

# DALLE PARTICELLE ELEMENTARI ALL'INTELLIGENZA ARTIFICIALE E RITORNO

STEFANO FORTE  
UNIVERSITÀ DI MILANO & INFN



UNIVERSITÀ DEGLI STUDI DI MILANO  
DIPARTIMENTO DI FISICA

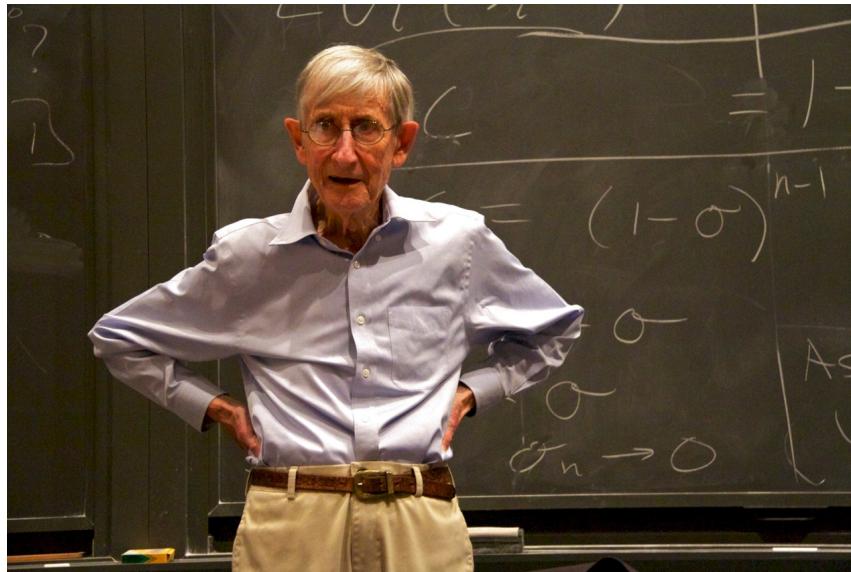


CISF 2019

MILANO, 6 MARZO 2019

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

# FISICA DELLE PARTICELLE E LHC



“There are historical reasons not to expect too much from the LHC. (...)

There have been sixteen important discoveries” (in HEP)

“between 1945 and 2008:

four discoveries on the energy frontier,

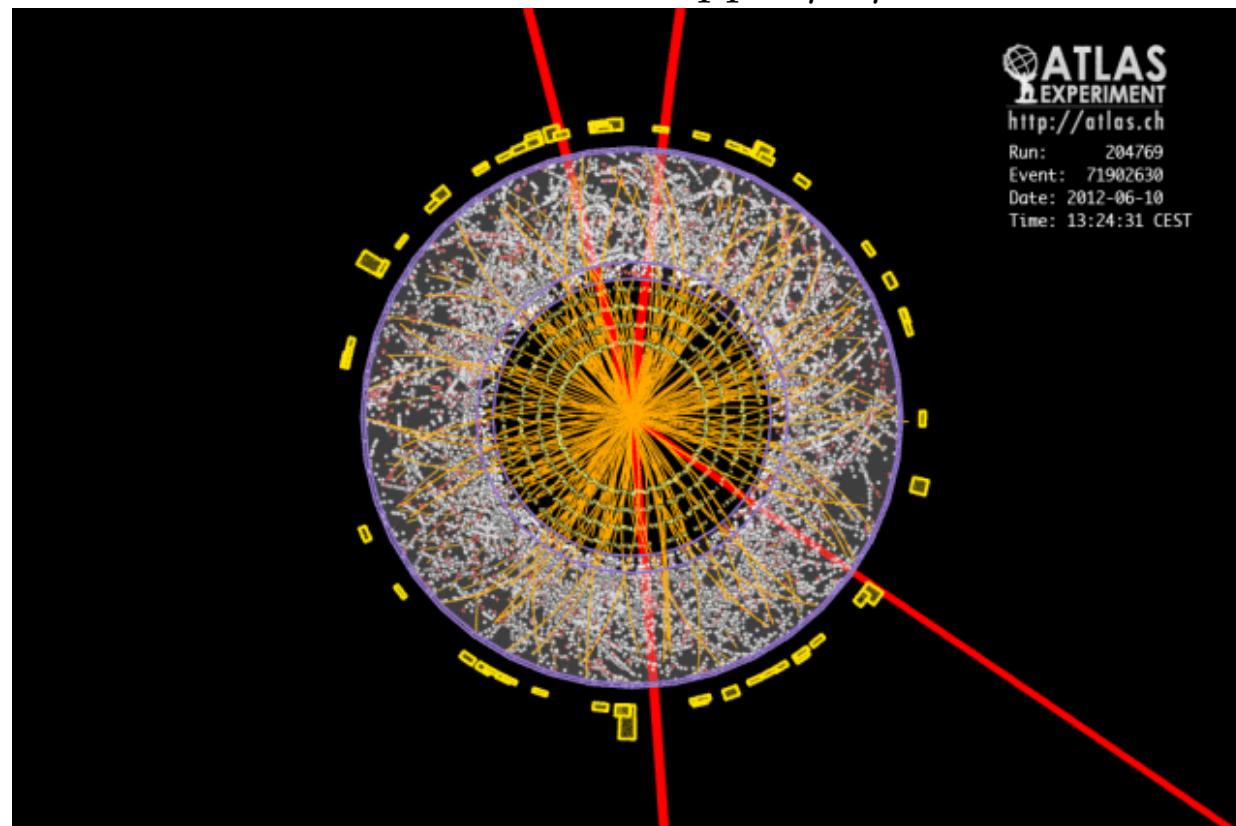
four on the rarity frontier,

eight on the accuracy frontier”

# QUAL È IL PROBLEMA?

“There are two reasons to be skeptical about the importance of the LHC:  
one technical and one historical”.

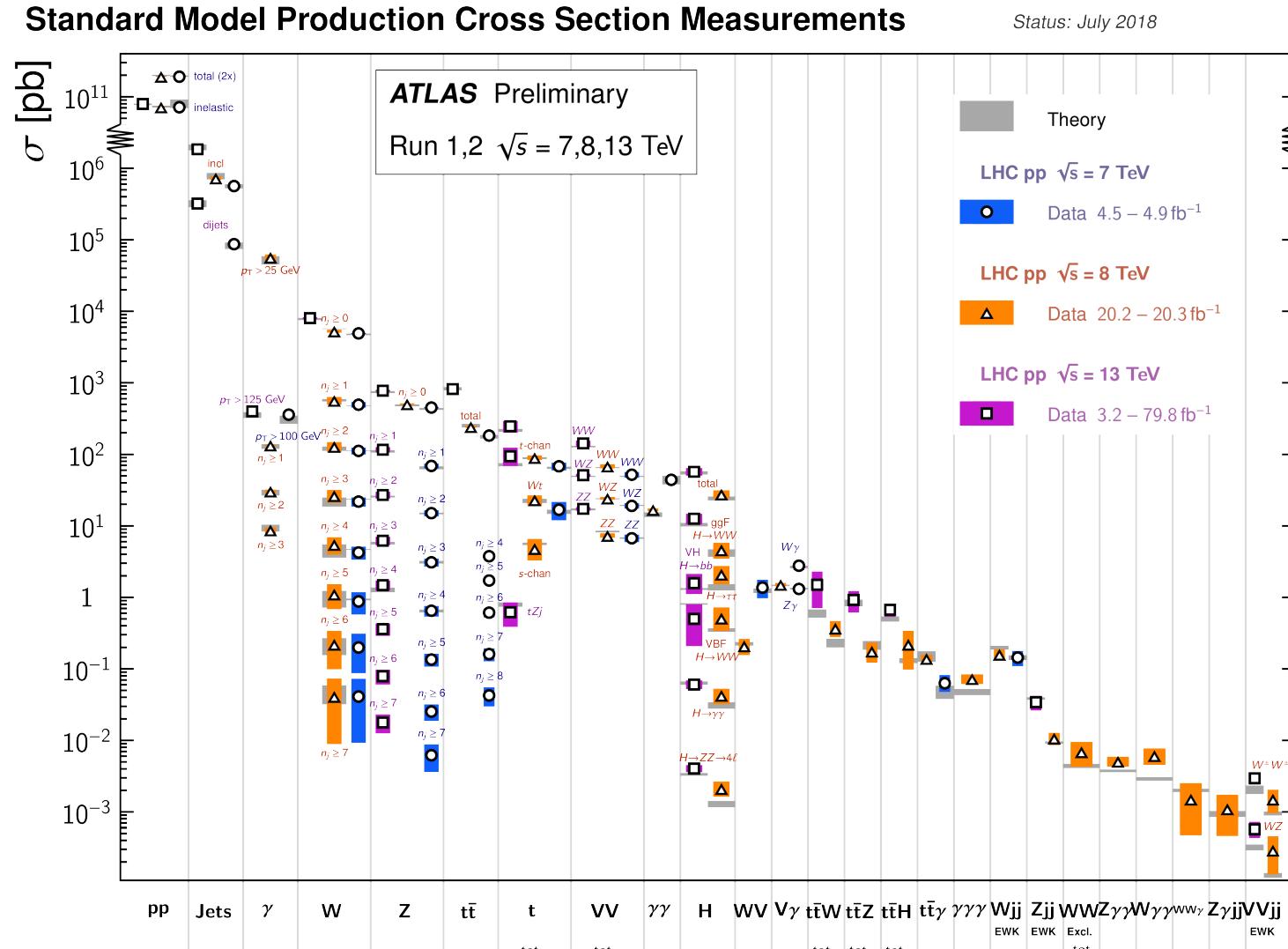
Un bosone di Higgs prodotto in collisioni  $pp$ ,  
decade in due coppie  $\mu^+\mu^-$



“The technical weakness of the LHC arises from the nature of the collisions that it studies.  
These are collisions of protons with protons, and they have the unfortunate habit of being  
messy”

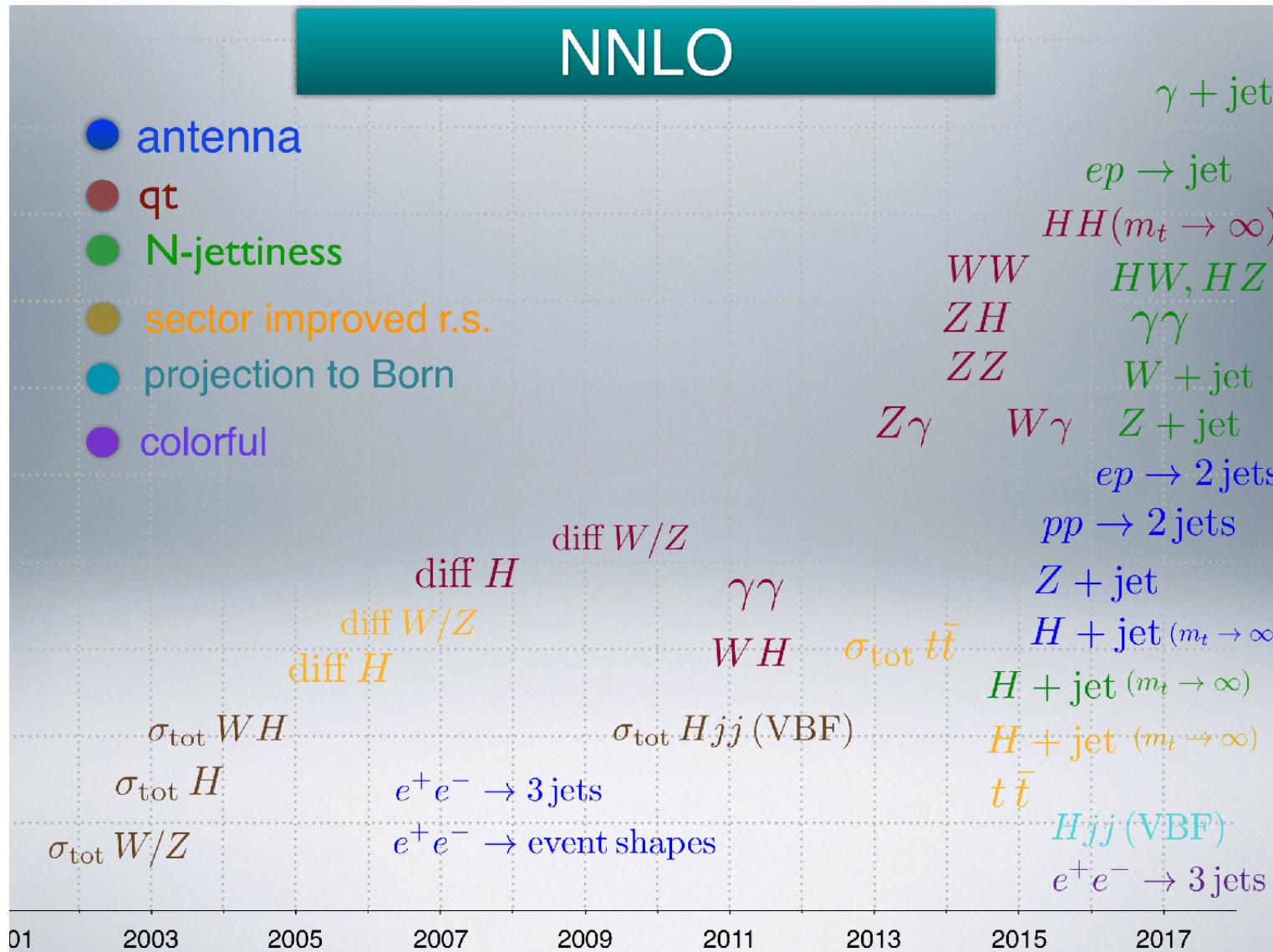
Freeman Dyson, 2008

# COME SONO ANDATE LE COSE VERAMENTE? MISURA DI PRECISIONE DI PROCESSI DEL MODELLO STANDARD A LHC



# DOVE HA SBAGLIATO DYSON?

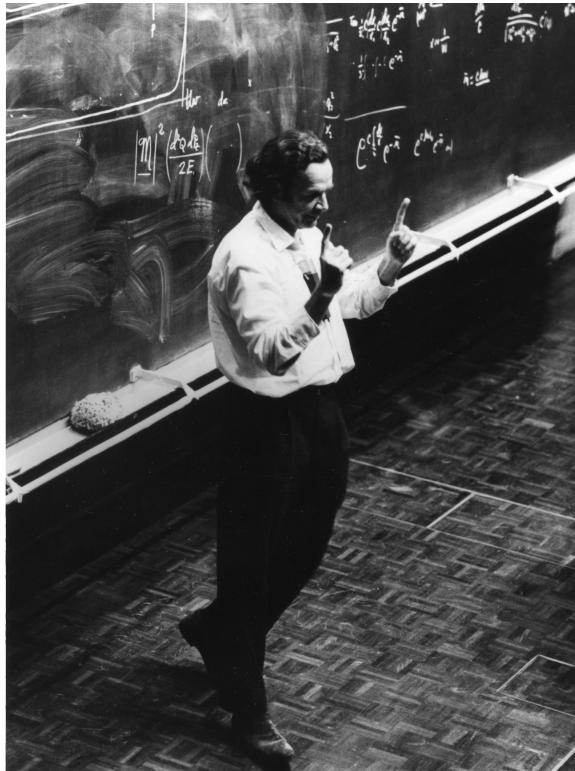
## CALCOLI AL TERZO ORDINE PERTURBATIVO NELL'INTERAZIONE FORTE



(G. Heinrich, 2017)

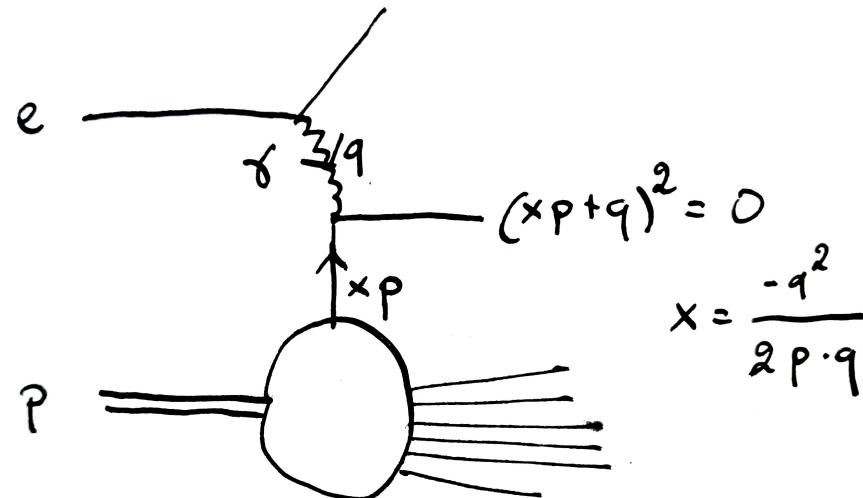
# CALCOLI PERTURBATIVI? IL “MIRACOLO” DELLA FATTORIZZAZIONE

## I “PARTONI” DI FEYNMAN



Feynman spiega il modello  
a partoni al CERN nel 1970

RADIOGRAFIA DEL PROTONE  
CON UN FOTONE DI PICCOLA LUNGHEZZA D’ONDA  $\lambda$ :

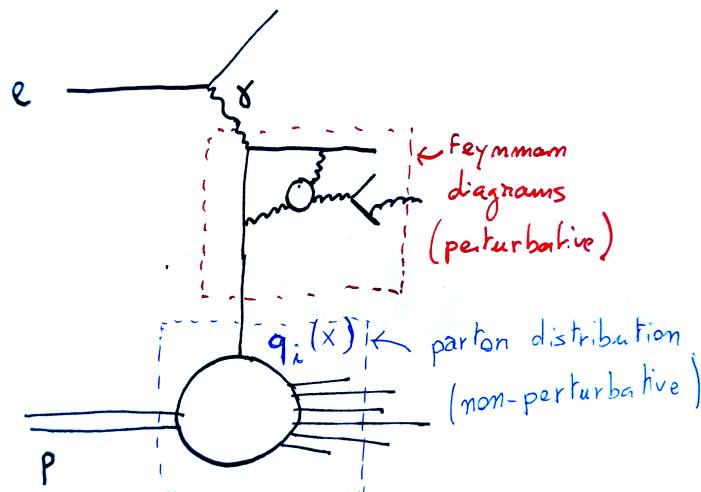


- SUPPONIAMO CHE IL FOTONE COLPISCA UN “**PARTONE LIBERO E SENZA MASSA** (QUARK, GLUONE) CHE **PORTA UNA FRAZIONE**  $x$  DELL’ENERGIA DEL PROTONE  $p_{\text{quark}} = x P_{\text{proton}}$
- IL VALORE DI  $x$  È **MISURABILE**: FISSATO DALLA CINEMATICA
- LA **SEZIONE D’URTO** È **PROPORZIONALE ALLA PROBABILITÀ**  $q_i(x)$  DI ESTRARRE UN PARTONE DI TIPO  $i$  CON FRAZIONE DI ENERGIA  $x$  NEL PROTONE

# IL “MIRACOLO” DELLA FATTORIZZAZIONE

## LA QCD PERTURBATIVA

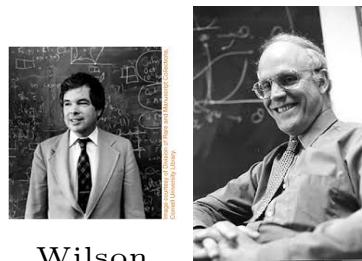
RADIOGRAFIA DEL PROTONE  
CON UN FOTONE  
DI PICCOLA LUNGHEZZA D’ONDA  $\lambda$ :



- IL MODELLO A PARTONI È IL PRIMO ORDINE DI UNA SERIE PERTURBATIVA
- A OGNI ORDINE PERTURBATIVO LA SEZIONE D’URTO FATTORIZZA:  
$$\sigma_{\text{misurata}} = \hat{\sigma}_{\text{perturbativa}} \otimes q_i$$
- SI PUÒ DEMONSTRARE RIGOROSAMENTE CON LO SVILUPPO DI WILSON

# LA QCD PERTURBATIVA

URTO PROTONE-PROTONE  
CON PICCOLA SCALA DI DISTANZA  $\lambda$ :



Wilson



Gross



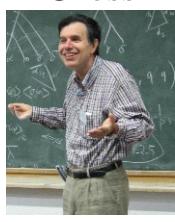
Wilczek



Politzer



Altarelli



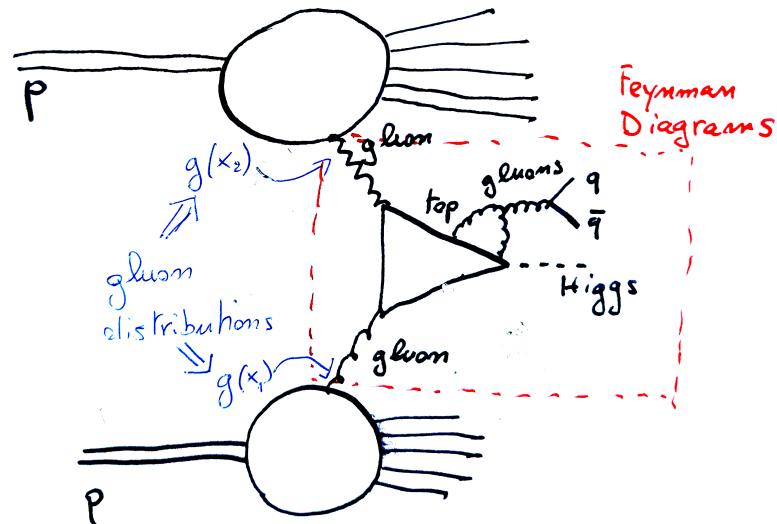
Parisi



Collins



Sterman

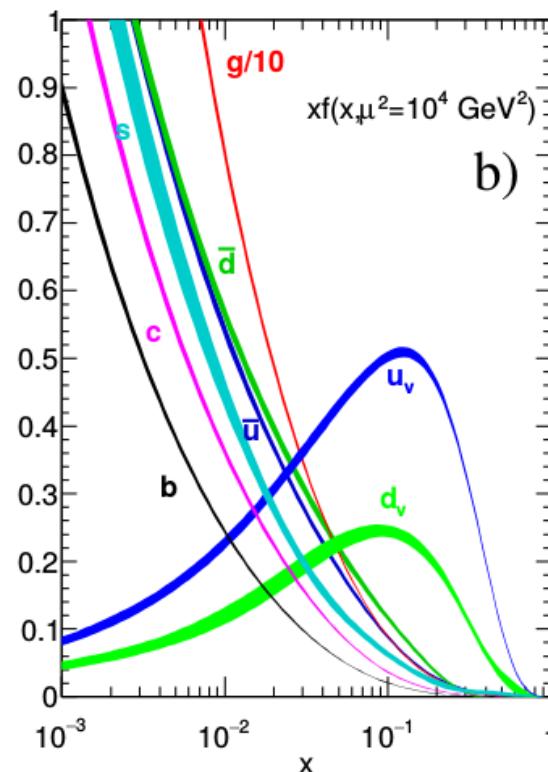
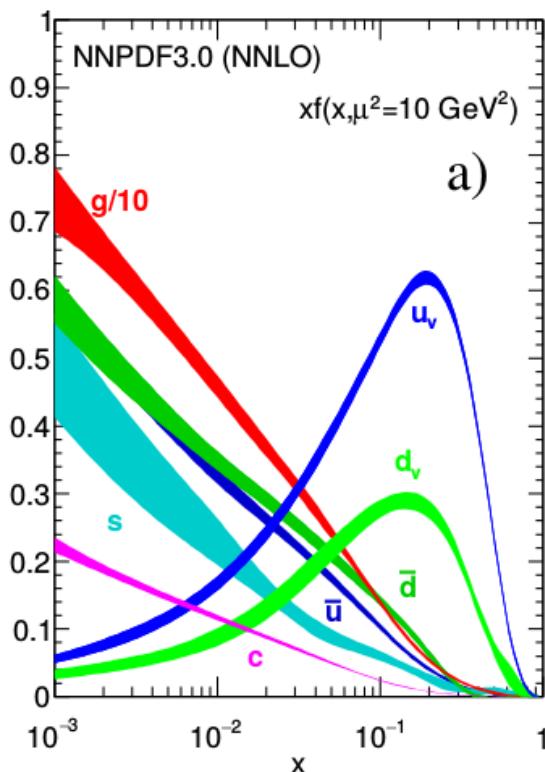


$$\lambda \sim \frac{1}{M_{\text{Higgs}}}$$

produzione di Higgs al **quarto** ordine perturbativo

- IL MODELLO A PARTONI È IL PRIMO ORDINE DI UNA SERIE PERTURBATIVA
- A OGNI ORDINE PERTURBATIVO LA SEZIONE D'URTO FATTORIZZA:  
 $\sigma_{\text{misurata}} = \hat{\sigma}_{\text{perturbativa}} \otimes q_i \otimes q_j$
- SI PUÒ DEMONSTRARE RIGOROSAMENTE CON LO SVILUPPO DI WILSON

# RITRATTO DEL PROTONE VISTO DA UN BOSONE DI HIGGS



(PDG 2018)

## LE “DISTRIBUZIONI PARTONICHE”

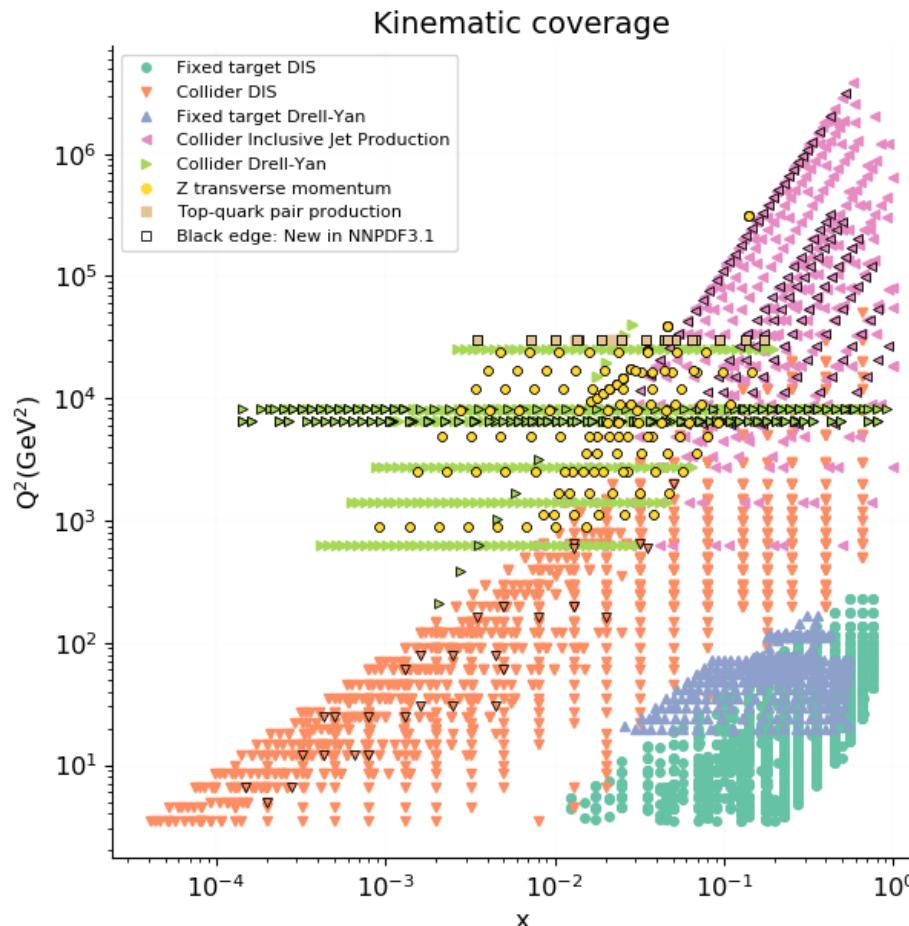
- DISTRIBUZIONI DI ENERGIA PER OGNI SPECIE DI QUARK, ANTIQUARK, E IL GLUONE
- $\int_0^1 f_{\text{up}}(x) dx = N_{\text{up}}, \dots$
- “VALENZA”  $N_{\text{up}} - N_{\text{antiup}} = 2; N_{\text{down}} - N_{\text{down}} = 2$
- DIPENDONO DALLA RISOLUZIONE  $\mu \sim \frac{1}{\lambda}$  (SCALA DI ENERGIA  $\sim \frac{1}{\text{distanza}}$ ); dipendenza dalla risoluzione calcolabile

# DOVE HA SBAGLIATO DYSON?

## COME SI FA IL RITRATTO?

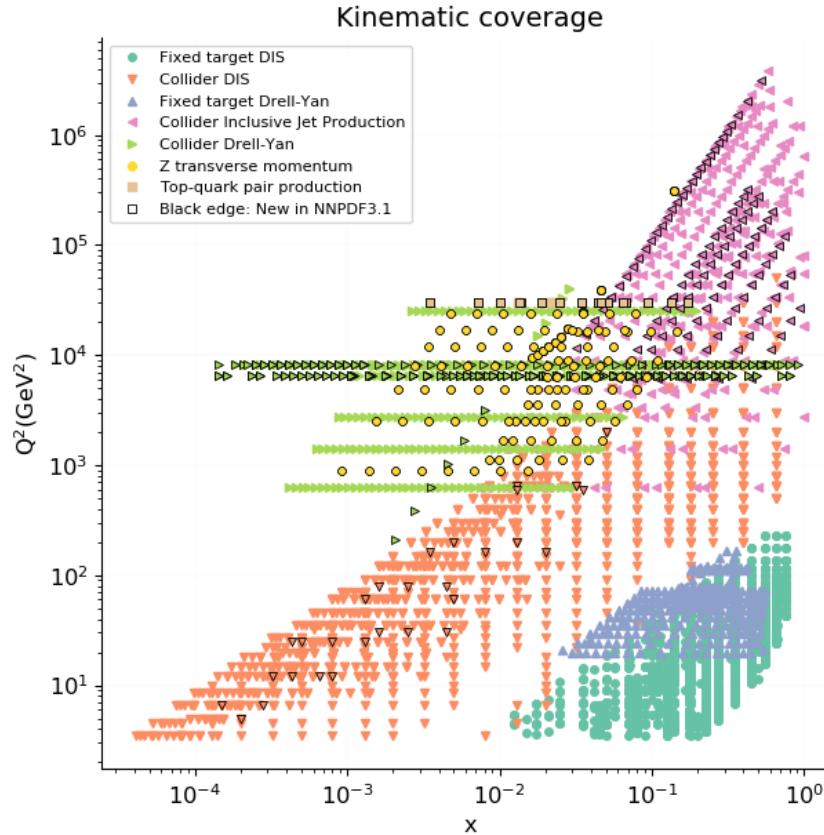
- CALCOLO DI UN PROCESSO FISICO  $\Rightarrow$  DIPENDE DALLA PDF
- CONFRONTO CON I DATI
- INVERSIONE

UN TIPICO INSIEME DI DATI (NNPDF3.1)



# DOVE HA SBAGLIATO DYSON?

## UN TIPICO INSIEME DI DATI(NNPDF3.1)

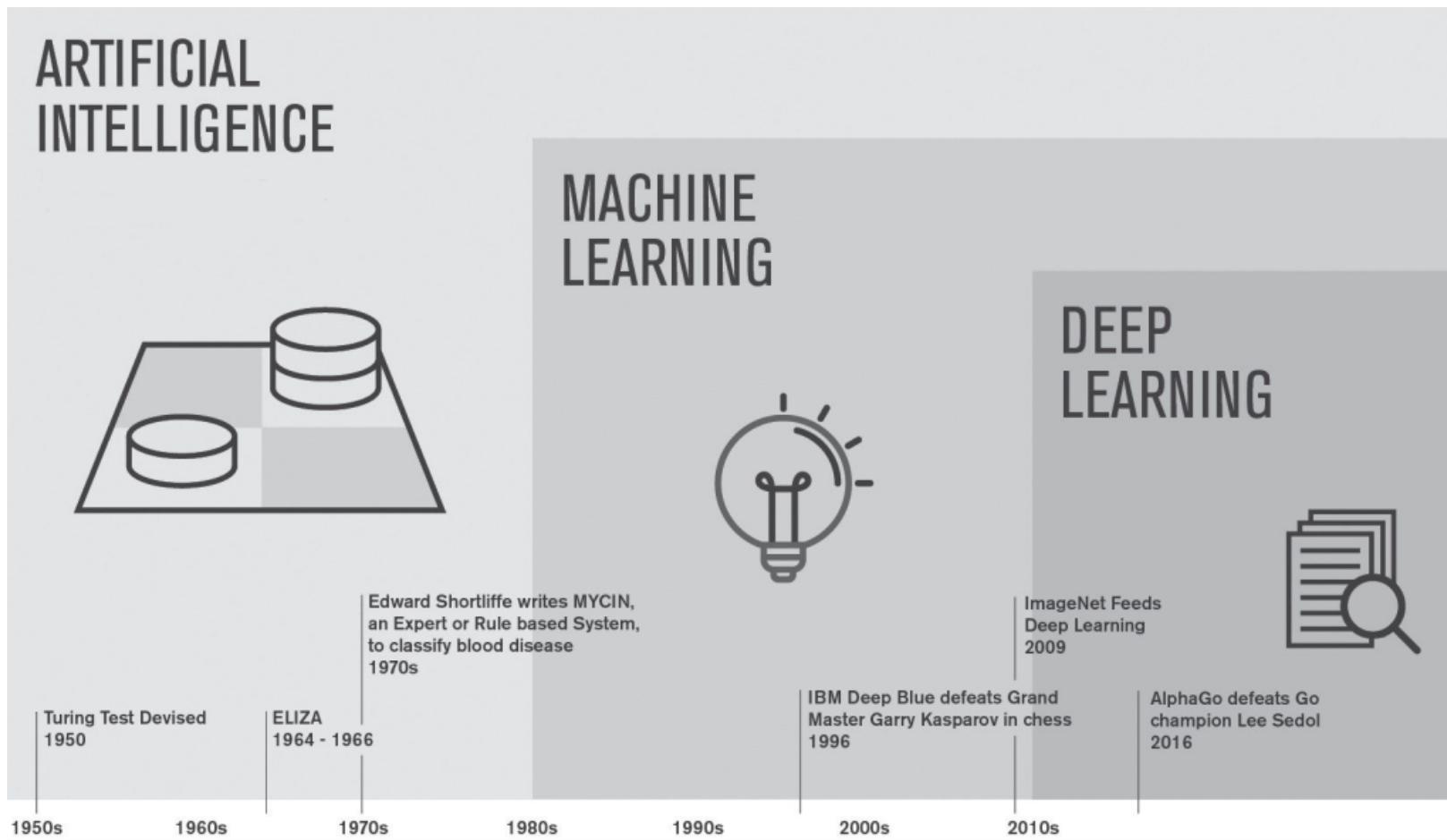


COME È POSSIBILE??

- OGNI PDF È UNA FUNZIONE una distribuzione, in realtà  
⇒ SI DEVE DETERMINARE UNA  
PROBABILITÀ IN UNO SPAZIO DI FUNZIONI
- SI DEVE DETERMINARE UN  
FUNZIONALE (CONTINUO) DA UN INSIEME DISCRETO DI DATI Sperimentali

# DOVE HA SBAGLIATO DYSON?

HA FATTO I CONTI SENZA L'INTELLIGENZA ARTIFICIALE!



# INTELLIGENZA ARTIFICIALE: PARADIGMI

## “KNOWLEDGE BASED” AI

- APPRENDIMENTO E IMPLEMENTAZIONE DI UN INSIEME DI REGOLE
- **BENE** PER GLI SCACCHI,  
**MALE** PER LA VITA VERA



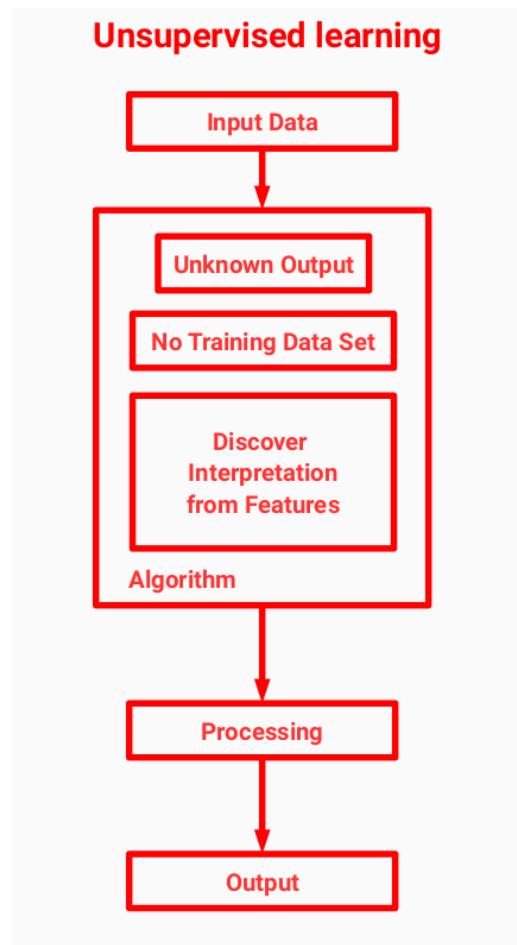
## MACHINE LEARNING

- RAPPRESENTAZIONE  
**“INTUITIVA”**
- IL SISTEMA **AI**  
ACQUISISCE LA  
PROPRIA CONOSCENZA

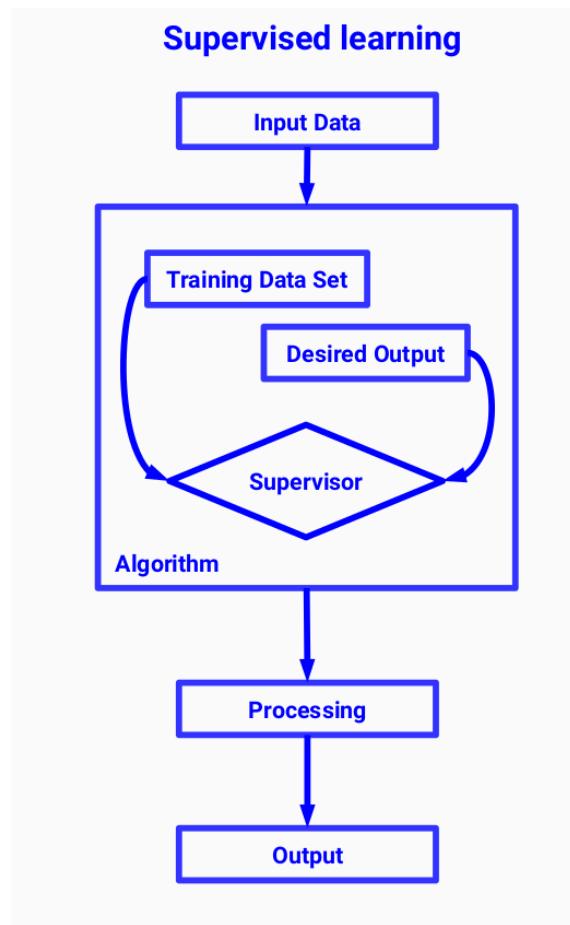
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3  
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4  
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5  
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6  
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7  
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8  
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9



# INTELLIGENZA ARTIFICIALE: ALGORITMI



ESTRAE E OTTIMIZZA  
LE CARATTERISTICHE DEI DATI



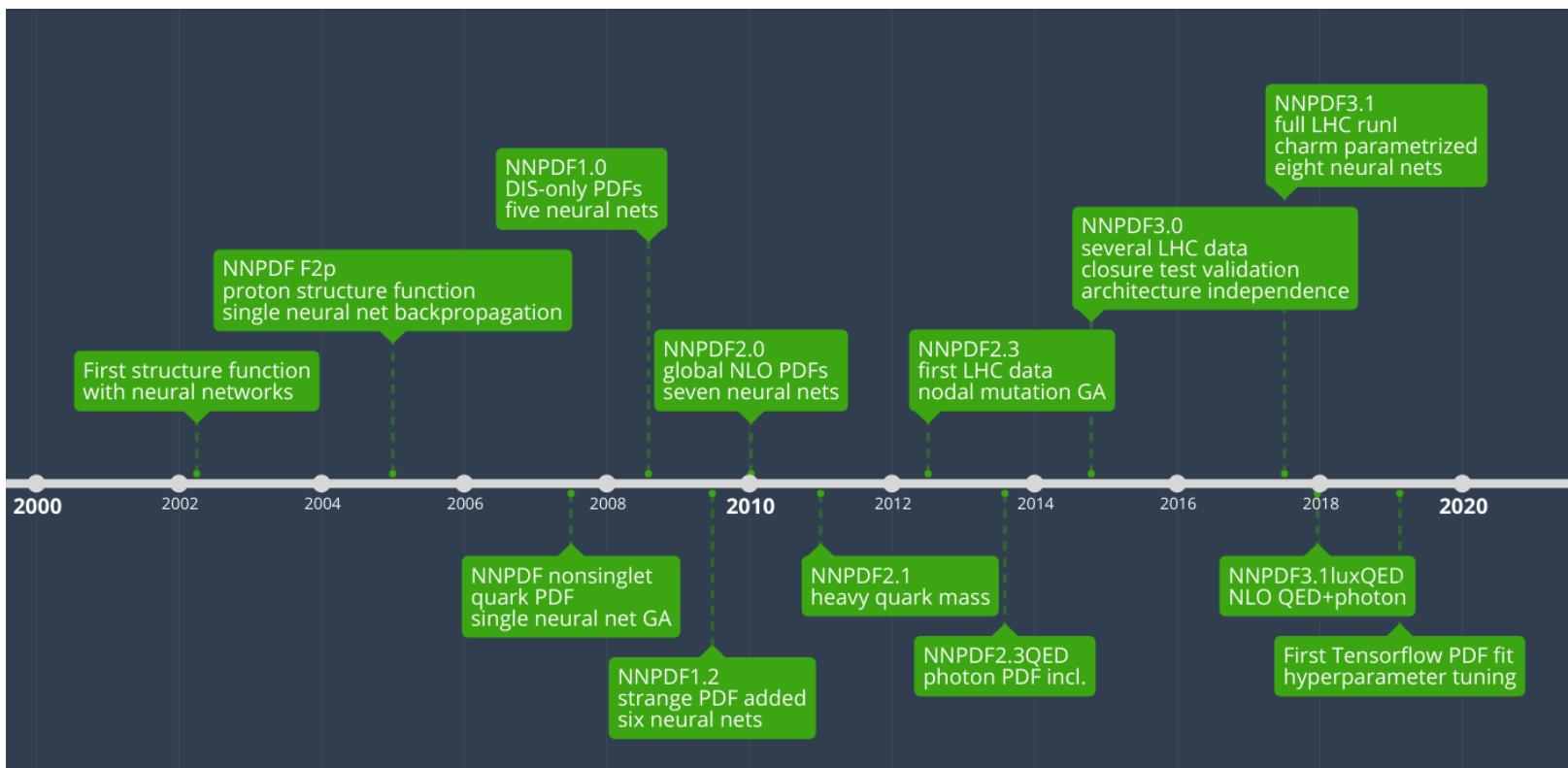
IMPARA DAI DATI OTTIMIZZANDO  
UNA PROPRIETÀ



IMPARA DAI DATI LA STRATEGIA  
DI APPRENDIMENTO

# LA STRUTTURA DEL PROTONE COME PROBLEMA DI AI: NNPDF

(s.f., Latorre: 2002)



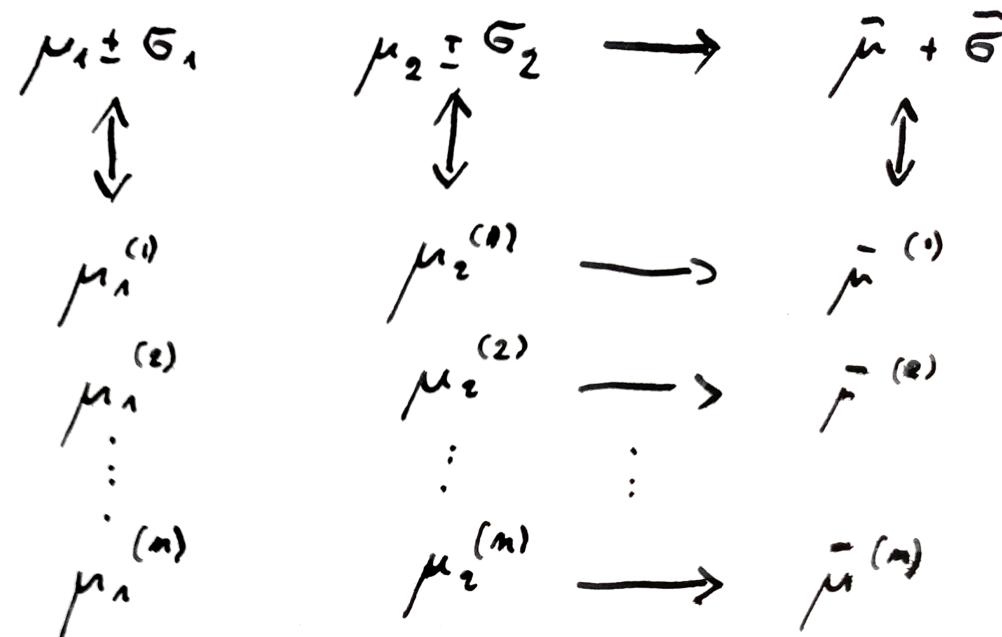
MONTECARLO + RETI NEURALI

## COMBINAZIONE DI DATI CON IL METODO MONTECARLO

DUE DATI:  $\mu_1 \pm \sigma_1$ ;  $\mu_2 \pm \sigma_2$

COMBINAZIONE DI MASSIMA VEROLOGIANZA:  $\bar{\mu} \pm \bar{\sigma}$ ;  $\bar{\mu} = \frac{\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$ ;  $\bar{\sigma}^2 = \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$

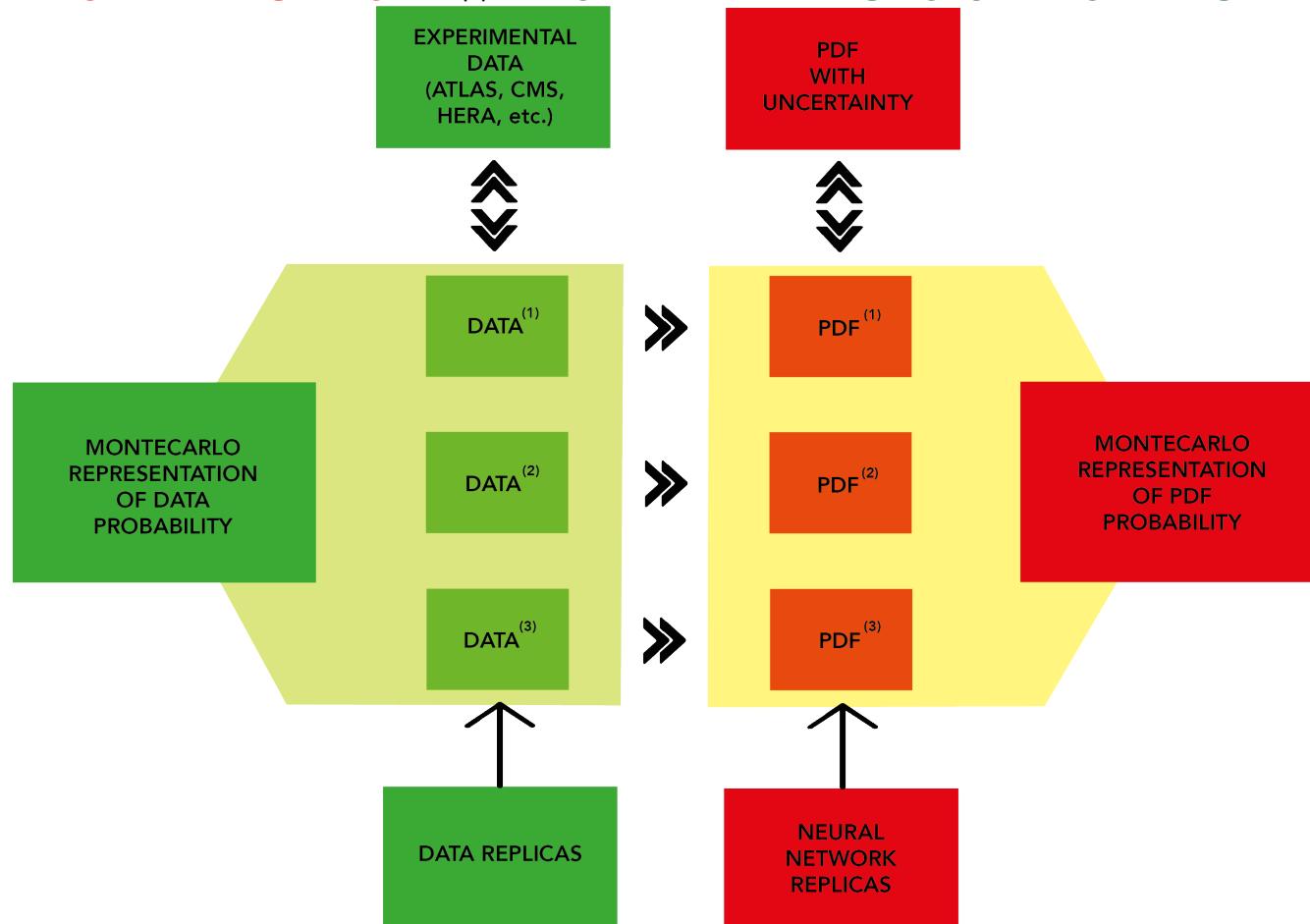
## RAPPRESENTAZIONE MONTECARLO



$\mu^{(i)} \Leftrightarrow$  REPLICHE  $\Leftrightarrow$  DISTRIBUZIONE DI PROBABILITÀ

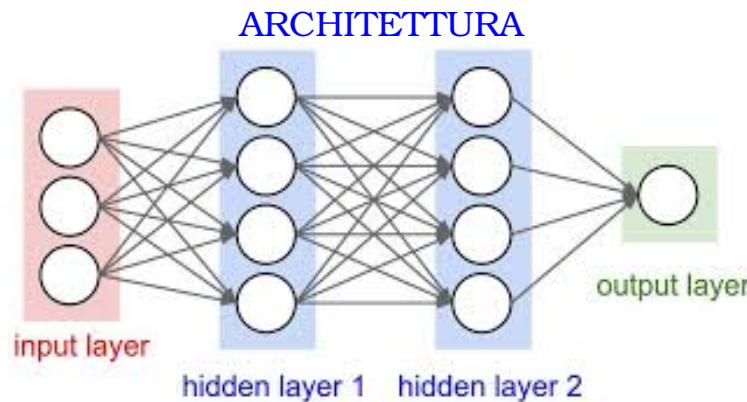
# LA STRUTTURA DEL PROTONE COME PROBLEMA DI AI IL MONTECARLO FUNZIONALE

REPLICHE DI FUNZIONI  $\Leftrightarrow$  PROBABILITÀ IN UNO SPAZIO DI FUNZIONI

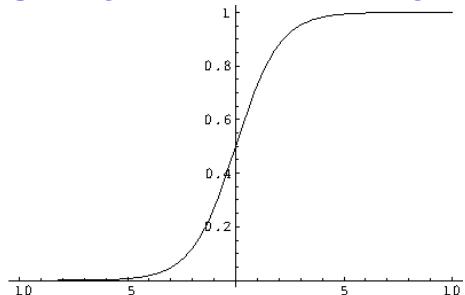


**PDF OUTPUT:**  $f_i^{(a)}(x, \mu)$ ;  $i = \text{up, antiup, down, antidown, strange, antistrange, charm, gluone}$ ;  
 $j = 1, 2, \dots N_{\text{rep}}$

# LA RETE NEURALE COME AGENTE AI UNIVERSALE



**FUNZIONE DI ATTIVAZIONE**



**PARAMETERI**

- **PESI**  $\omega_{ij}$
- **SOGLIE**  $\theta_i$

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

**IL PIÙ SEMPLICE ESEMPIO**

**1-2-1**

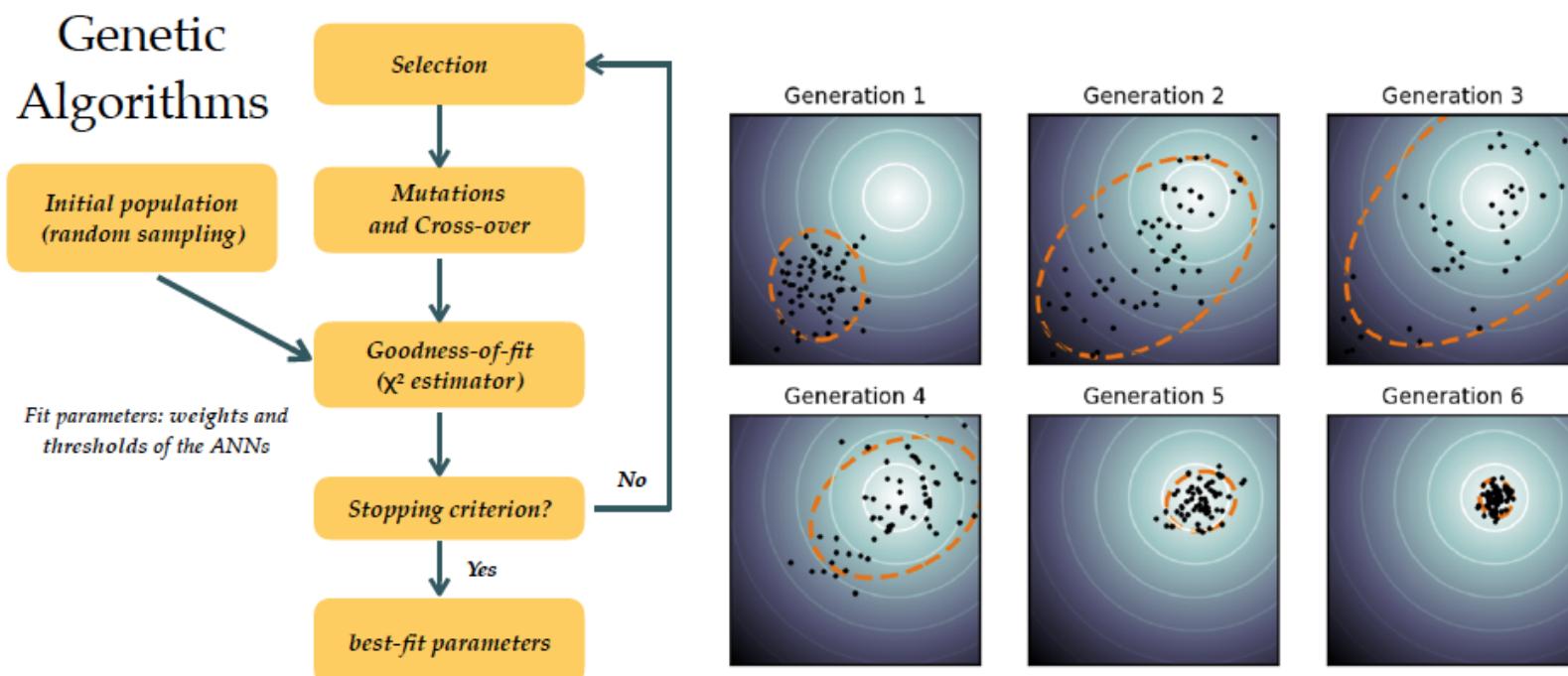
$$f(x) = \frac{1}{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1 + e^{\theta_1^{(2)} - x\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1 + e^{\theta_2^{(2)} - x\omega_{21}^{(1)}}}}$$

NNPDF: 2 – 5 – 3 – 1 NN PER OGNI PDF:  $37 \times 8 = 296$  PARAMETRI

# SUPERVISED LEARNING L'ALGORITMO GENETICO

## IDEA:

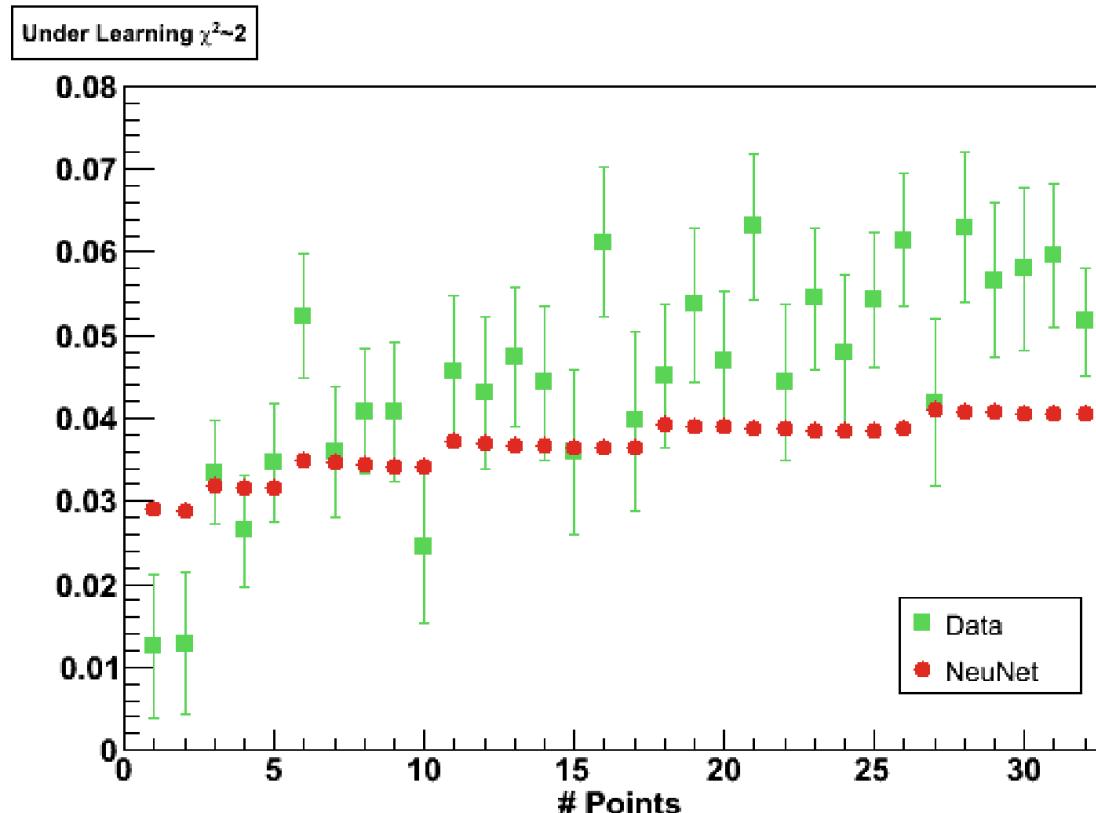
- MUTAZIONE CASUALE DEI PARAMETRI DELLA RETE
- SELEZIONE DEL PIÙ ADATTO



# RETI NEURALI: APPRENDIMENTO DI UNA LEGGE

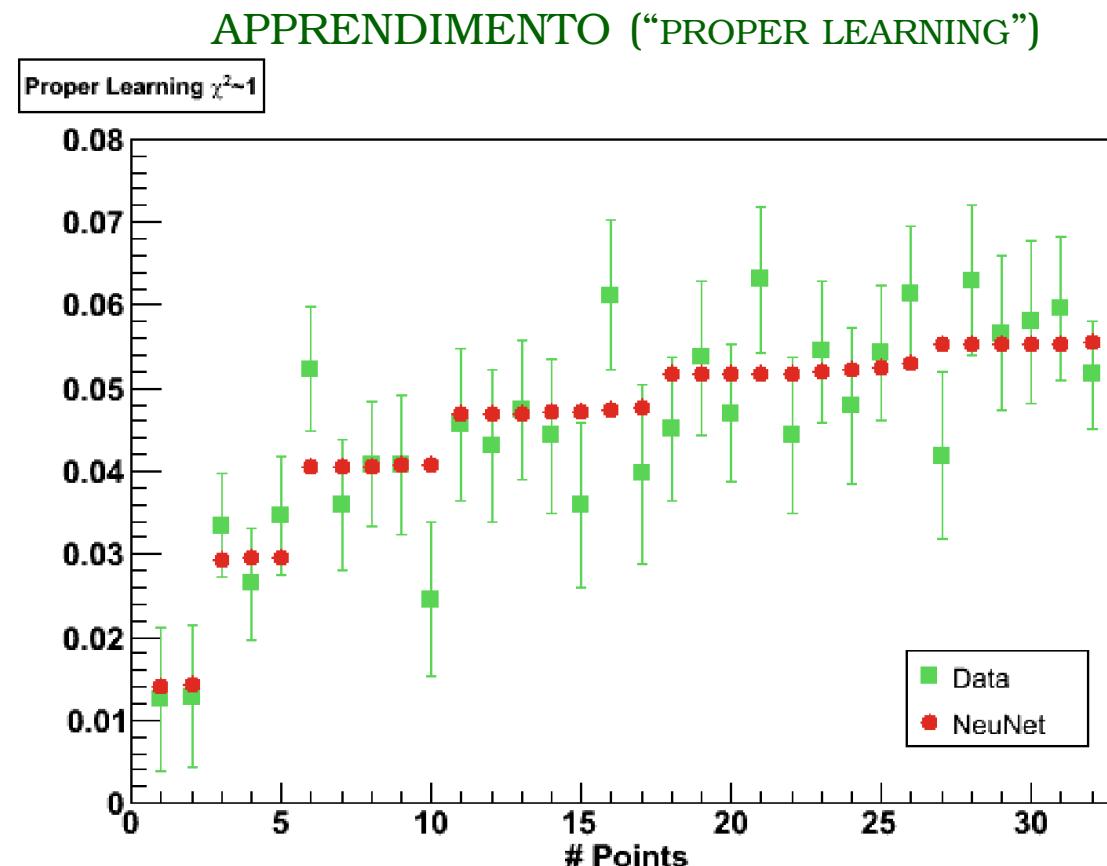
- LA COMPLESSITÀ AUMENTA DURANTE L'APPRENDIMENTO
- FINO A IMPARARE IL RUMORE
- QUANDO FERMARSI?

SOTTO-APPRENDIMENTO (“UNDERLEARNING”)



# RETI NEURALI: APPRENDIMENTO DI UNA LEGGE

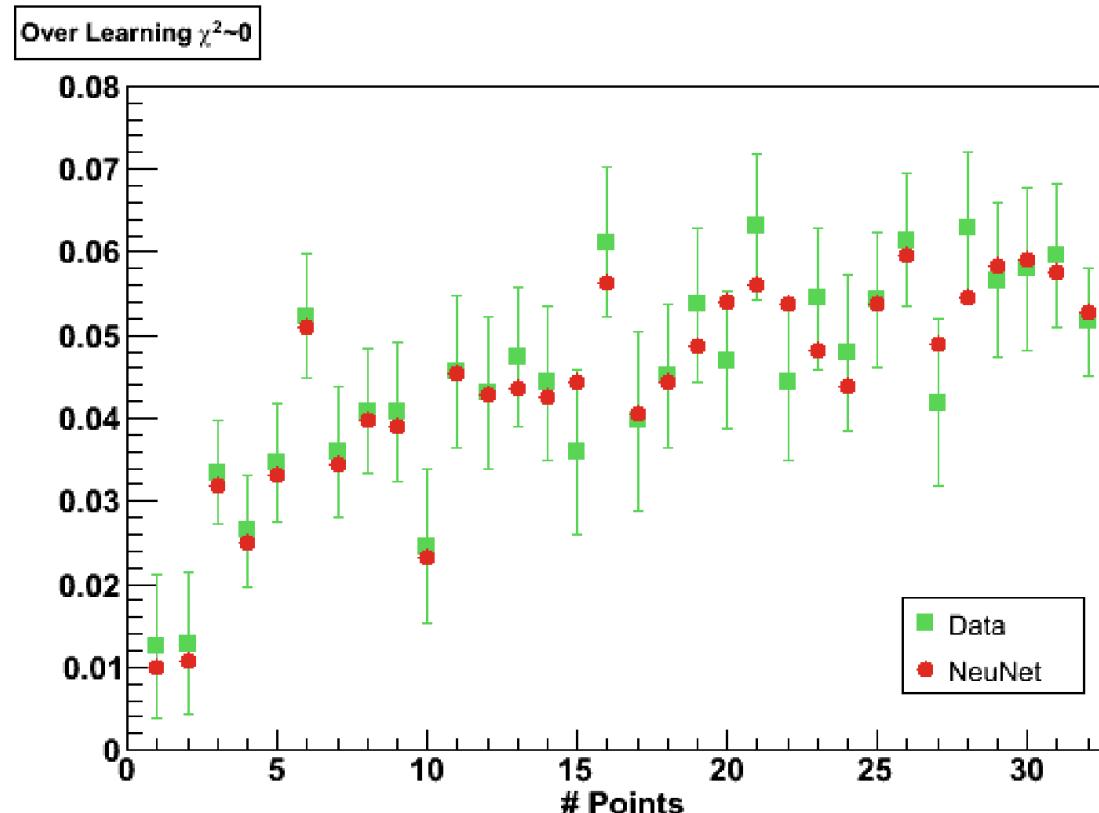
- LA COMPLESSITÀ AUMENTA DURANTE L'APPRENDIMENTO
- FINO A IMPARARE IL RUMORE
- QUANDO FERMARSI?



# RETI NEURALI: APPRENDIMENTO DI UNA LEGGE

- LA COMPLESSITÀ AUMENTA DURANTE L'APPRENDIMENTO
- FINO A IMPARARE IL RUMORE
- QUANDO FERMARSI?

SOVRA-APPRENDIMENTO (“OVERLEARNING”)

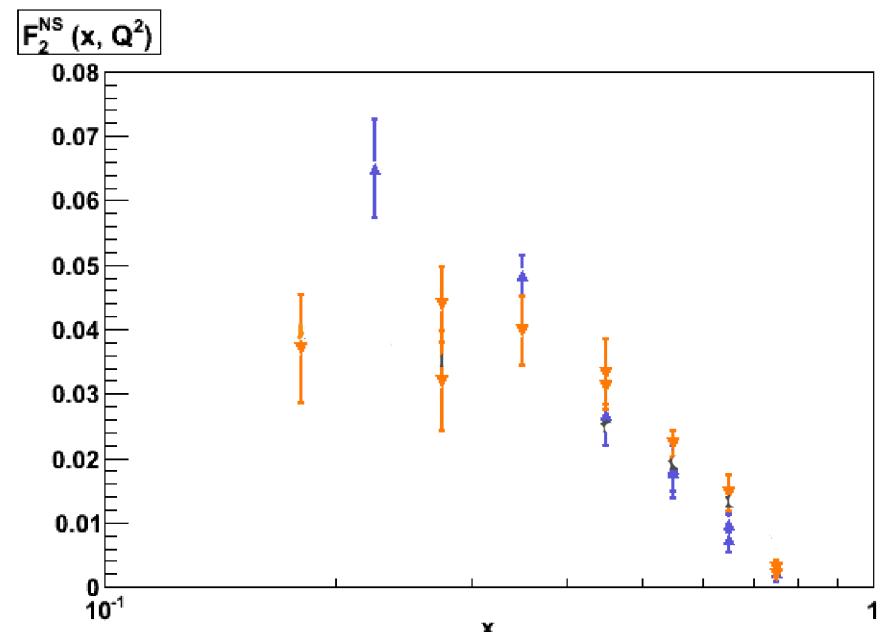


# CONVALIDA INCROCIATA!

ALGORITMO GENETICO:

AD OGNI GENERAZIONE, IL COSTO ( $\chi^2$ ) SCENDE O RESTA INVARIATO

- DIVIDI L'INPUT IN DUE: TRAINING CONVALIDA
- MINIMIZZA IL COSTO DELL'INSIEME DI TRAINING
- AD OGNI GENERAZIONE CALCOLA IL COSTO DELL'INSIEME DI CONVALIDA (NON USATO PER LA MINIMIZZAZIONE)
- QUANDO IL COSTO DI CONVALIDA NON SCENDE PIÙ L'APPRENDIMENTO È COMPLETO



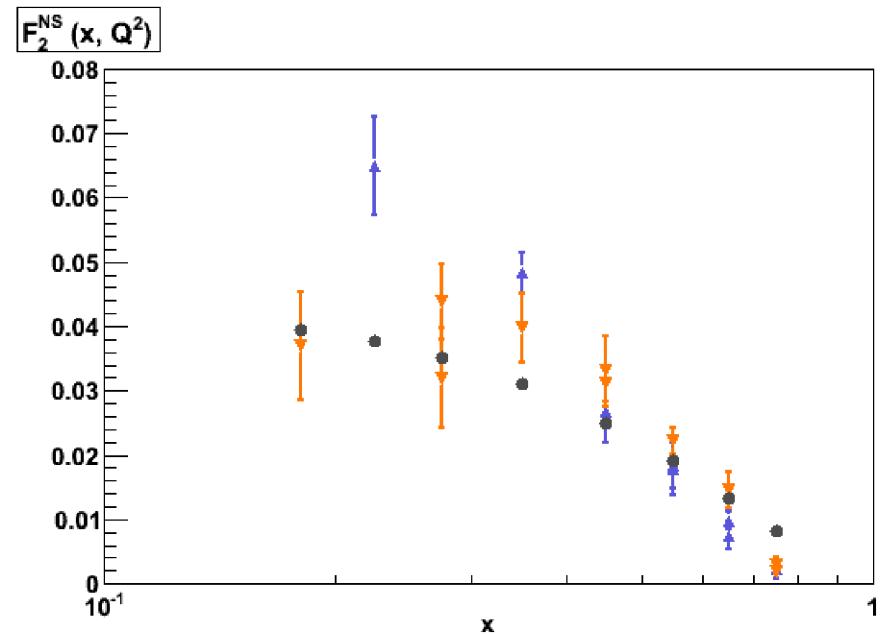
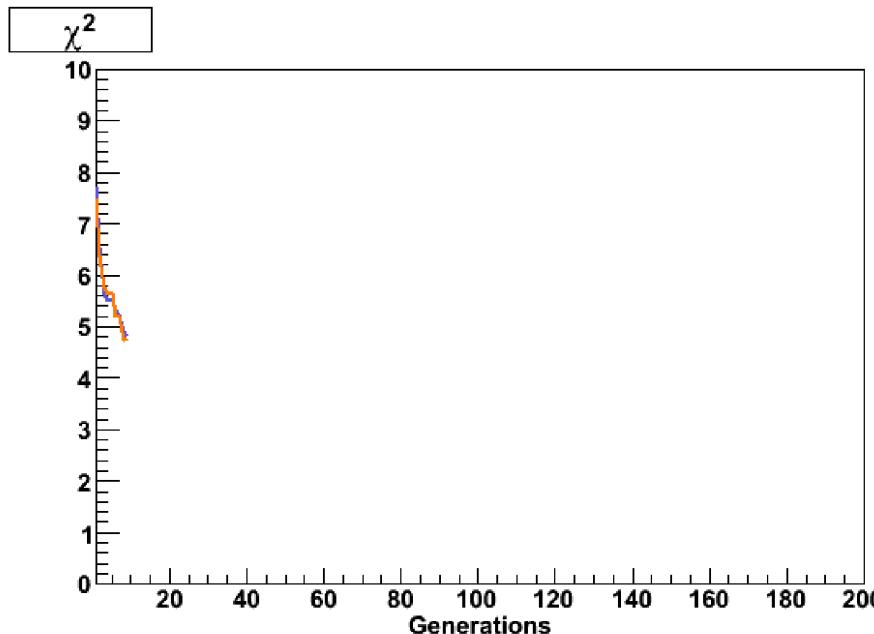
# CONVALIDA INCROCIATA!

ALGORITMO GENETICO:

AD OGNI GENERAZIONE, IL COSTO ( $\chi^2$ ) SCENDE O RESTA INVARIATO

- DIVIDI L'INPUT IN DUE: TRAINING CONVALIDA
- MINIMIZZA IL COSTO DELL'INSIEME DI TRAINING
- AD OGNI GENERAZIONE CALCOLA IL COSTO DELL'INSIEME DI CONVALIDA (NON USATO PER LA MINIMIZZAZIONE)
- QUANDO IL COSTO DI CONVALIDA NON SCENDE PIÙ L'APPRENDIMENTO È COMPLETO

GO!



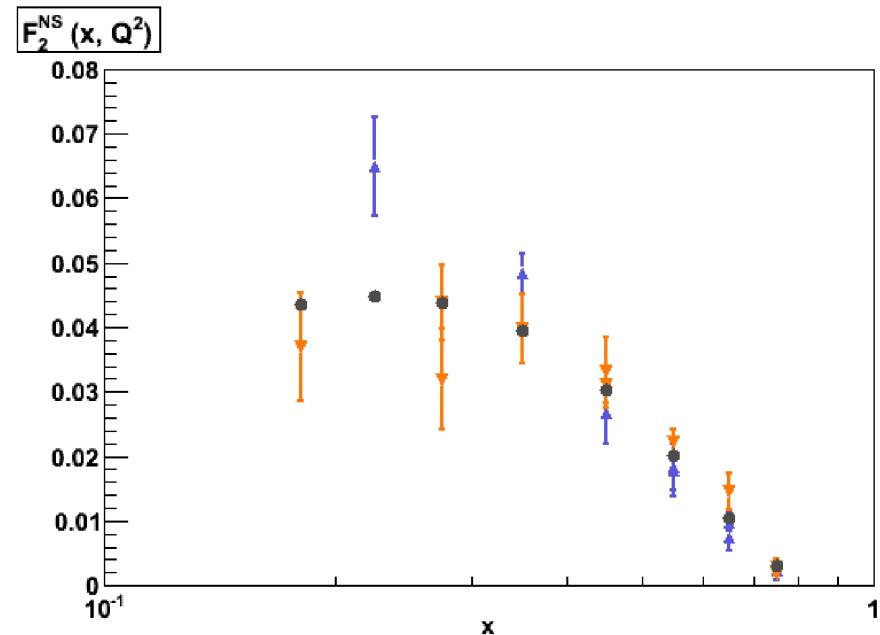
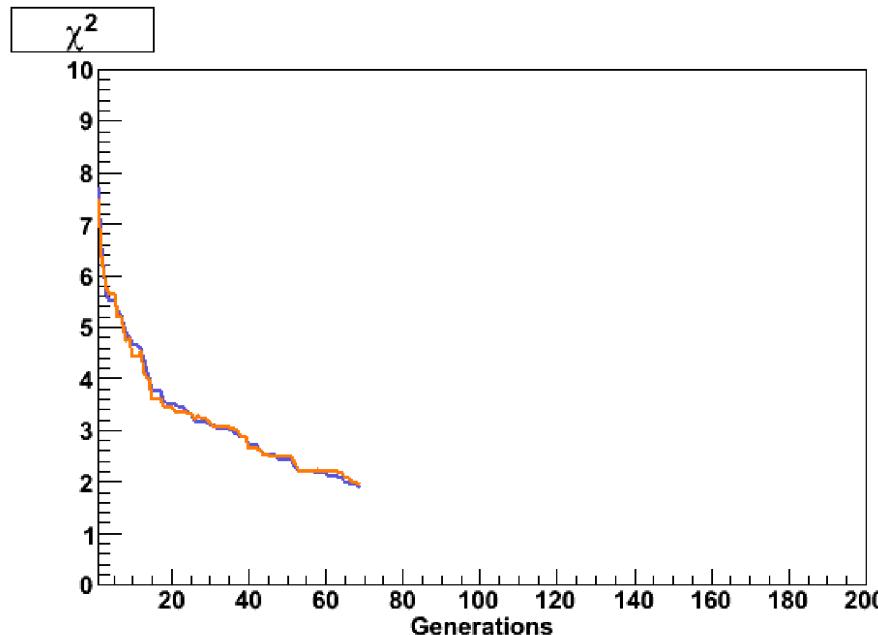
# Q CONVALIDA INCROCIATA!

ALGORITMO GENETICO:

AD OGNI GENERAZIONE, IL COSTO ( $\chi^2$ ) SCENDE O RESTA INVARIATO

- DIVIDI L'INPUT IN DUE: TRAINING CONVALIDA
- MINIMIZZA IL COSTO DELL'INSIEME DI TRAINING
- AD OGNI GENERAZIONE CALCOLA IL COSTO DELL'INSIEME DI CONVALIDA (NON USATO PER LA MINIMIZZAZIONE)
- QUANDO IL COSTO DI CONVALIDA NON SCENDE PIÙ L'APPRENDIMENTO È COMPLETO

STOP!



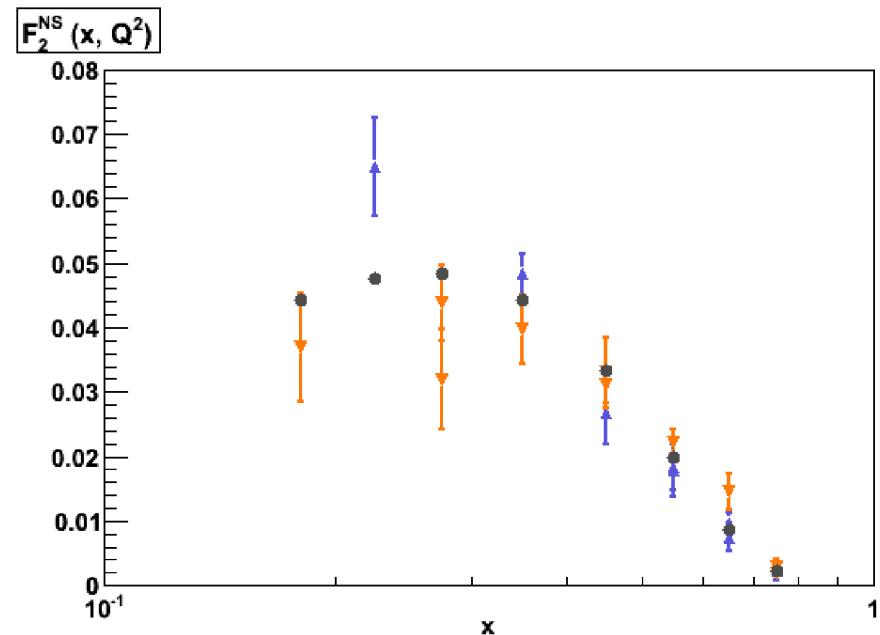
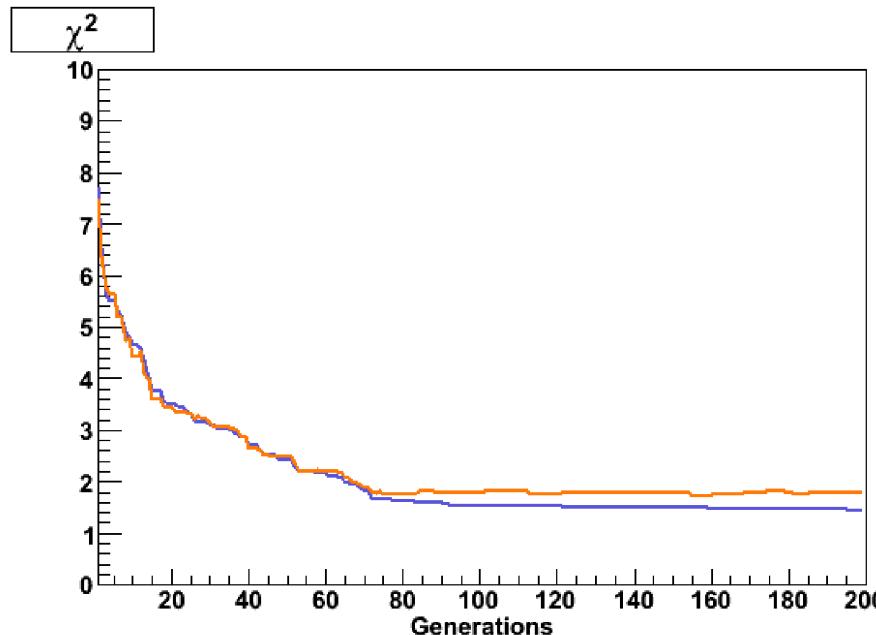
# CONVALIDA INCROCIATA!

ALGORITMO GENETICO:

AD OGNI GENERAZIONE, IL COSTO ( $\chi^2$ ) SCENDE O RESTA INVARIATO

- DIVIDI L'INPUT IN DUE: TRAINING CONVALIDA
- MINIMIZZA IL COSTO DELL'INSIEME DI TRAINING
- AD OGNI GENERAZIONE CALCOLA IL COSTO DELL'INSIEME DI CONVALIDA (NON USATO PER LA MINIMIZZAZIONE)
- QUANDO IL COSTO DI CONVALIDA NON SCENDE PIÙ L'APPRENDIMENTO È COMPLETO

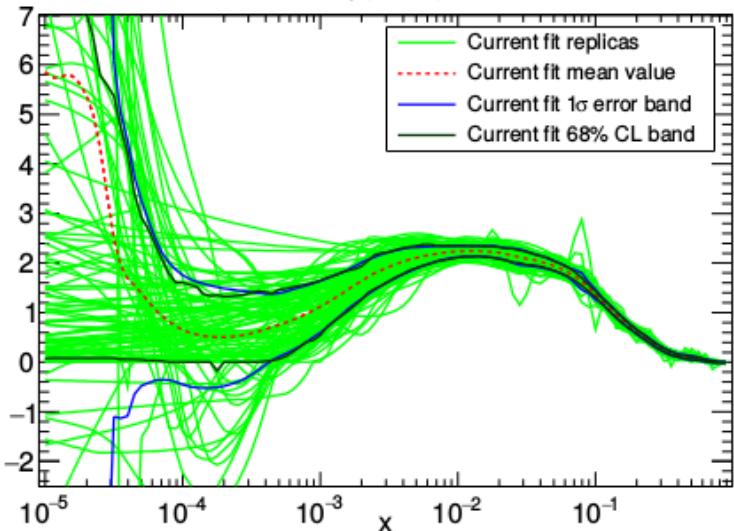
TROPPO TARDI!



# IL RISULTATO IL RITRATTO DEL PROTONE (100 REPLICHE)

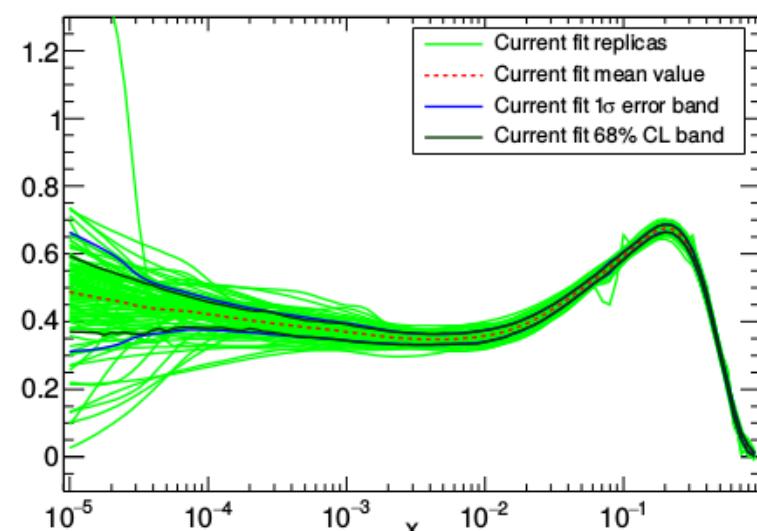
GLUON

$xg(x, Q^2)$



UP

$xu(x, Q^2)$

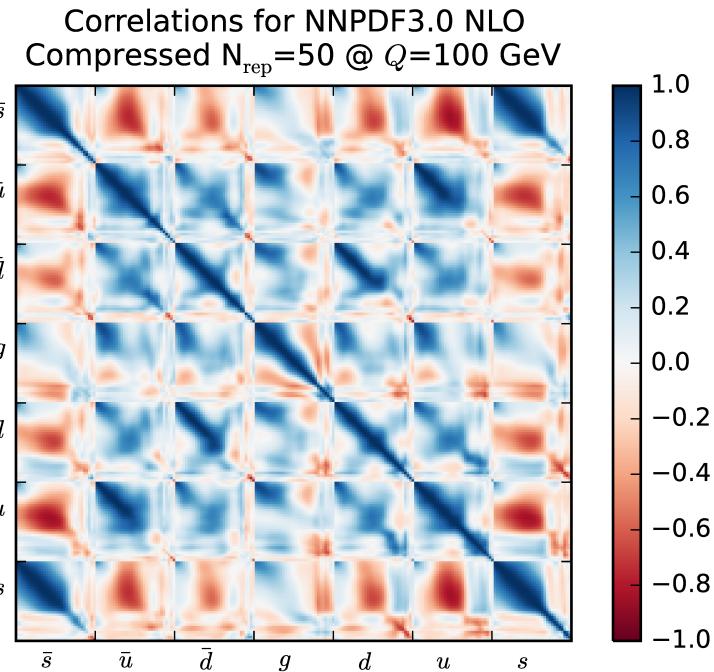
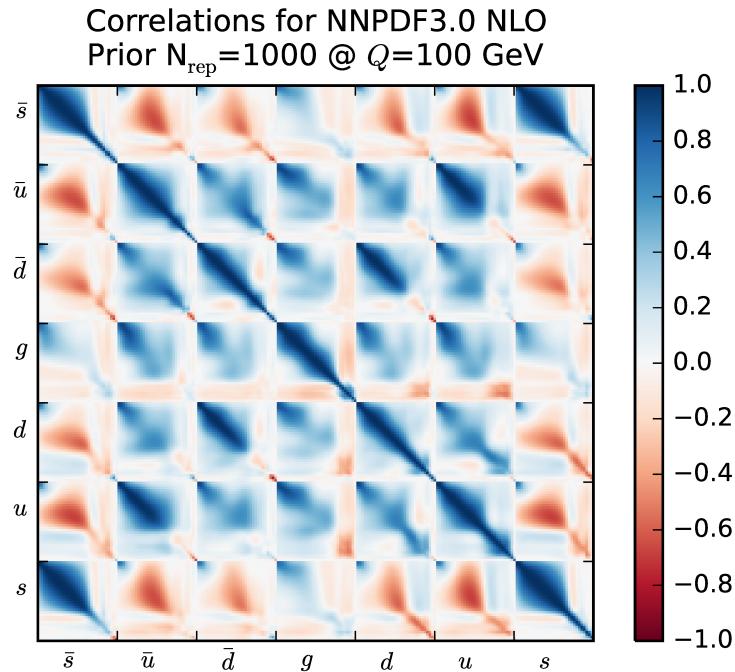


- DISTRIBUZIONE DI PROBABILITÀ  $\leftrightarrow$  REPLICHE
- RISULTATO PIÙ PROBABILE  $\leftrightarrow$  MEDIA
- INCERTEZZA  $\leftrightarrow$  DEVIAZIONE STANDARD
- OPPURE 68% C.L.
- MA BASTANO 100 REPLICHE?

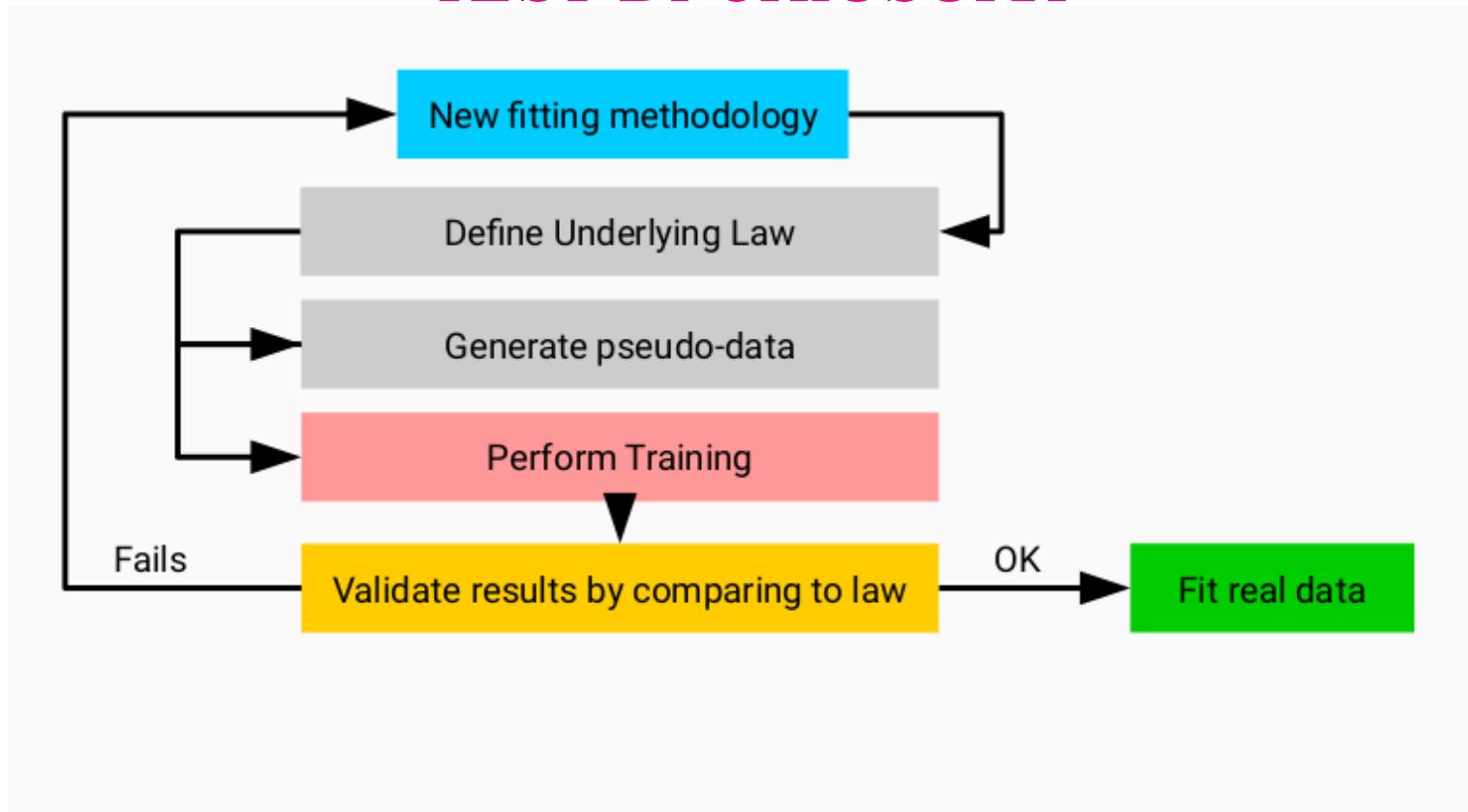
# UNSUPERVISED LEARNING OTTIMIZZAZIONE

- COME MASSIMIZZARE L'ACCURATEZZA?
- INSIEME (PRIOR) DI MOLTE REPLICHE
- SELEZIONE GENETICA  $\Rightarrow$  OTTIMIZZAZIONE INDICATORI STATISTICI (DIVERGENZA KULLBACK-LEIBLER)
- 50 REPLICHE OTTIMIZZATE  $\Leftrightarrow$  1000 REPLICHE DI PARTENZA

MATRICE DI CORRELAZIONE



## COME FACCIAMO A SAPERE CHE IL RISULTATO È GIUSTO? TEST DI CHIUSURA

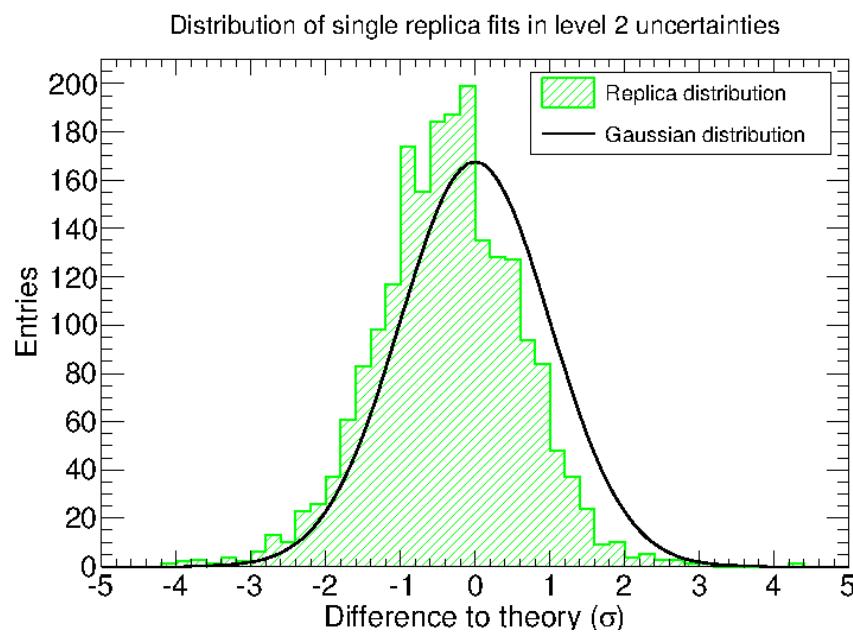
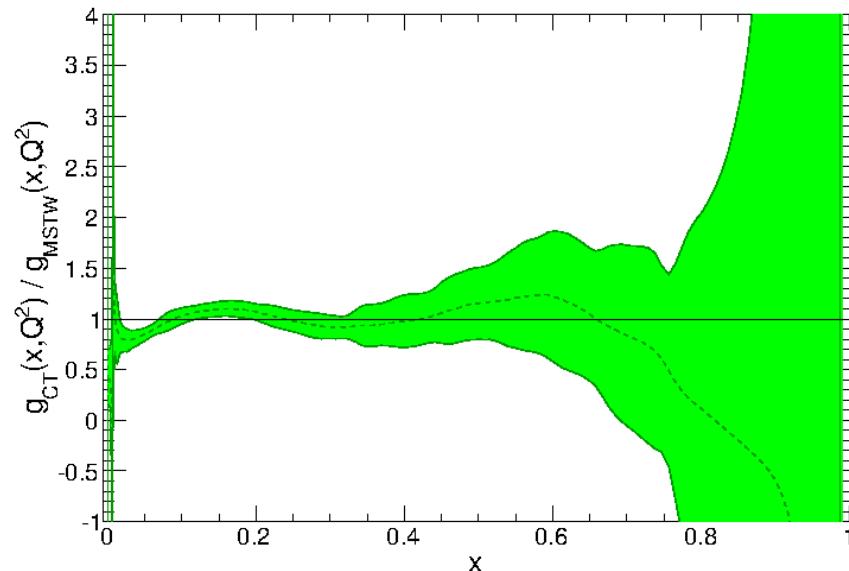


# PRIMO TEST DI CHIUSURA

## RISULTATI

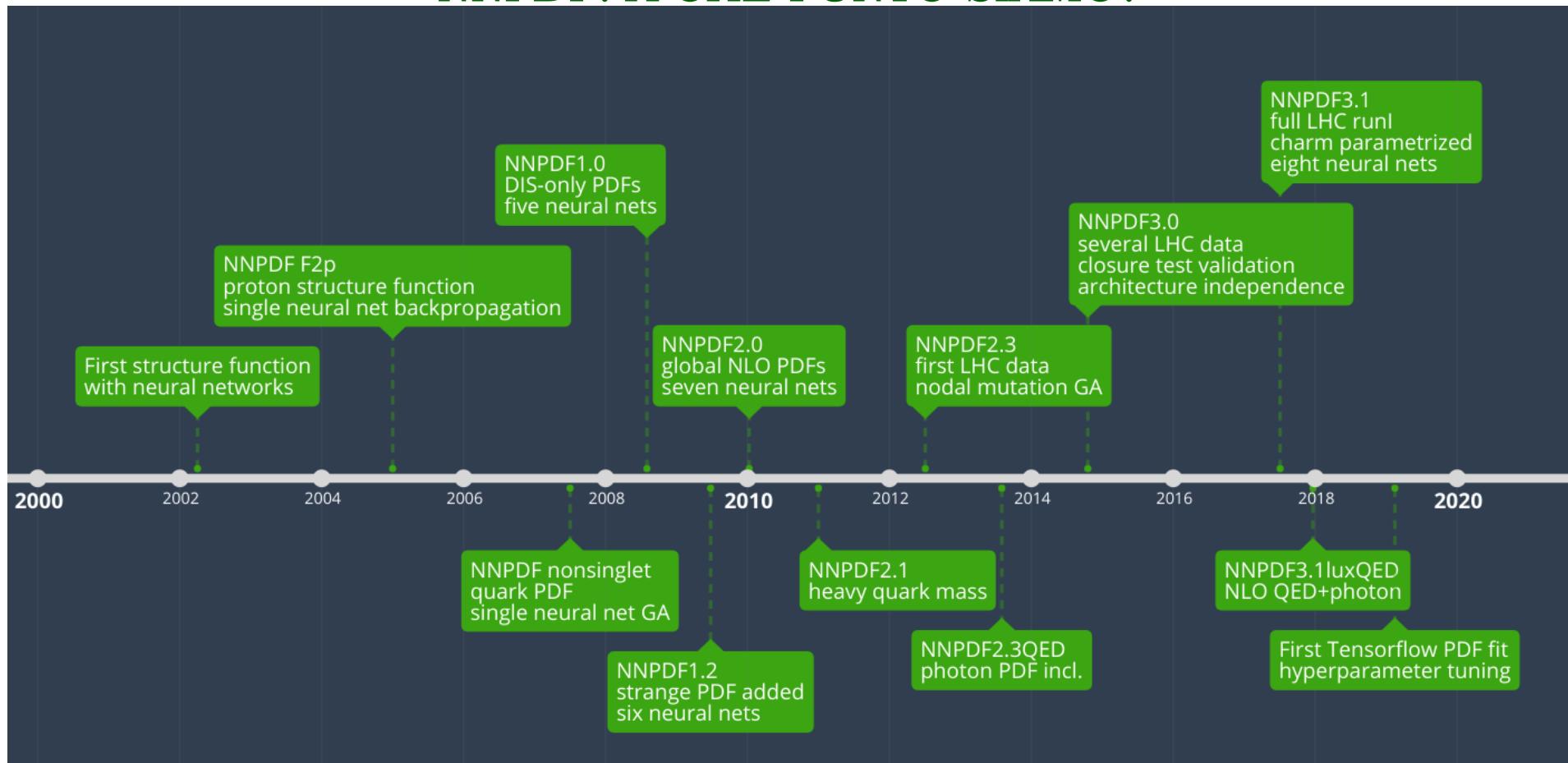
### DISTRIBUZIONE NORMALE DELLE DEVIAZIONI

IL GLUONE: RISULTATO / "VERITÀ"  
Ratio of Closure Test g to MSTW2008



REINFORCEMENT LEARNING?

## NNPDF: A CHE PUNTO SIAMO?



# VERSO IL “DEEP LEARNING”

## NUOVI STRUMENTI

- KERAS, TENSORFLOW, . . .
- LIBRERIE OPEN-SOURCE PER IL DEEP LEARNING,  
ELEVATA EFFICIENZA
- PROGRAMMAZIONE AD ALTO LIVELLO IN PYTHON



SET NNPDF: 15 ORE  $\Rightarrow$  10 MINUTI

### OBIETTIVI:

- OTTIMIZZAZIONE DEI PARAMETRI
- “REINFORCEMENT LEARNING” AUTOMATICO DELLA METODOLGIA OTTIMALE

# VERSO IL “DEEP LEARNING”

## OTTIMIZZAZIONE DEI PARAMETRI



- VELOCITÀ ⇒ POSSIBILE PROVARE STRATEGIE ALTERNATIVE
- SCAN DELLO SPAZIO DI PARAMETRI ⇒ AUTOMATIZZAZIONE

# “DEEP LEARNING” CHE COSA ABBIAMO IMPARATO?

IMPARARE UNA LEGGE SCONOSCIUTA  $\Leftrightarrow$  PROBABILITÀ  
**IL PROGETTO NUOTO**



IMPARARE LA **QUALITÀ** DI UN NUOTATORE DAL **PROFILO DI VELOCITÀ**  
 $\Rightarrow$  STRATEGIA DI ALLENAMENTO **OTTIMIZZATA**

UN LAVORO DI MOLTI!  
LA COLLABORAZIONE NNPDF  
**NNPDF**



IL TEAM N<sup>3</sup>PDF

