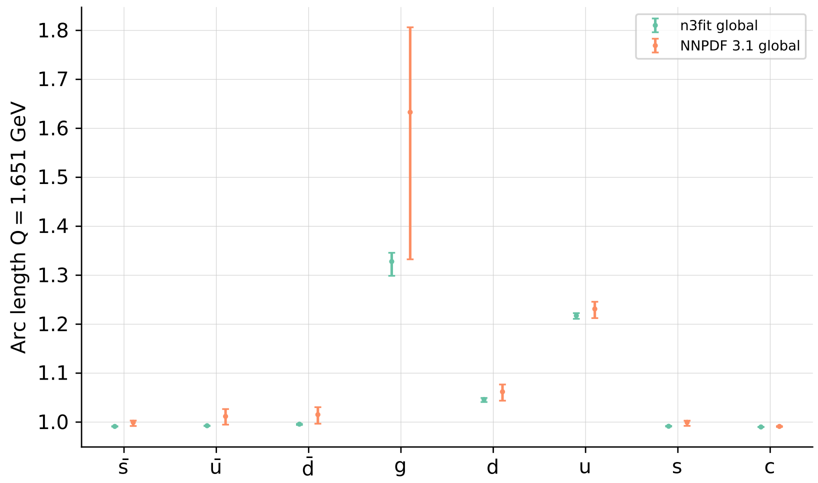
N3 OVERFIT vs NNPDF3.1

N3FIT vs NNPDF3.1

d QUARK



ARCLENGTHS



- NO OVERFITTING
- COMPARED TO NNPDF3.1
  - MUCH GREATER STABILITY $\Rightarrow$ FEWER REPLICAS FOR EQUAL ACCURACY
  - UNCERTAINTIES SOMEWHAT REDUCED

FITTING THE METHODOLOGY

WHAT IS "PROPER LEARNING"?

FORECASTING AN UNKNOWN TRUTH $\Rightarrow$ WHAT IS "OPTIMAL"?

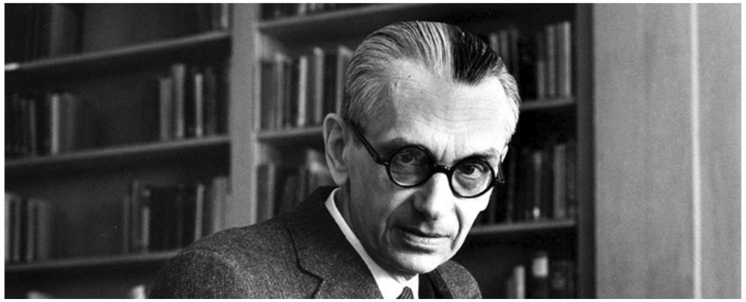
MENU nature

NEWS · 08 JANUARY 2019

Machine learning leads mathematicians to unsolvable problem

Simple artificial-intelligence problem puts researchers up against a logical parc
by famed mathematician Kurt Gödel.

Davide Castelvecchi



SOME POSSIBLE ANSWERS/CRITERIA

- PASS A CLOSURE TEST
- PASS A "FUTURE TEST":
  GENERALIZE TO CURRENT DATA BASED ON PAST DATA
- REPRODUCE THE EXPECTED STATISTICAL PROPERTIES:
  ONE $\sigma \Leftrightarrow \Delta\chi^2 = 1$
- SATISFY THEORETICAL PREJUDICE?

REINFORCEMENT LEARNING?

THE WORK OF MANY PEOPLE

NNPDF collaboration and N^{3}PDF team meeting,
Varenna, Italy, September 2019

“Io stimo più il trovare un vero, benché di cosa leggiera, che il disputar lungamente delle massime questioni senza verità nissuna”

“I am more interested in uncovering a fact, however trifling, than to dispute at length about profound questions devoid of any truth”

Galileo Galilei, letter to Tommaso Campanella