MACHINE LEARNING TECHNIQUES IN HIGH ENERGY COLLIDERS

Roger J. Hernández Pinto
FCFM-UAS
roger@uas.edu.mx

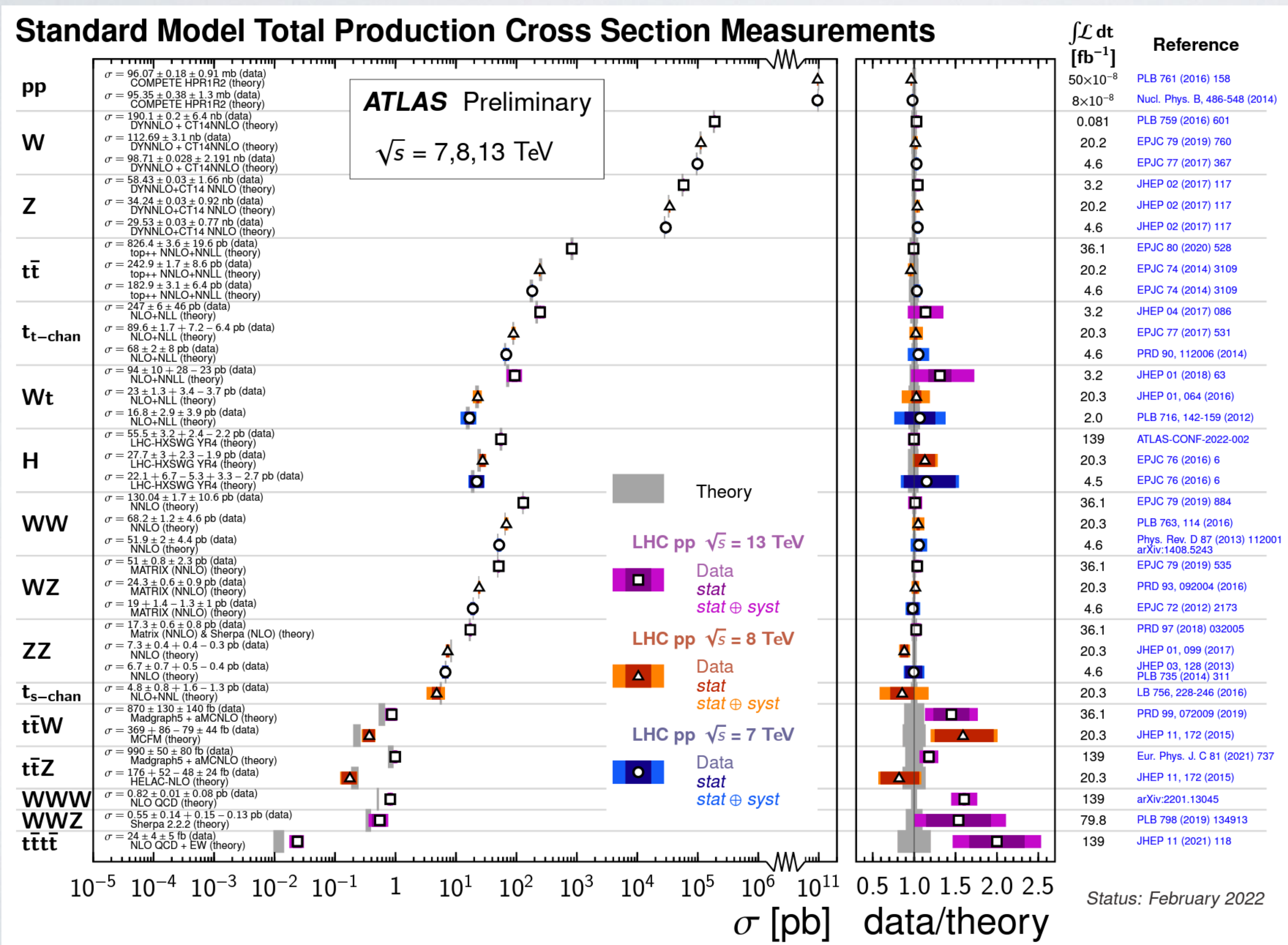
In collaboration with:
D. F. Rentería-Estrada, G.F.R. Sborlini and P. Zurita

XXXVII Annual Meeting of the Mexican Division of Particles and Fields
Physics Department, CINVESTAV-Zacatenco, June 12th, 2023

INTRODUCTION

- The description of the fundamental interactions rely on unitary and local quantum field theories.
- Multiloop scattering amplitudes describe the quantum fluctuations at high-energy scattering processes are the main bottleneck.
- Accurate theoretical predictions in High Energy Physics require to deal with multiloop and multileg scattering amplitudes.
- Even if we can compute all loops, numerical integrations are needed to compare with experimental results.

PRECISE MEASUREMENTS



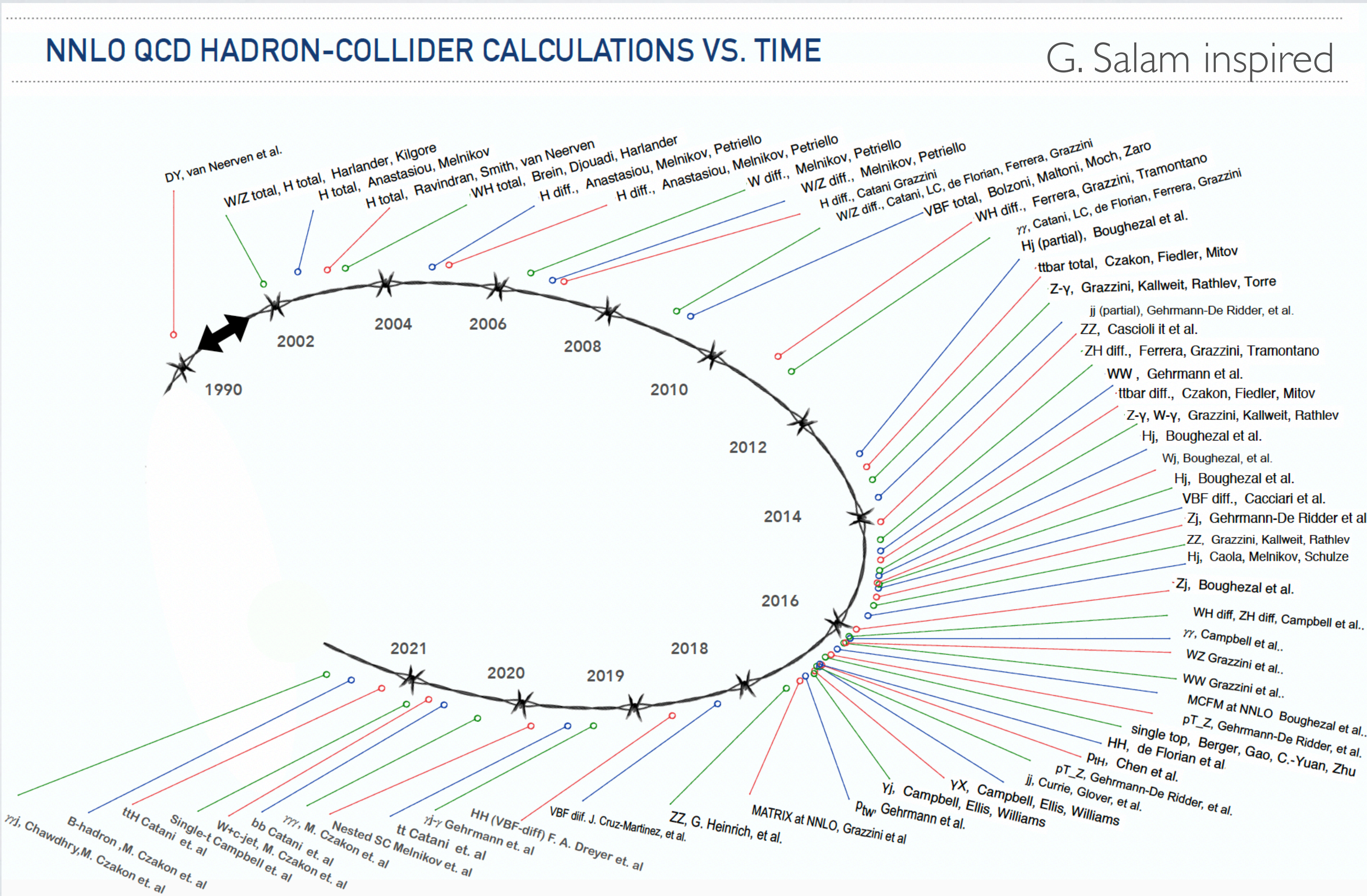
Experiments are guiding the precision frontier

Most observables are known to the highest possible accuracy

What is the highest theoretical precision ?

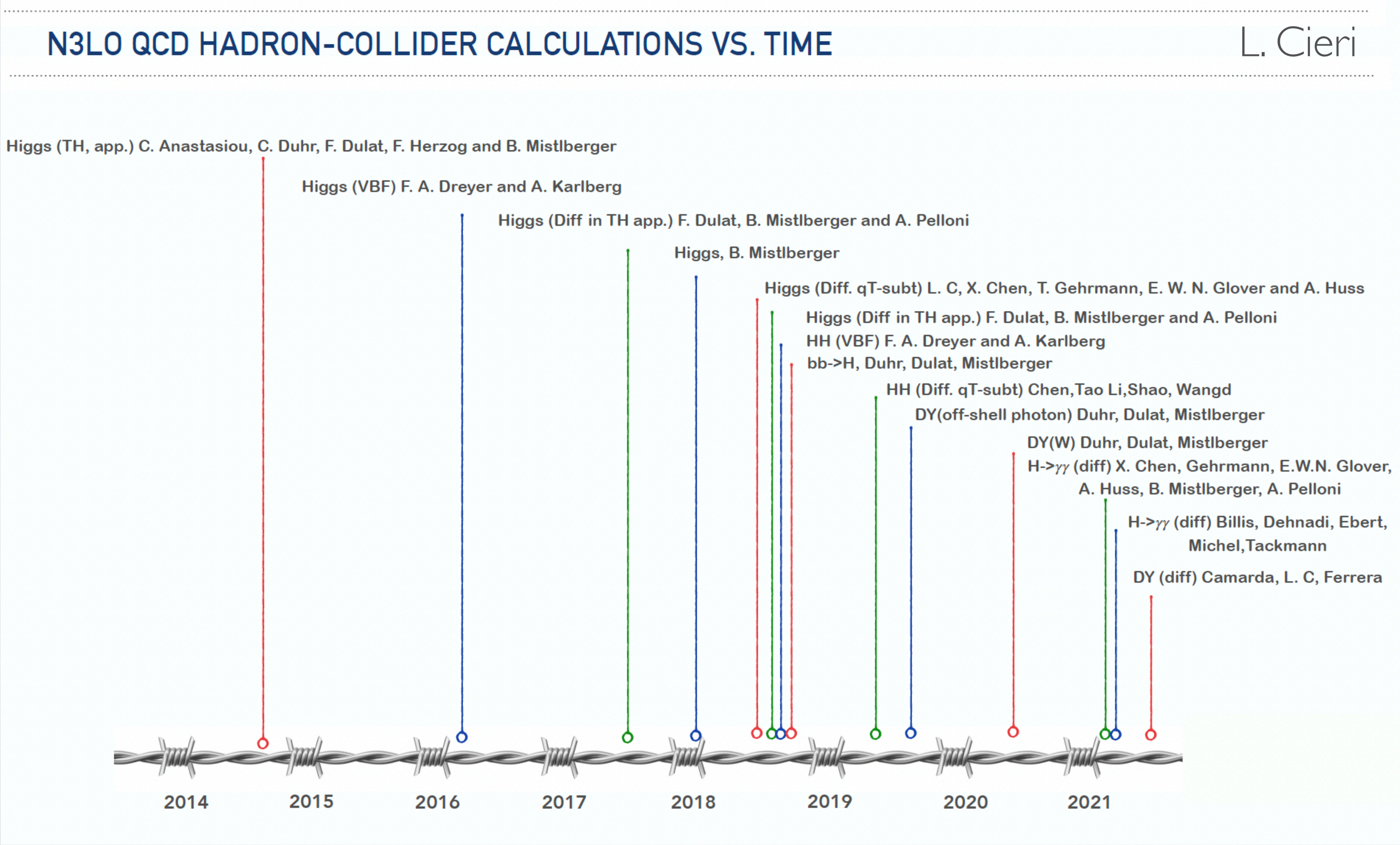
Where is the theory precision frontier ?

STATUS AT NNLO



N3LO

How difficult could it be adding and extra "N" to the game?

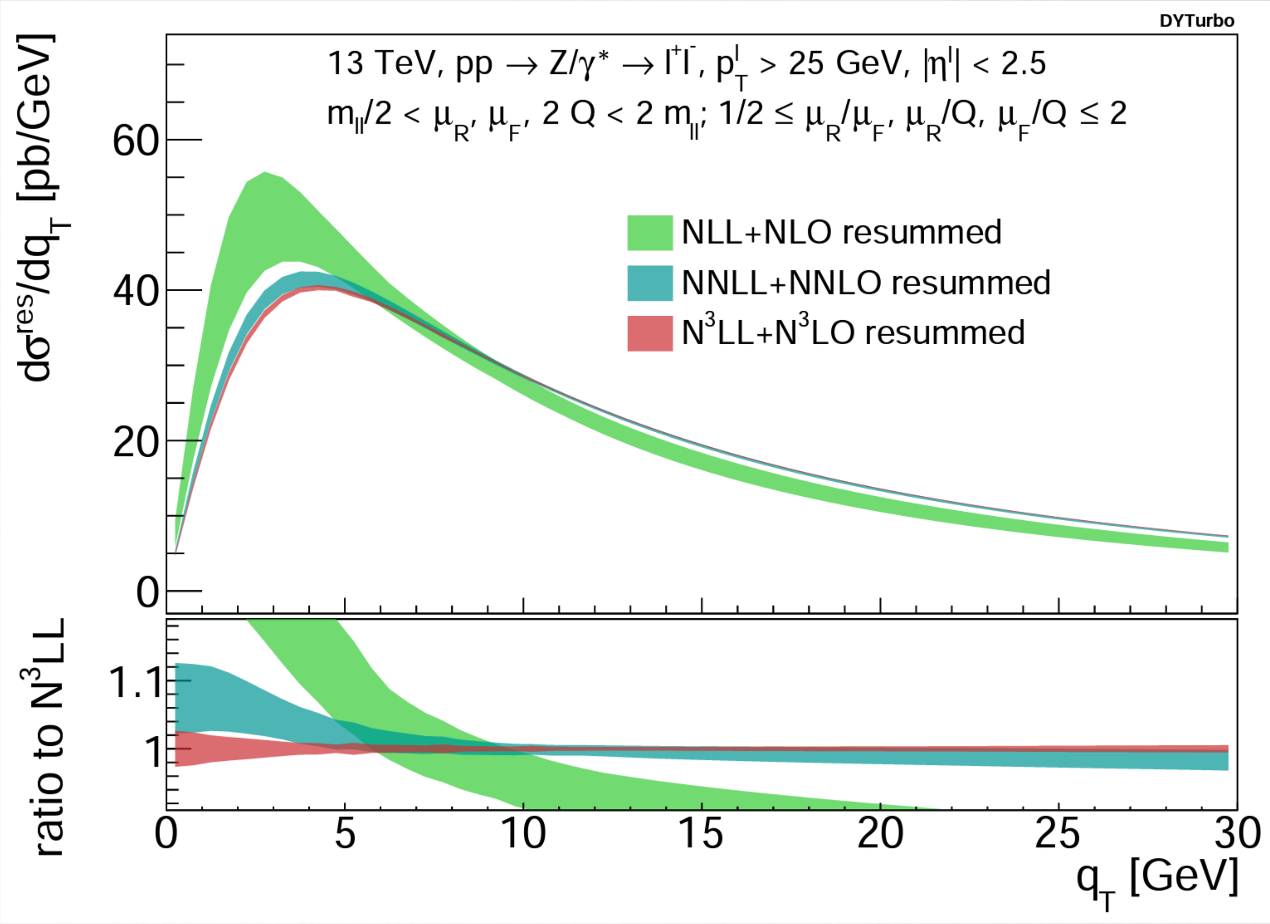


THEORETICAL ISSUES

- Integrands are usually lengthy.
- The number of Feynman diagrams increase enormously when high accuracy is required.
- Feynman amplitudes have divergences everywhere. We need to deal with all of them !!
- Experimental results are always real numbers so, how we extract real numbers from divergent quantities numerically ?
- Furthermore, exact comparisons must be done at Monte Carlo level and a vast of numerical instabilities shall appear in the calculations.

THEORY MEETS EXPERIMENT

- However, experimental data require the theoretical input pushing the precision frontier to a highest order in QCD.
- Physics searches have to be done at the highest posible accuracy.
- New methods for higher order calculations are extremely important.
- Experimental results needs Monte Carlo simulations in order to compare with nature.



Phys. Rev. D 104 (2021) 11, L111503

HANDLING INFINITIES

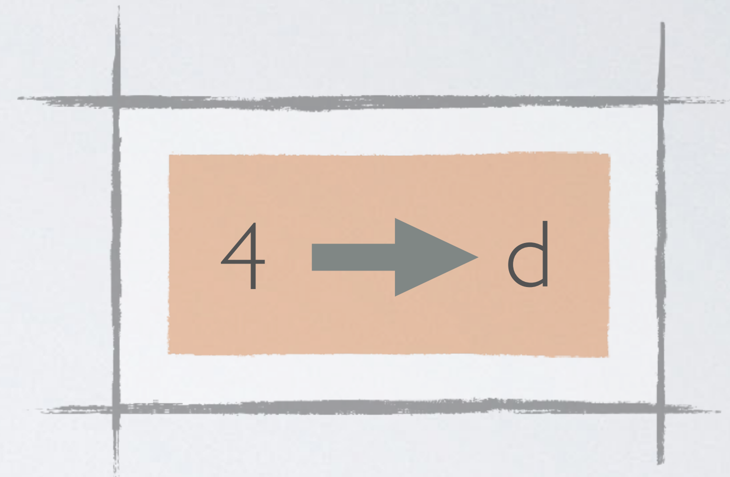
- The standard definition for observables rely on QFTs.
- The basic algorithm are the Feynman rules which provides the theoretical predictions of nature through complicated expressions.
- First, build the scattering amplitude $|\mathscr{M}\rangle$, and then

$$\sigma \sim \int d\mathrm{PS} \, |\mathscr{M}|^2$$

- However, the validity of QFT is extrapolated to *infinity energy,* when loops are computed, and also to *zero energy* when parallel particles mimic the behavior to a single particle emitted.

DIMENSIONAL METHODS

- Dimensional schemes
    - 't-Hooft and Veltman (HV) '72
    - Conventional Dim. Reg. (CDR) '73
    - Dimensional Red. (DRED) '79
    - Four-dimensional helicity (FDH) '92
- Reformulation of dimensional schemes
    - Six-Dim. Formalism (SDF) '09
    - Four-Dim. Formalism (FDF) '14



Desired properties:
- Mathematical consistency
- Unitarity and causality

Can I add:
- Symmetries ?
- Computational efficiency ?

DIM. SCHEMES - ISSUES

- Virtual corrections are, in some sense, doable up to certain level, but real contributions are a hard task to perform.

- UV renormalization is tedious.

- Pseudoscalars are complicated to study since the γ_5 is not well defined in d-dimensions.

-

By construction:	Algebraic definition:
$\gamma_5^{\mathrm{BM}} \equiv \frac{i}{4!}(\varepsilon^{\mu\nu\rho\sigma}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma)_{[4]}$ $\equiv \frac{i}{4!}\varepsilon^{\mu\nu\rho\sigma}_{[4]}(\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma)_{[d]}$	$\{\gamma_5^{\mathrm{AC}}, \gamma^\mu_{[d]}\} \equiv 0$

$d \neq 4$ →

$$\{\gamma_5^{\mathrm{BM}}, \gamma^\mu_{[d]}\} = 2\gamma^\mu_{[d-4]}\gamma_5^{\mathrm{BM}}$$

↓

$$\{\gamma_5^{\mathrm{BM}}, \gamma^\mu_{[d]}\} \neq \{\gamma_5^{\mathrm{AC}}, \gamma^\mu_{[d]}\}$$

Different results !

NON-DIMENSIONAL METHODS

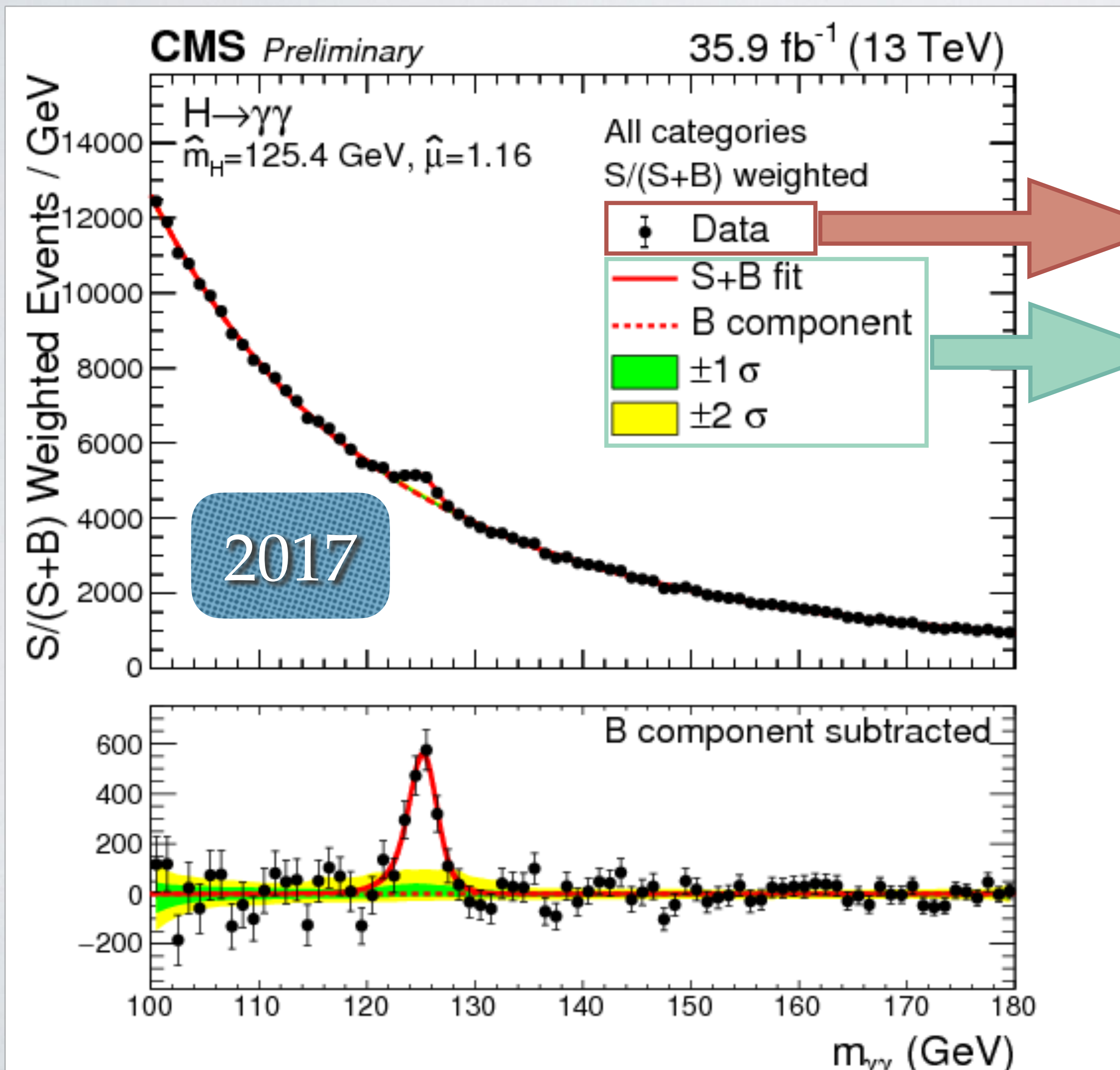
- Non-dimensional schemes
  - Implicit reg. (IREG) '98
  - Four-Dim. Reg/Ren. (FDR) '12
  - Four-Dim. Unsubtraction (FDU) '16
- Towards cross sections in 4D (LTD)
  - Removing threshold singularities, '21
  - Causality of multi-loop amplitudes, '21
  - LTD and Quantum computing, '21

FOUR

Desired properties:
- Mathematical consistency
- Unitarity and causality

Can I add:
- Symmetries ?
- Computational efficiency ?

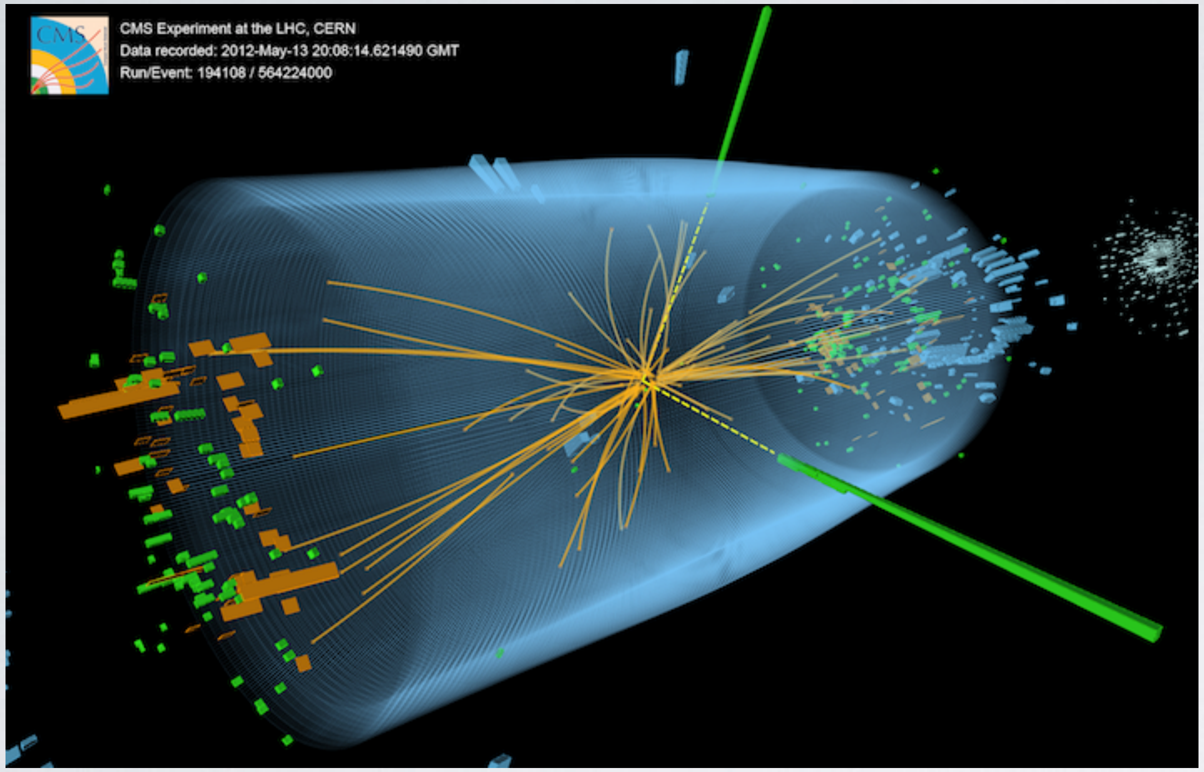
PRECISION PHENOMENOLOGY

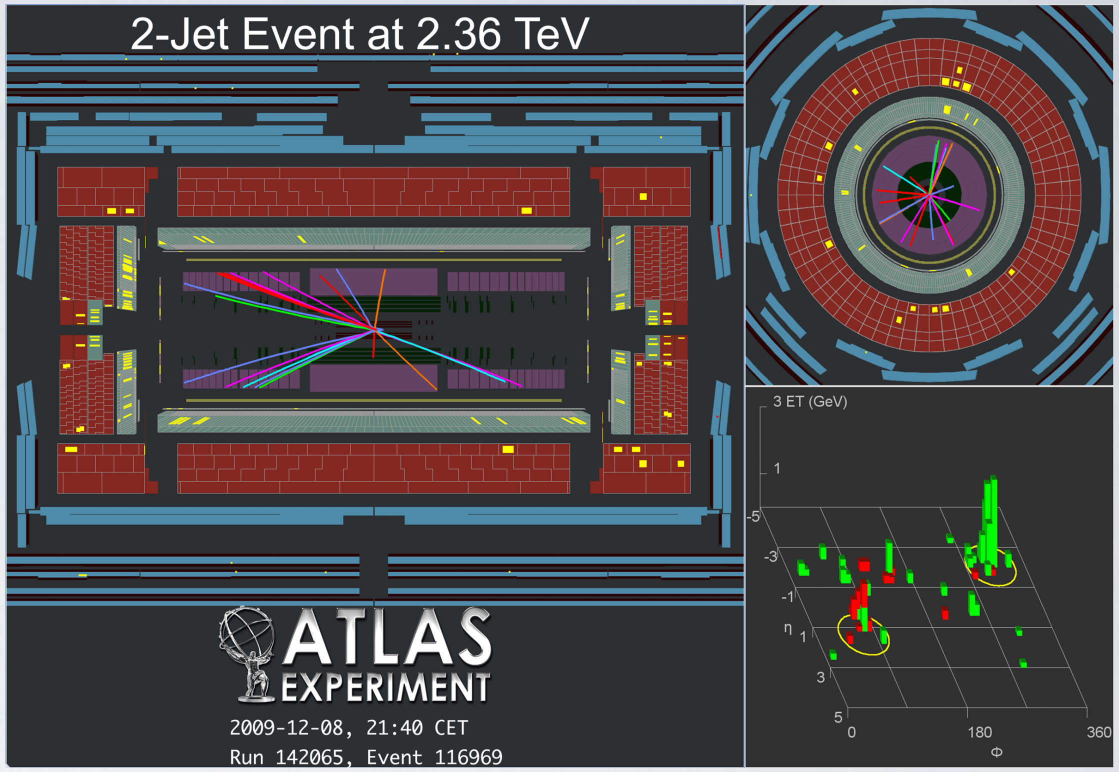


Experiment

Phenomenology

Strong knowledge in :
-Quantum Field Theory
-Mathematical Methods
-Programming, …

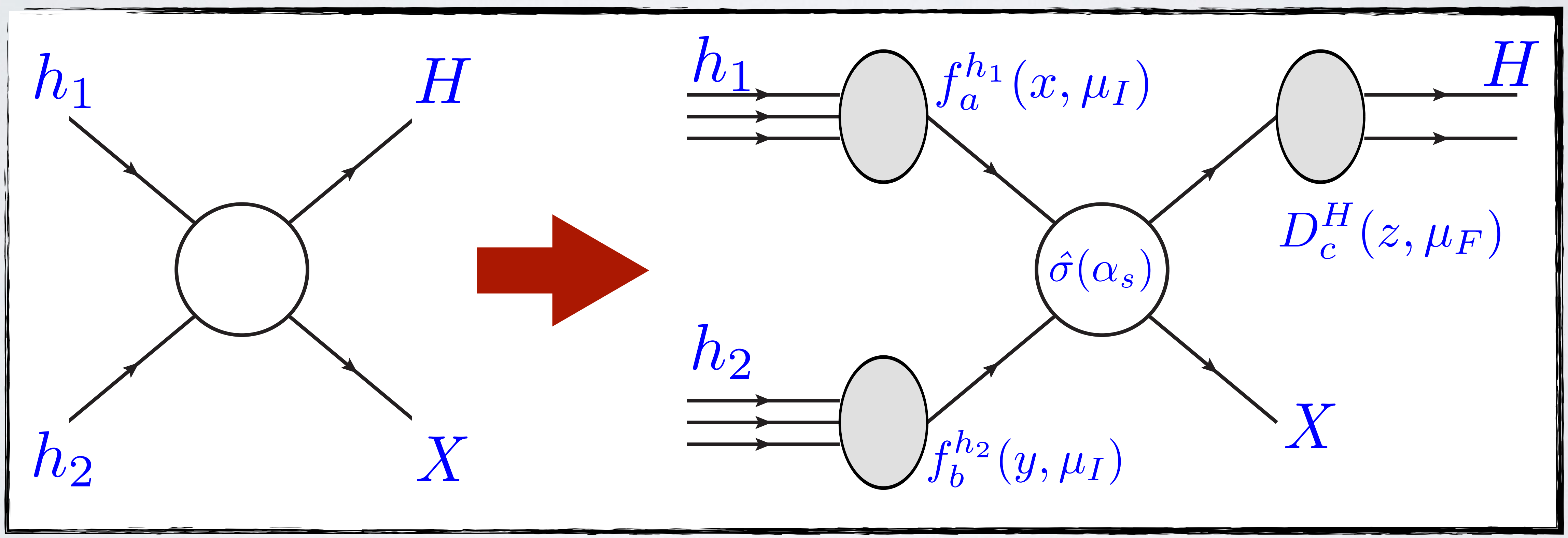




Phenomenology in Hadron Colliders

THEORETICAL PREDICTIONS

To fully determine (or to try to understand) how elementary particles interact, we need to compute a process like the following, by means of the Factorization Formula (probabilities everywhere).

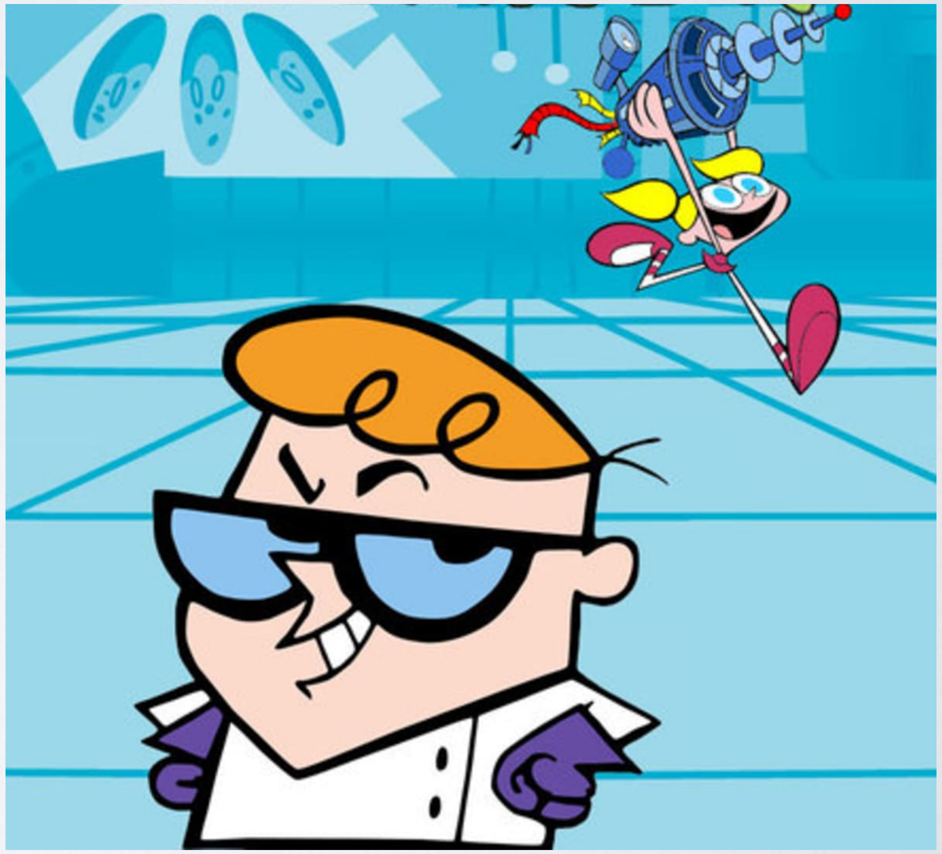


$$d\sigma^{h_1 h_2 \to HX} = \sum_{a,b,c} \int_0^1 dx \int_0^1 dy \int_0^1 dz\, f_a^{h_1}(x, \mu_I) f_b^{h_2}(y, \mu_I) D_c^H(z, \mu_F) d\hat{\sigma}_{ab \to cX}$$

HOWEVER IN THE LAB…

Certainly, it is possible to compute this complicated object by some complicated numerical methods, new algorithms, etc., but it is enough the computation of it to claim that we understand the physics of fundamental particles ?

The issue is that experimental data can only provide us informations of the particles that interact with the detectors.



In general, we only know completely the information of the inicial and final state particles.

What happens with fundamental partons ?

MACHINE LEARNING

Machine Learning is a form of Artificial Intelligence that allows a system to learn from data instead of learning through explicit programming.

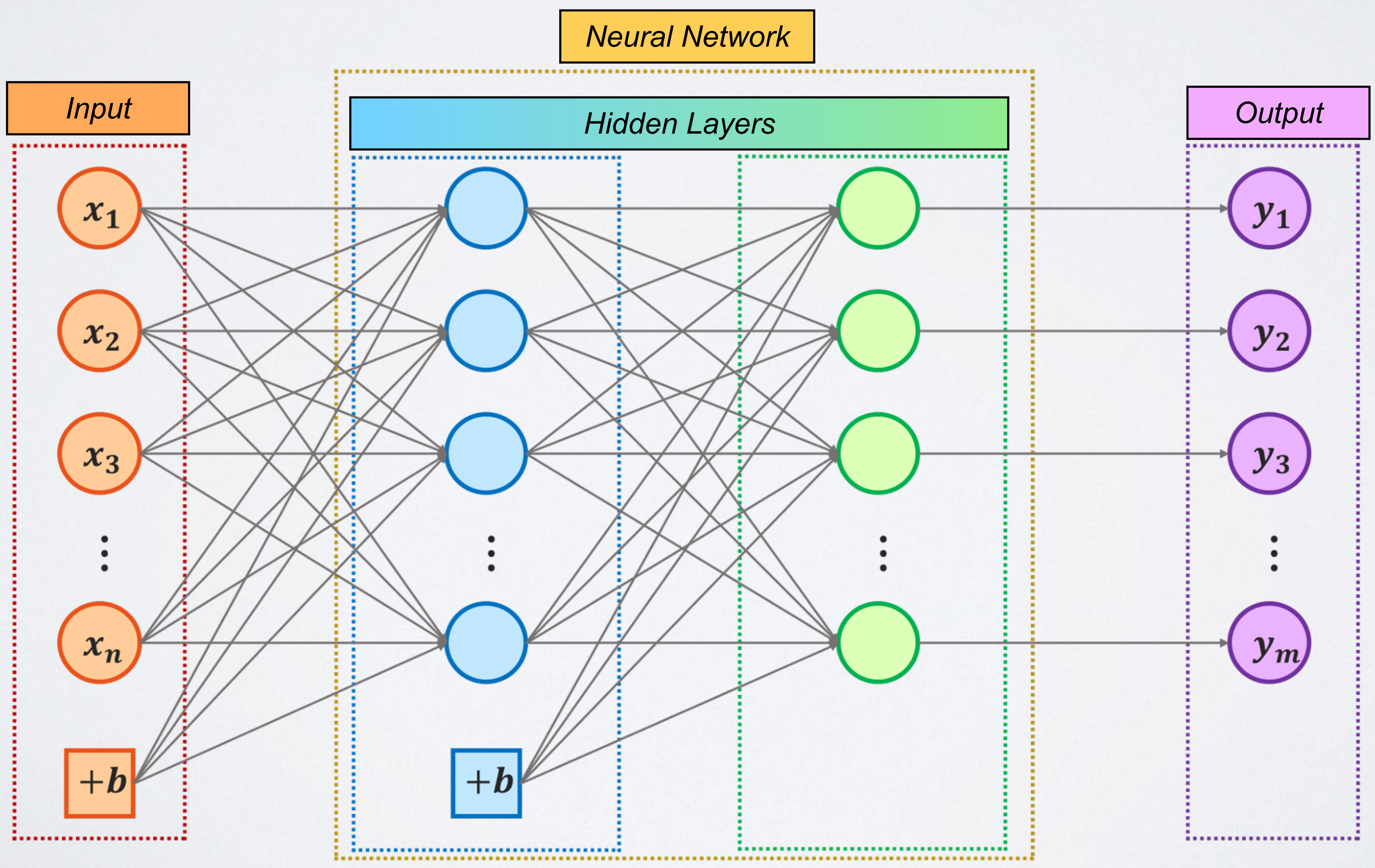
One of these systems that are widely used are neural networks. Can we use this new paradigm to know the properties of fundamental partons (at certain level of accuracy) ?

Input
Experimental data

Output
Theoretical variables

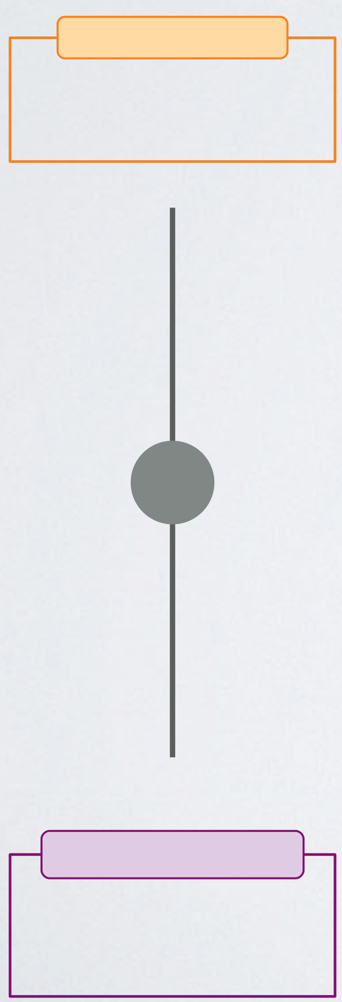
NEURAL NETWORKS

In practice, neural networks are organized by layers and neurons. All connected in order to learn to recognize patterns or structures that allow us to find the output faster.



ANALYZING *DATA* IN HEP USING ML

- What would happen if we left you with a *known data* set and asked predictive methods to learn and predict the unknown *theoretical variables*?



Linear Regression

$$Y_{REC} = \sum_{i=k}^{i=0} \alpha_i x_i \text{ for } x_i \in X_j$$

Gaussian Regression

$$Y_{REC} = \prod_i \exp\left(-||x-\mu_i||^2/2l^2\right)$$

STUDY OF $pp \to \pi^+\gamma$ @ NLO QCD + LO QED

- A first attempt was made to study the production of a pion in association with a photon.
- Since the kinematics of a $2 \to 2$ process at LO is fully determined, we are interested in quantum corrections: NLO QCD + LO QED.
- It is important to point out that at very high energies pions may be the product of a hadronization process, which is poorly understood at the fundamental level but which can be characterized by a kinematic variable, z.

EXTRACTING *DATA*

- The *experimental data* were simulated by Monte Carlo simulation of different channels.
- There are many integrals but we can reduce them by applying the experimental isolation algorithms.

$$
\begin{aligned}
d\hat{\sigma}^{\rm ISO}_{a_1\,a_2\to a_3\,\gamma} \quad &= \quad \frac{\alpha_s}{2\pi}\frac{\alpha}{2\pi}\int d{\rm PS}^{2\to 2}\,\frac{|\mathcal{M}^{(0)}|^2(x_1K_1,x_2K_2,K_3/z,K_4)}{2\hat{s}}\,\mathcal{S}_2 \\
&+ \quad \frac{\alpha_s^2}{4\pi^2}\frac{\alpha}{2\pi}\int d{\rm PS}^{2\to 2}\,\frac{|\mathcal{M}^{(1)}|^2(x_1K_1,x_2K_2,K_3/z,K_4)}{2\hat{s}}\,\mathcal{S}_2 \\
&+ \quad \frac{\alpha_s^2}{4\pi^2}\frac{\alpha}{2\pi}\sum_{a_5}\int d{\rm PS}^{2\to 3}\,\frac{|\mathcal{M}^{(0)}|^2(x_1K_1,x_2K_2,K_3/z,K_4,k_5)}{2\hat{s}}\,\mathcal{S}_3
\end{aligned}
$$

$$
d\hat{\sigma}^{\rm ISO,QED}_{a_1\,a_2\to a_3\,\gamma} = \frac{\alpha^2}{4\pi^2}\int d{\rm PS}^{2\to 2}\,\frac{|\mathcal{M}^{(0)}_{QED}|^2(x_1K_1,x_2K_2,K_3/z,K_4)}{2\hat{s}}\,\mathcal{S}_2.
$$

Generated channels

LO QCD	LO QED	NLO QCD
$q\bar{q}\to\gamma g$	$q\gamma\to\gamma q$	$q\bar{q}\to\gamma gg$
$qg\to\gamma q$	~~$q\bar{q}\to\gamma\gamma$~~	$qg\to\gamma gq$
		$gg\to\gamma q\bar{q}$
		$q\bar{q}\to\gamma Q\bar{Q}$
		$qQ\to\gamma qQ$

NON PERTURBATIVE QCD

Non perturbative objects (PDF and FF) enters in the Factorization Formula and they are known from fitting data with theoretical inputs.

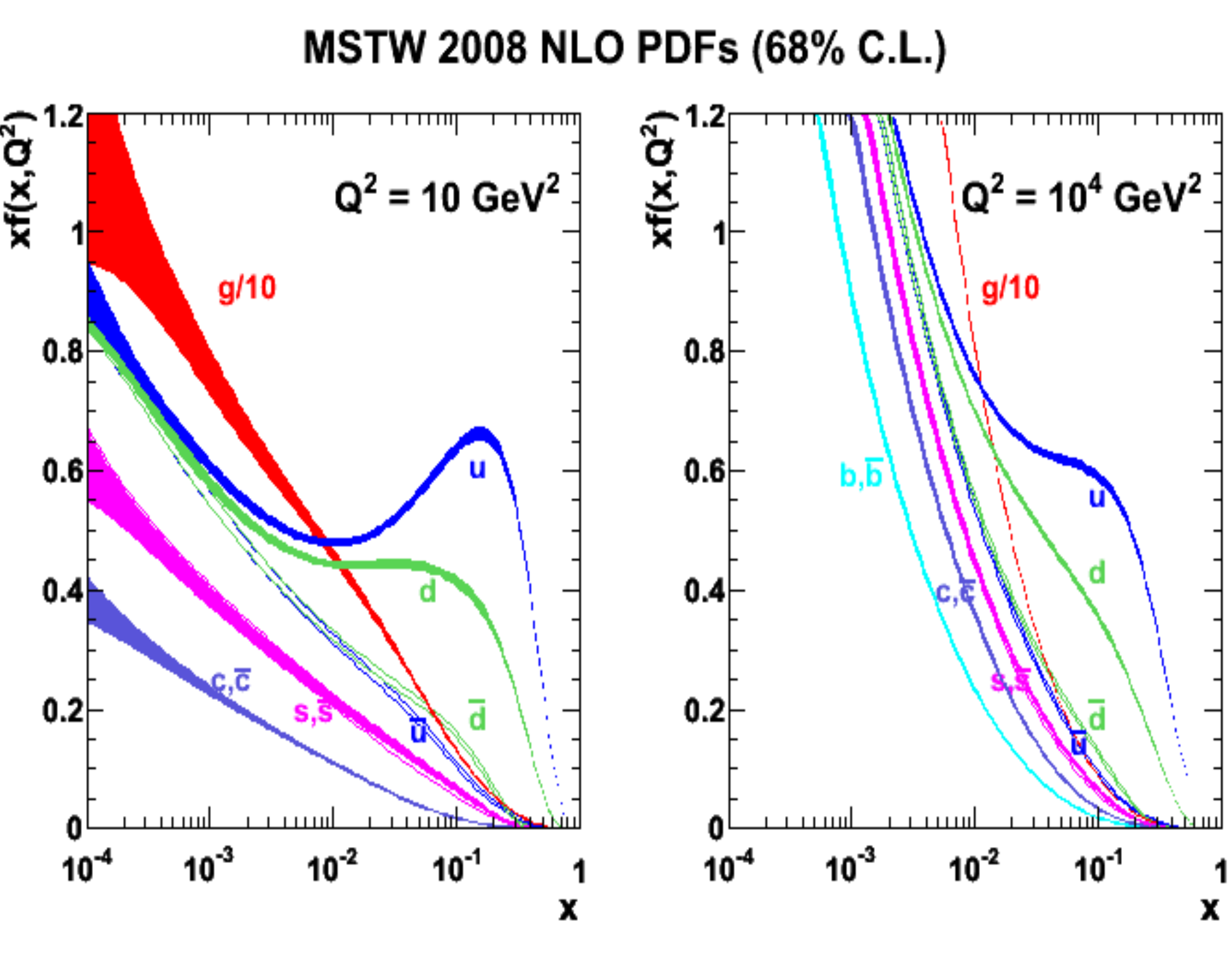
The parton distribution function (PDF) is the probability density of finding a parton a with momentum fraction x inside an h hadron, $f_a^h(x)$.

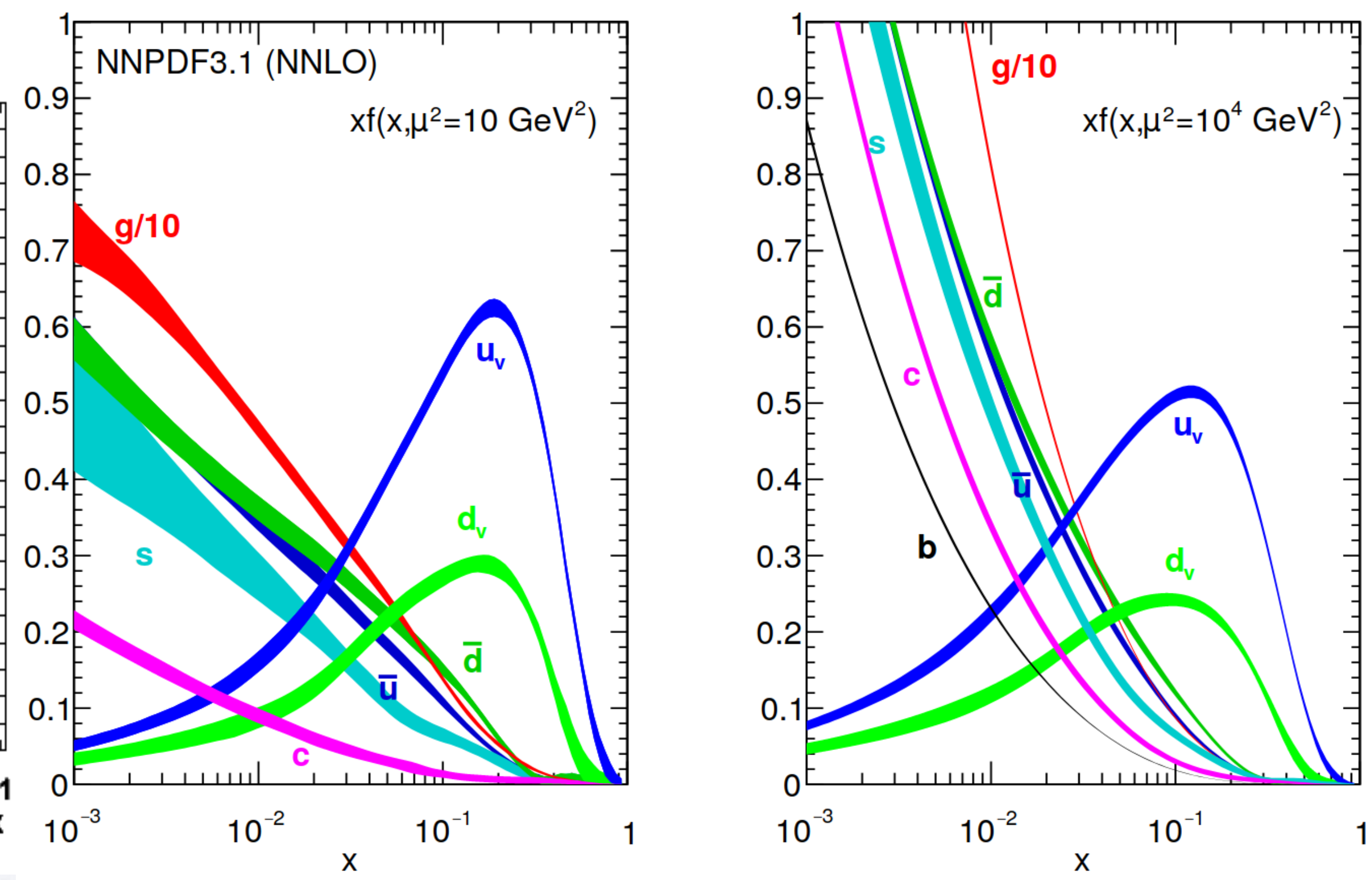
$$\sum_a \int dx\, x\, f_a^h(x) + \int dx\, x\, f_g^h(x) = 1$$

The fragmentation function (FF) is the probability density of generating a hadron h with momentum fraction z of the b-parton, $D_b^h(z)$.

$$\sum_h \int dx\, x\, D_b^h(z) = 1$$

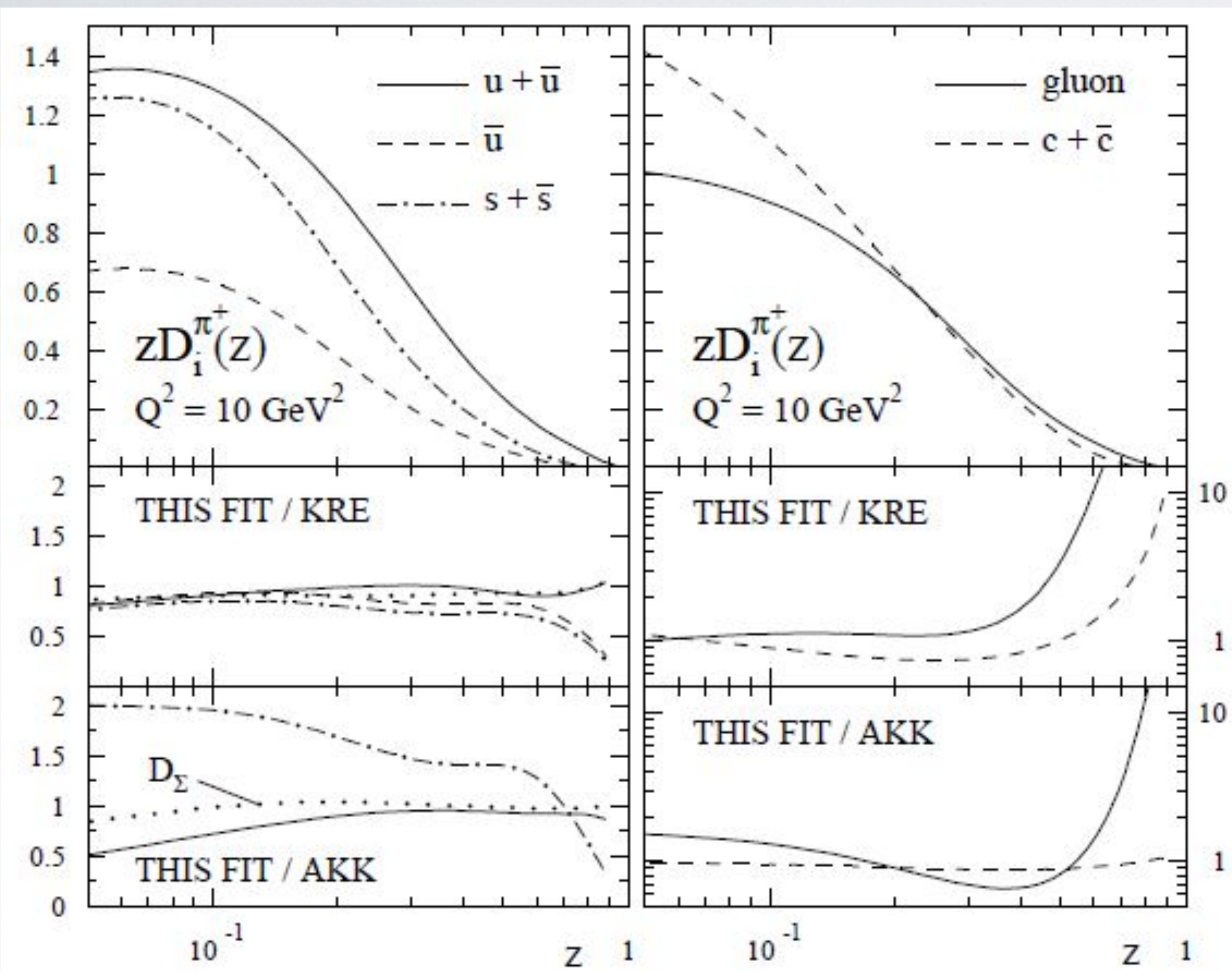
PARTON DISTRIBUTION FUNCTIONS

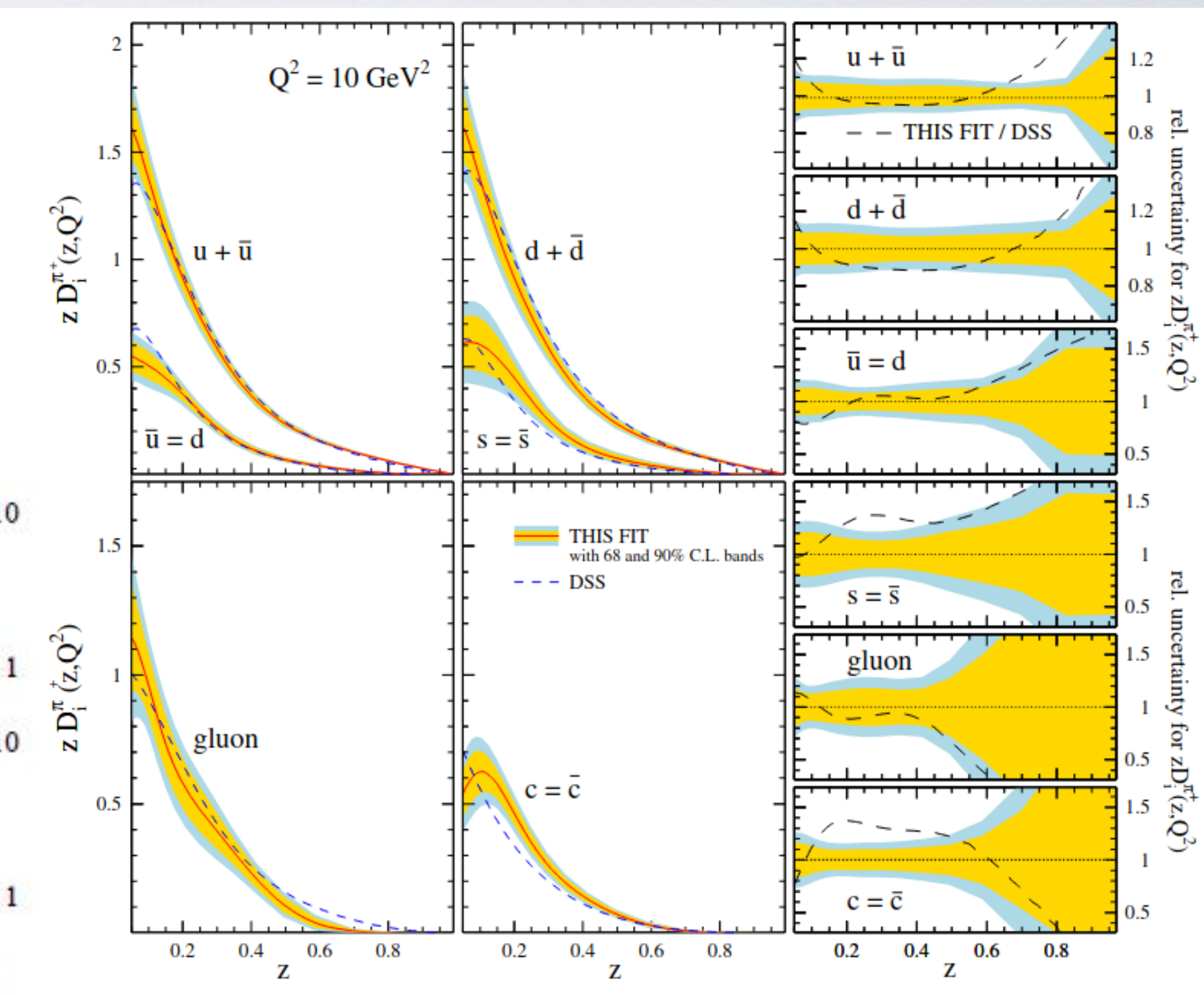




We are interested in studying the impact of the new set of PDFs and FFs. For this reason, we compare MSTW2008 [arXiv:0901.0002] with NNPDF3.1 [arXiv:2009.00014].

FRAGMENTATION FUNCTIONS

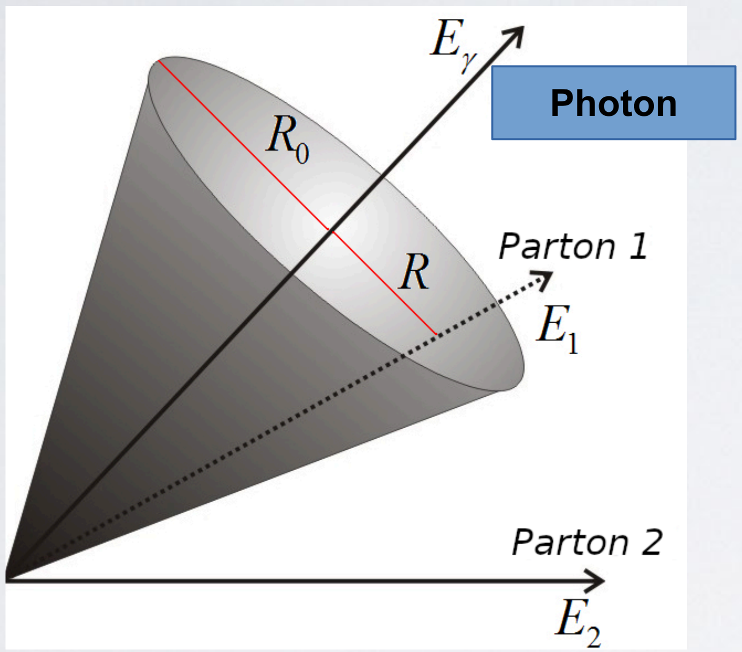




For FF, we compared fragmentation DSS-2007 [arXiv:0707.1506] with DSS-2014 [arXiv:1410.6027].

ISOLATION CRITERIA

- The event selection procedure is given by the Smooth Cone Isolation algorithm.



$$R = r_j = \sqrt{(\eta_j - \eta_0)^2 + (\phi_j - \phi_0)^2}$$

Smooth cone isolation

Smooth function: $\xi(r) = \epsilon_\gamma E_T^\gamma \left(\frac{1-\cos(r)}{1-\cos r_0} \right)^4$

Selection criteria

Define

$$\mathrm{E}_T(r) = \sum_j E_{T_j} \Theta(r - r_j)$$

If, $\mathrm{E}_T(r) < \xi(r)$

γ is isolated

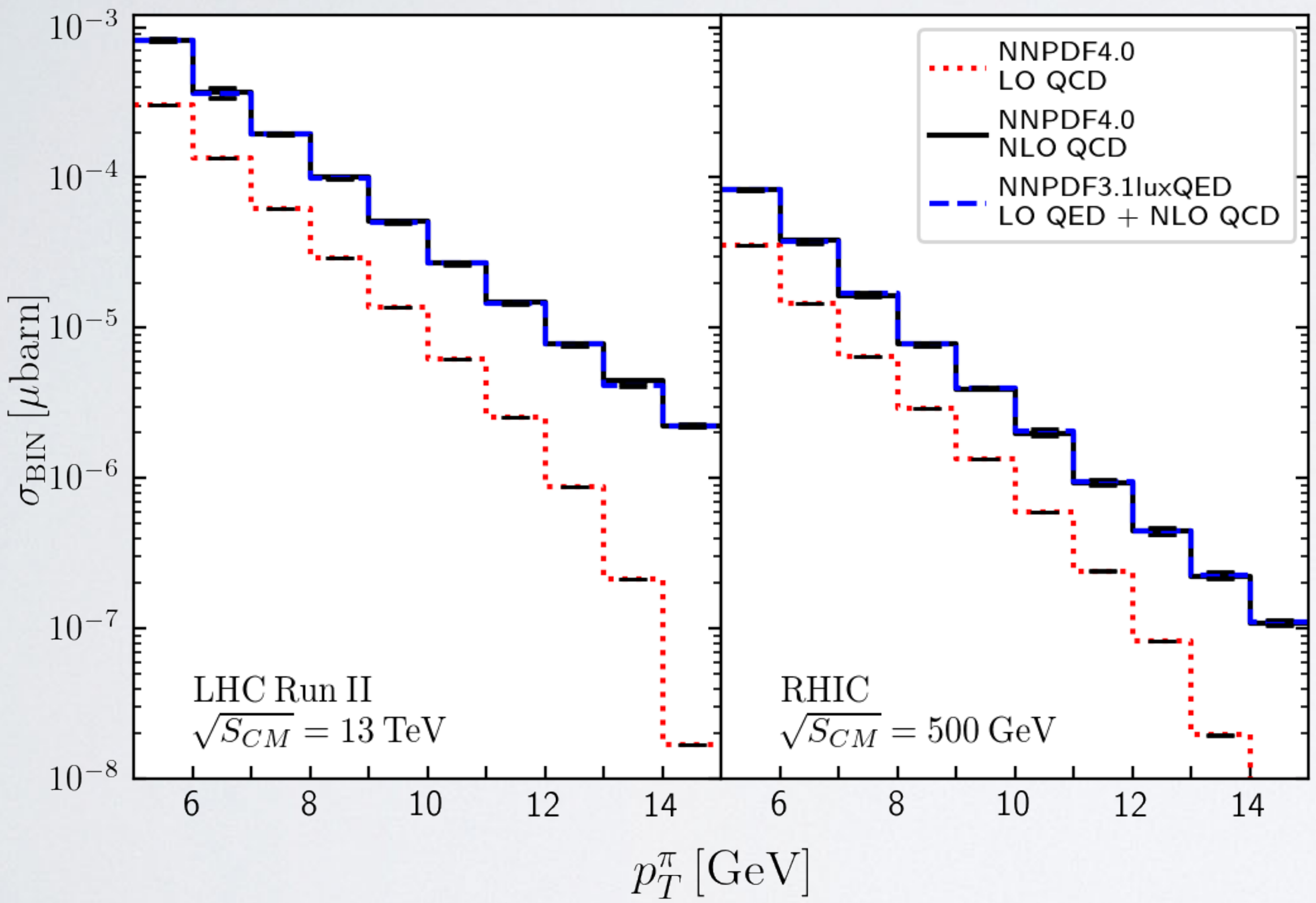
Else, γ is not isolated

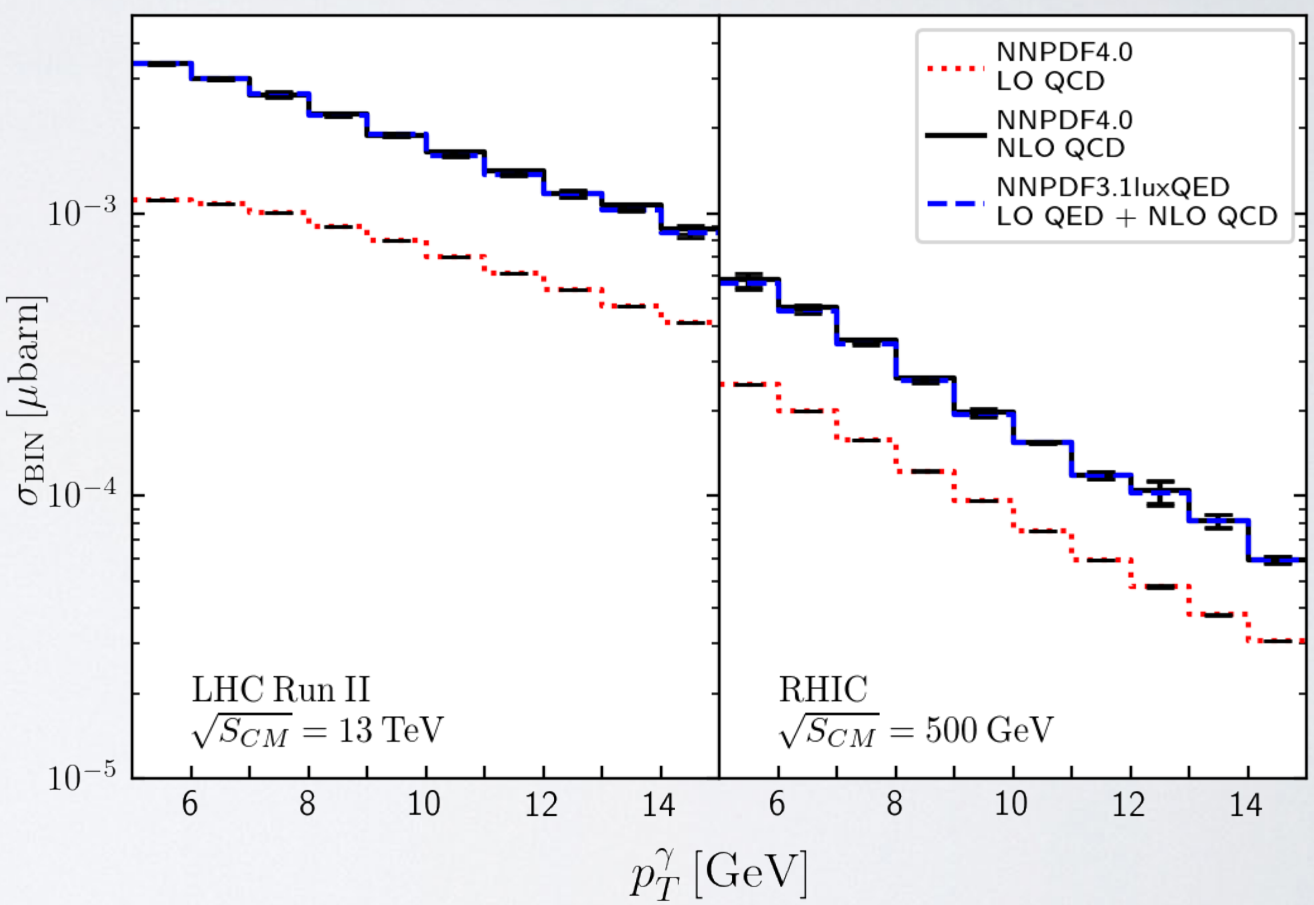
We apply to the MC simulation kinematical cuts from STAR/PHENIX @ RHIC

$$|\eta^h| \leq 0.35\,, \quad |\eta^\gamma| \leq 0.35\,, \quad p_T^h \geq 2\,\mathrm{GeV}\,, \quad 5\,\mathrm{GeV} \leq p_T^\gamma \leq 15\,\mathrm{GeV}$$

HADRON AND PHOTON DISTRIBUTIONS

Differential distributions of transverse momentum of pion and photon at LHC and RHIC energies.



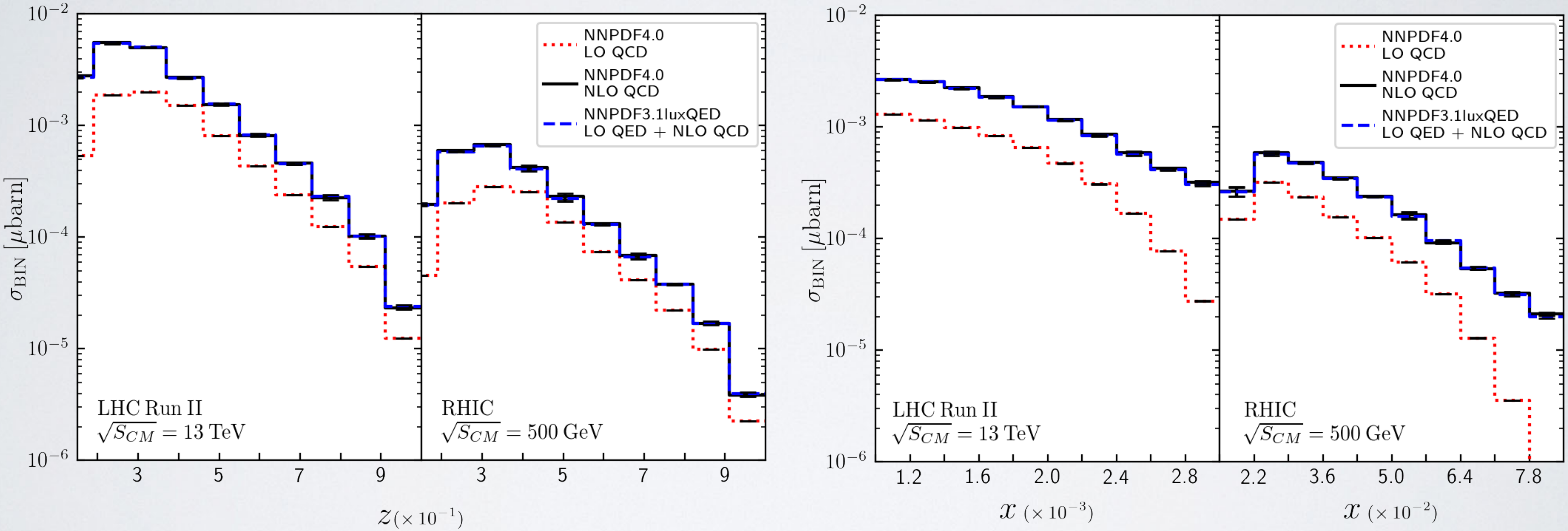


Transverse momentum of the pion increase faster than the photon transverse momentum due to FFs.

Cross sections is larger for LHC energies.

HADRON AND PHOTON DISTRIBUTIONS

Differential distributions of initial and final momentum fractions at LHC and RHIC energies.



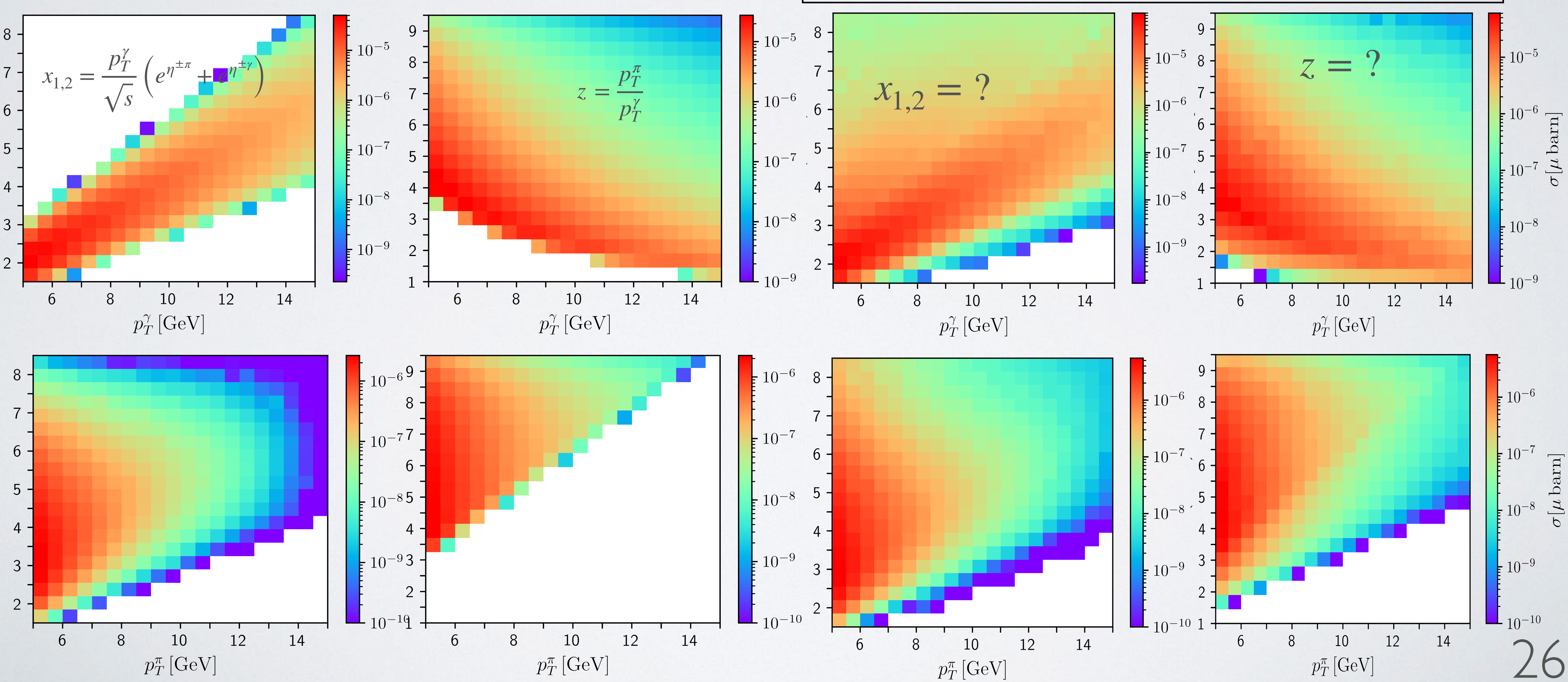
There are important corrections at NLO QCD and small deviations at LO QED.
Experimental cuts in p_T^γ, induce constraints on the maximum value of *x*.
The distribution averages around $z_{peak} \approx 0.35$ for RHIC and $z_{peak} \approx 0.25$ for LHC.

CORRELATIONS OF VARIABLES

In order to consider *experimental* environment, we average the our *experimental data* sets, as $\bar{\mathcal{V}}_{\rm Exp} = \{\bar{p}_T^\gamma, \bar{p}_T^\pi, \bar{\eta}^\gamma, \bar{\eta}^\pi, \overline{\cos}(\phi^\pi - \phi^\gamma)\}$

Momentum fractions at LO

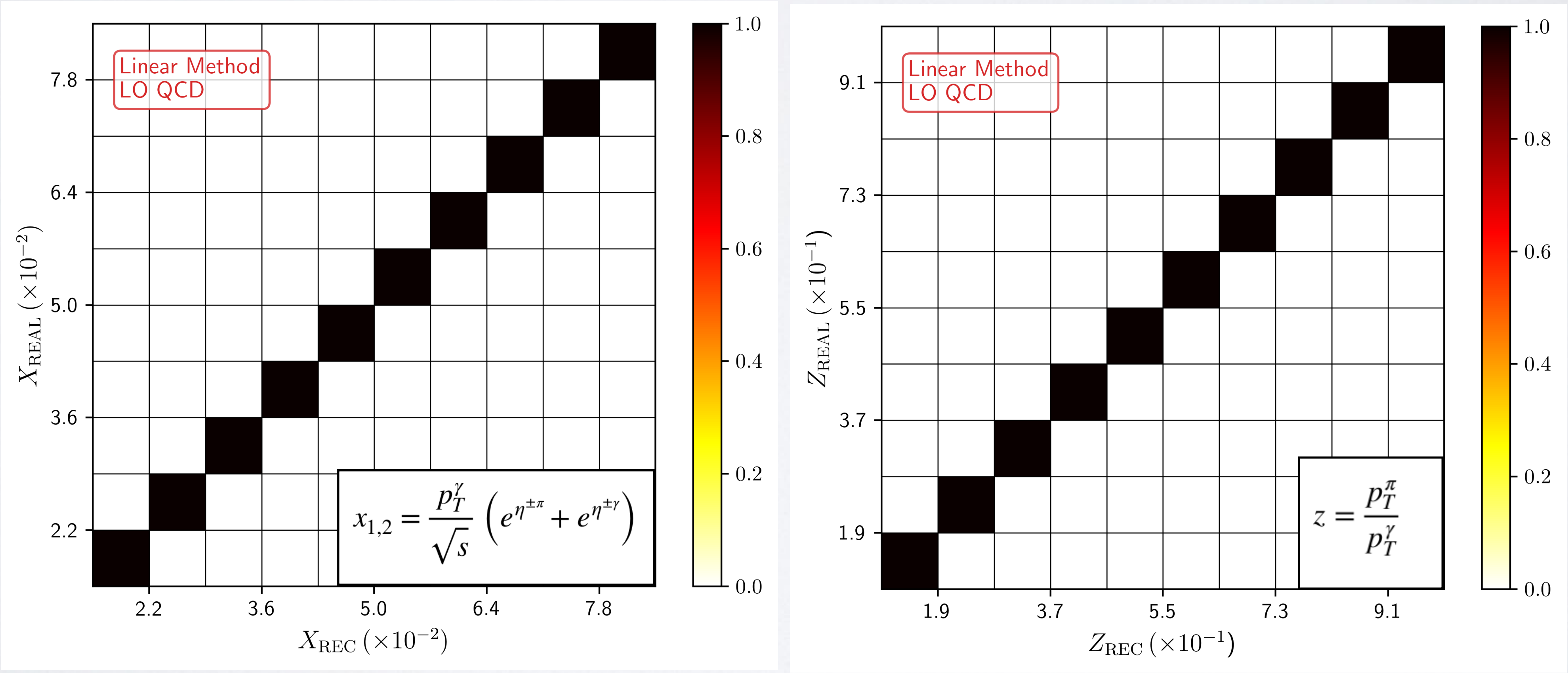
Momentum fractions at NLO QCD + LO QED ?



PARTON KINEMATICS AT LO

Since the kinematics at LO is fully known, we used simple Linear Regression to test our approach.
The basis is inspired by the analytical results,

$$\mathrm{B_{LO}} = \{ \frac{p_T^\gamma}{\sqrt{S_{CM}}} \exp(\eta^\pi), \ \frac{p_T^\gamma}{\sqrt{S_{CM}}} \exp(\eta^\gamma), \ \frac{p_T^\gamma}{\sqrt{S_{CM}}} \exp(-\eta^\pi), \ \frac{p_T^\gamma}{\sqrt{S_{CM}}} \exp(-\eta^\gamma), \ p_T^\pi / p_T^\gamma \}$$



There is no news, the agreement is perfect.

PARTON KINEMATICS AT NLO

For NLO kinematics, for Linear and Gaussian Regression, the LO basis is not enough.

General basis:

$$\mathcal{K} = \{\frac{p_T^\gamma}{\sqrt{S_{CM}}}, \frac{p_T^\pi}{\sqrt{S_{CM}}}, \exp(\eta^\gamma), \exp(\eta^\pi), \cos(\phi^\pi - \phi^\gamma), (\frac{p_T^\gamma}{\sqrt{S_{CM}}})^{-1}, (\frac{p_T^\pi}{\sqrt{S_{CM}}})^{-1}, (\exp(\eta^\gamma))^{-1}, (\exp(\eta^\pi))^{-1}\}$$

$$Y_{\text{REC}} = \sum_{i=1, i\neq 5}^{9} (a_i^Y + b_i^Y \mathcal{K}_5)\,\mathcal{K}_i + \sum_{i\leq j, \{i,j\}\neq 5, j-i\neq 5} (c_{ij}^Y + d_{ij}^Y \mathcal{K}_5)\,\mathcal{K}_i \mathcal{K}_j$$

LO-inspired basis:

$$\mathcal{B}_{\text{NLO}}^{X_1} = \{\frac{p_T^\gamma}{\sqrt{S_{CM}}}\exp(\eta^\gamma), \frac{p_T^\gamma}{\sqrt{S_{CM}}}\exp(\eta^\pi), \frac{p_T^\pi}{\sqrt{S_{CM}}}\exp(\eta^\gamma), \frac{p_T^\pi}{\sqrt{S_{CM}}}\exp(\eta^\pi), \frac{p_T^\gamma \mathcal{K}_5}{\sqrt{S_{CM}}}\exp(\eta^\gamma), \frac{p_T^\gamma \mathcal{K}_5}{\sqrt{S_{CM}}}\exp(\eta^\pi), \frac{p_T^\pi \mathcal{K}_5}{\sqrt{S_{CM}}}\exp(\eta^\gamma), \frac{p_T^\pi \mathcal{K}_5}{\sqrt{S_{CM}}}\exp(\eta^\pi)\}$$

$$\mathcal{B}_{\text{NLO}}^{Z} = \{p_T^\pi/p_T^\gamma, \mathcal{K}_5\, p_T^\pi/p_T^\gamma, \mathcal{K}_5\, p_T^\pi/\sqrt{S_{CM}}, \mathcal{K}_5 \sqrt{S_{CM}}/p_T^\gamma\}$$

Physically-motivated basis:

$$b_6^{X_1} = 0$$

$$c_{6,j}^{X_1} = d_{6,j}^{X_1} = c_{i,6}^{X_1} = d_{i,6}^{X_1} = 0, \quad \{i,j\} \in \{1,\ldots,9\}$$

$$b_1^Z = b_7^Z = 0$$

$$c_{1,j}^Z = d_{1,j}^Z = 0, \quad j \in \{1,\ldots,9\}, \quad j \neq \{5,7\}$$

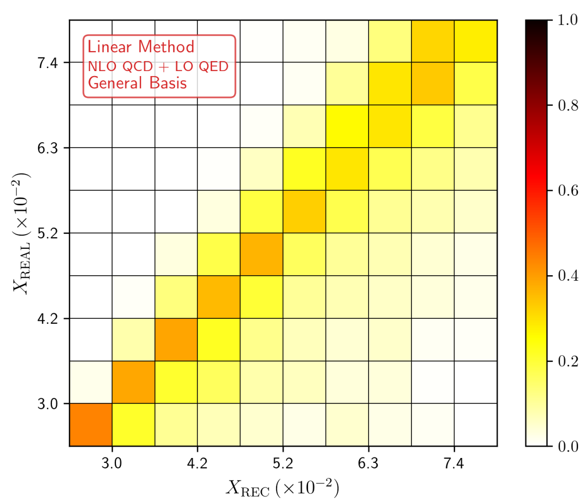
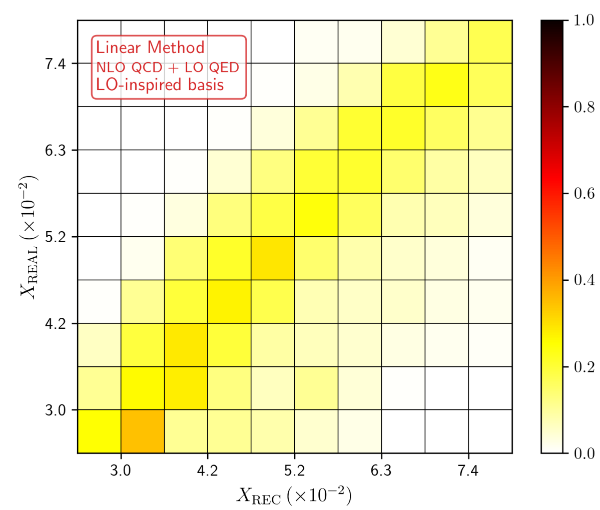
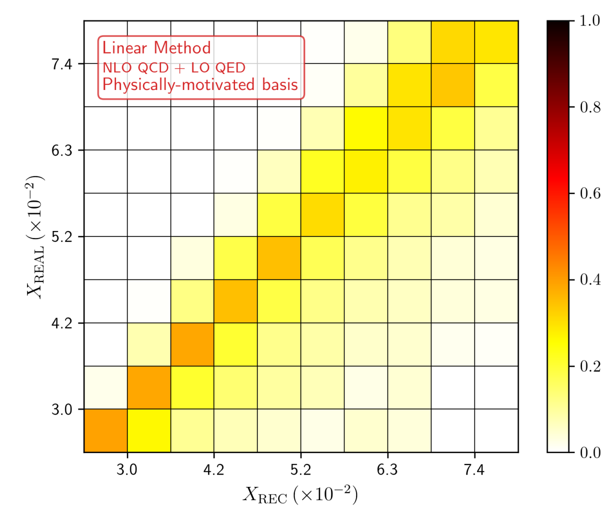
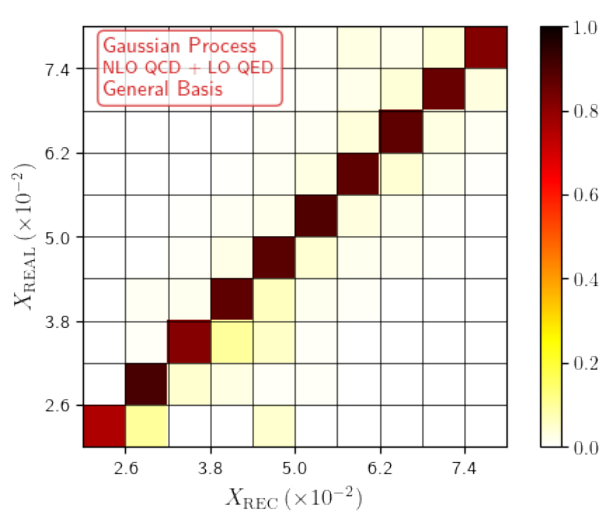
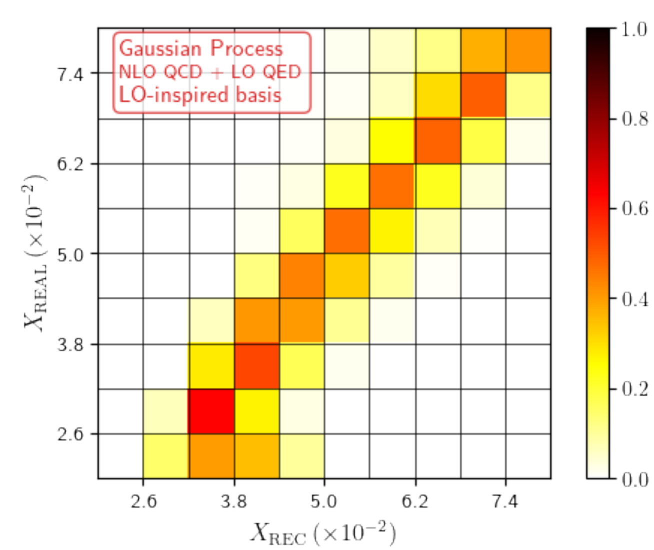
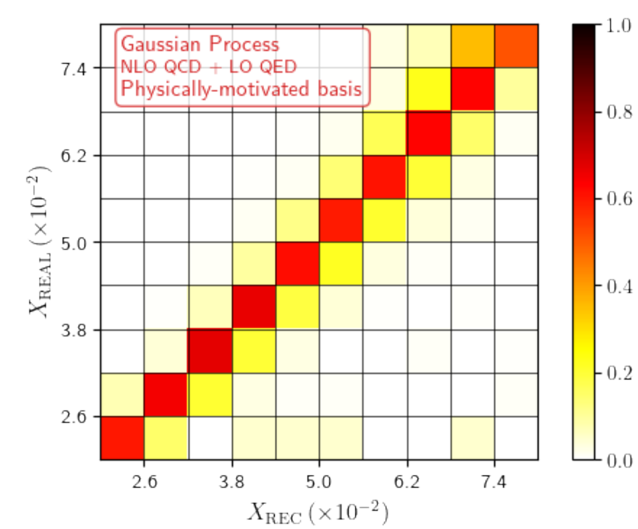
$$c_{i,7}^Z = d_{i,7}^Z = 0, \quad j \in \{1,\ldots,9\}, \quad j \neq \{1,5\}$$

$$x_{1,2} = \frac{p_T^\gamma}{\sqrt{s}}\left(e^{\eta^{\pm\pi}} + e^{\eta^{\pm\gamma}}\right)$$

$$z = \frac{p_T^\pi}{p_T^\gamma}$$

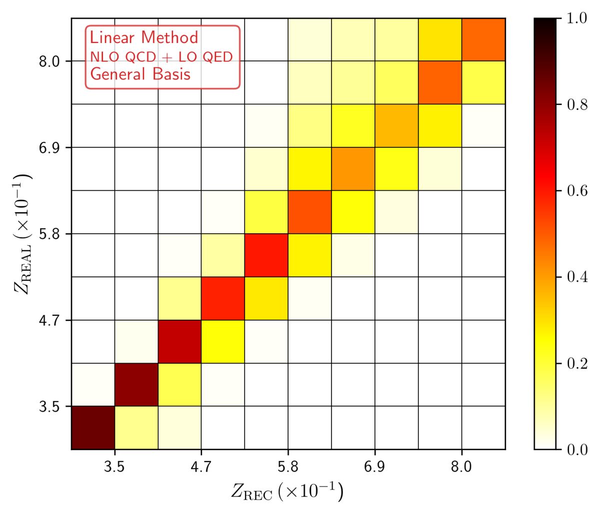
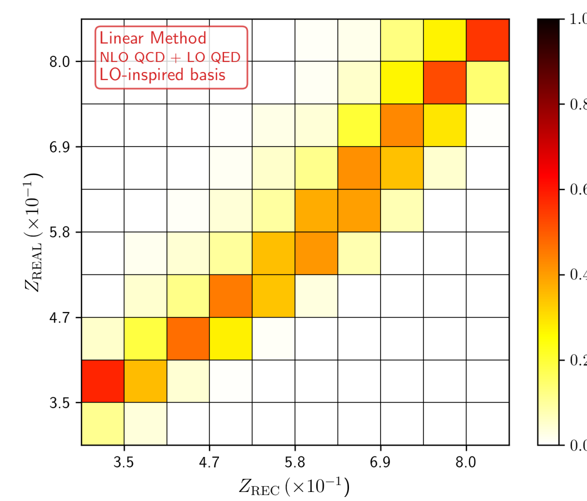
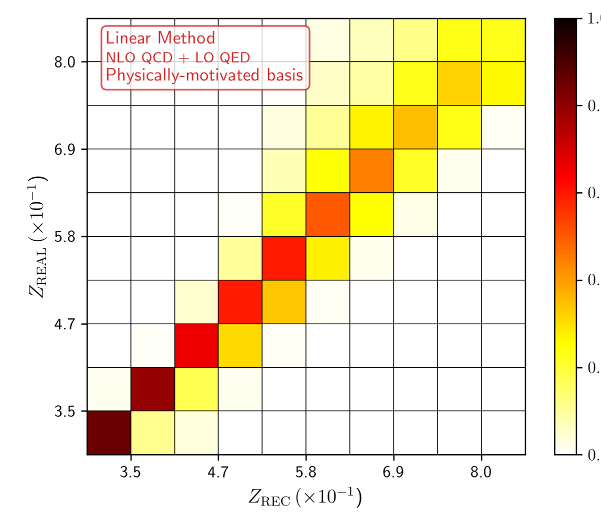
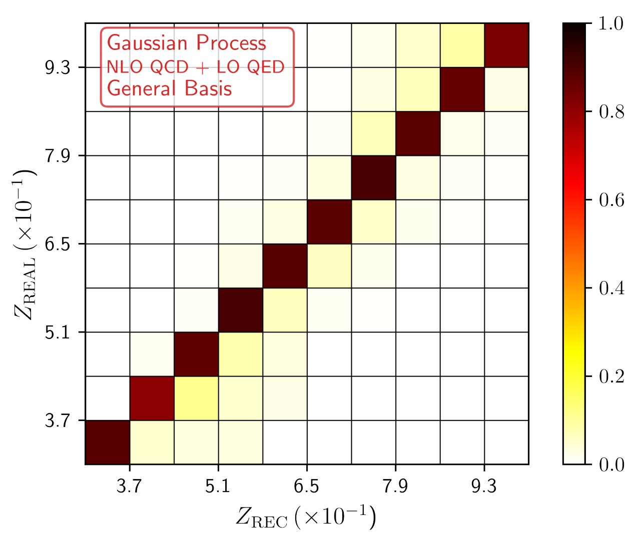
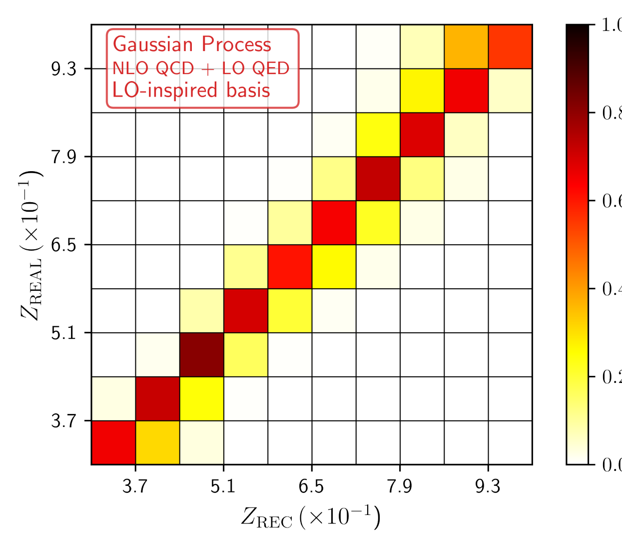
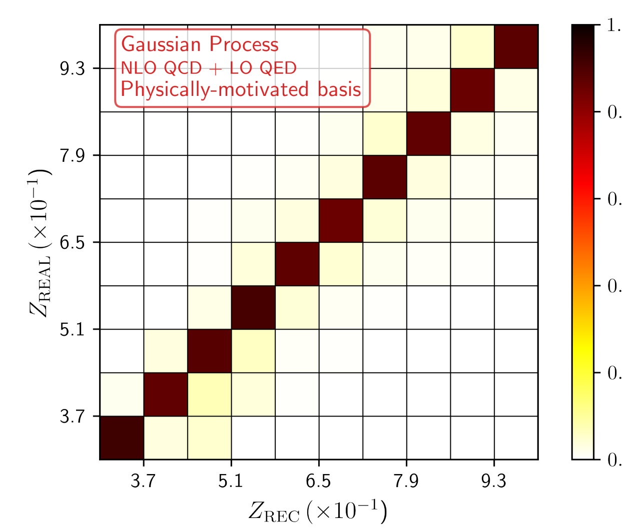
X RECONSTRUCTION AT NLO

There are different bases that can be used. The choice can be inspired by physics or not, it depends on the programmer. [Rentería-Estrada, H.P., Sborlini, Zurita].

<table>
<tr><th></th><th>Base I</th><th>Base II</th><th>Base III</th></tr>
<tr><td>Linear Regression</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Gaussian Regression</td><td>
</td><td>
</td><td>
</td></tr>
</table>

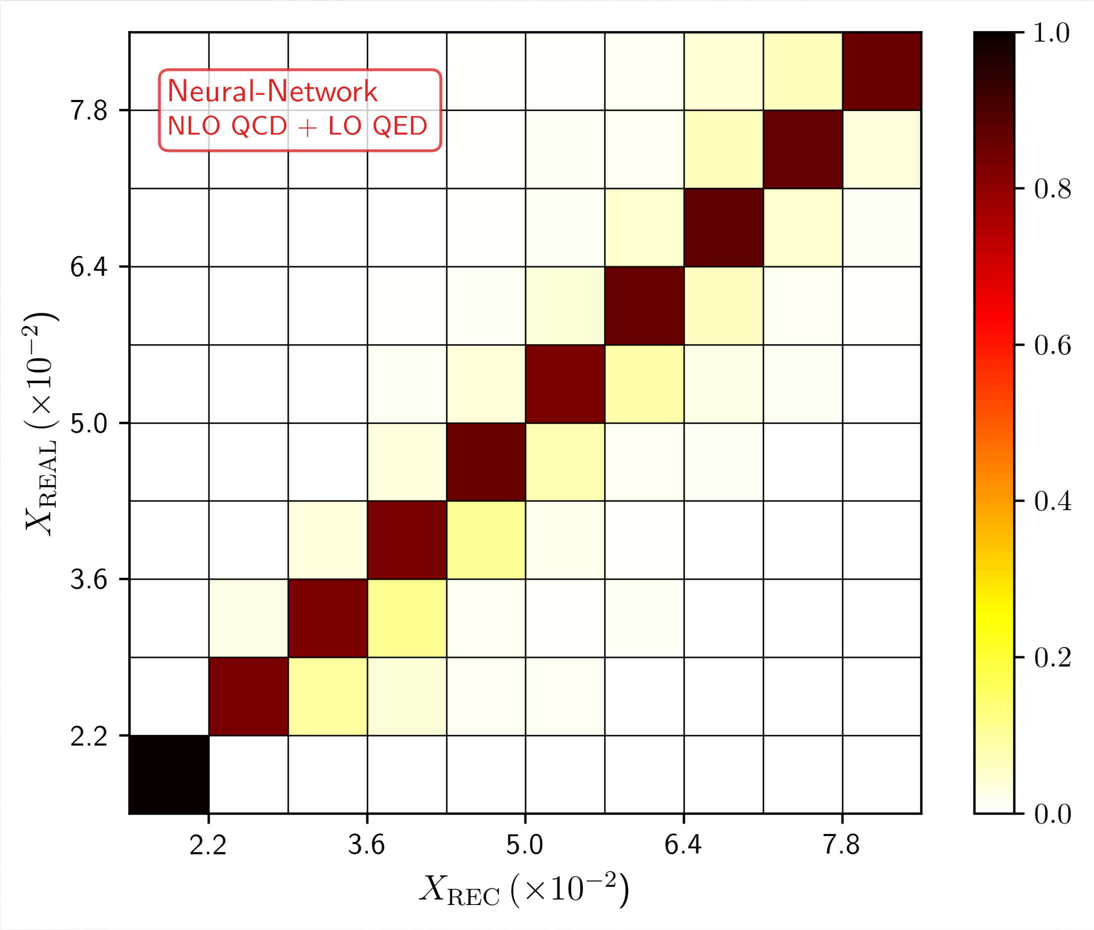
Z RECONSTRUCTION AT NLO

For the momentum fraction of final state partons we find a good agreement in the reconstruction

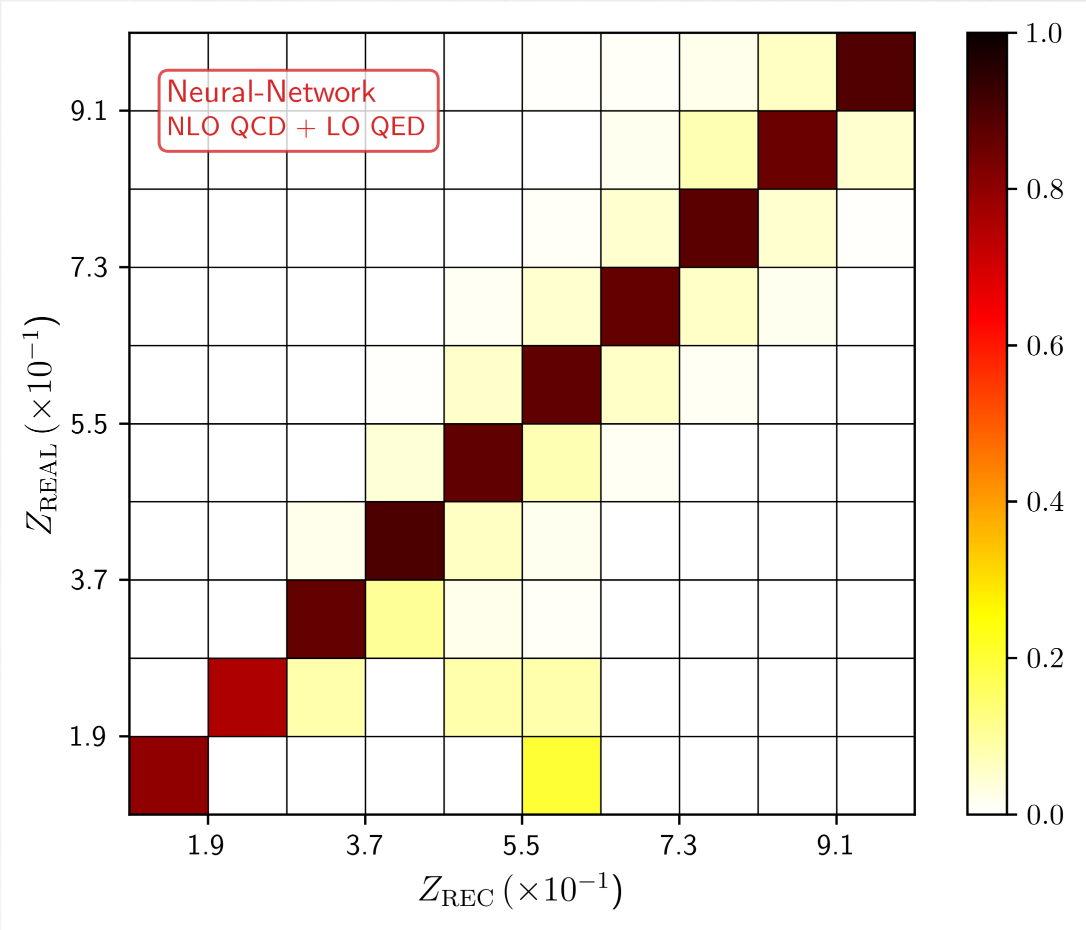
<table>
<tr><th></th><th>Base I</th><th>Base II</th><th>Base III</th></tr>
<tr><td>Linear Regression</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>Gaussian Regression</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
</table>

USING A NEURAL NETWORK

In the case of NN, there is no need to specify the basis to reconstruct momentum fractions of initial and final states partons.



Prediction of *x*
Initial state



Prediction of *z*
Final state

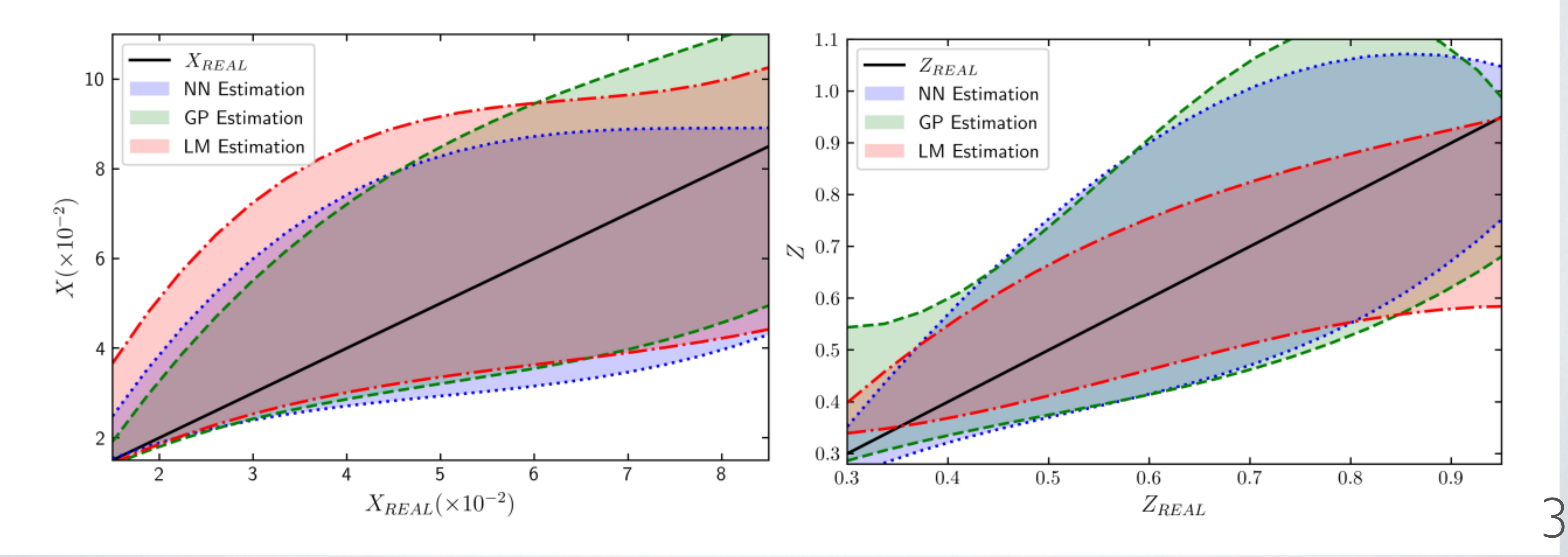
UNCERTAINTIES

The procedure presented leads to three functions for reconstructing each momentum fraction. Given a kinematic point in the grid, $p_j \in \bar{\mathscr{V}}_{\text{Exp}}$, we have

$$X(p_j) \equiv \left\{ X_{\text{REC}}^{\xi=2}(p_j), X_{\text{REC}}^{\xi=1}(p_j), X_{\text{REC}}^{\xi=1/2}(p_j) \right\} ,$$

and define

$$X_{\text{REC}}(p_j) = \overline{X(p_j)} \pm \frac{\max(X(p_j)) - \min(X(p_j))}{2} \equiv \overline{X(p_j)} \pm \Delta X(p_j) .$$



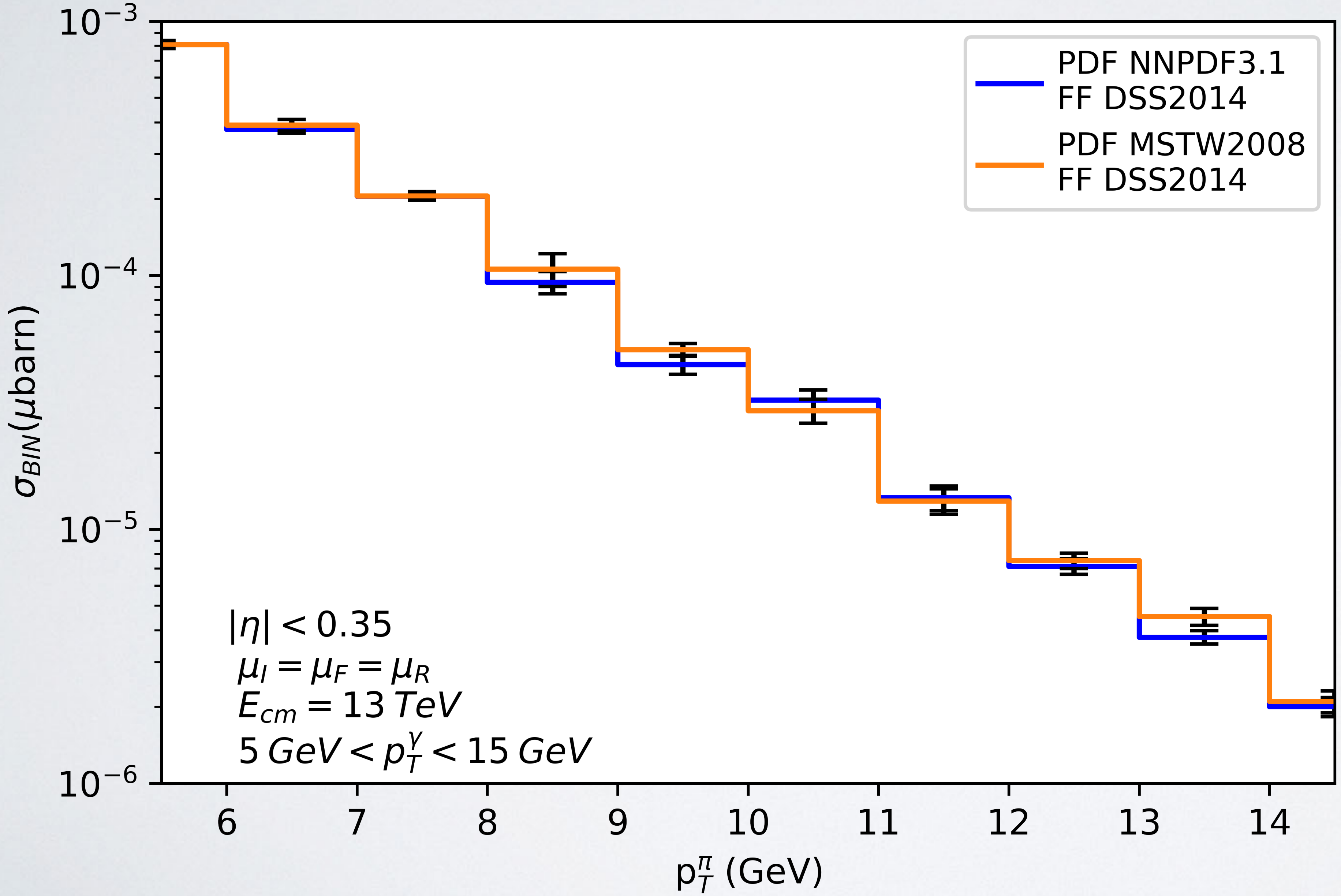
CONCLUSIONS

- Precision phenomenology is needed for the forthcoming precision era.
- Theoretical calculations needs new paradigms to avoid the handling of ill-defined quantities (infinities).
- Machine Learning is now guiding some studies to achieve the understanding of fundamental properties of nature.
- Monte Carlo simulations takes long time to compute large number of integrals due to infinities but also from the running of non-perturbative kernels, can we do something there with ML ? Stay tuned…

GRACIAS

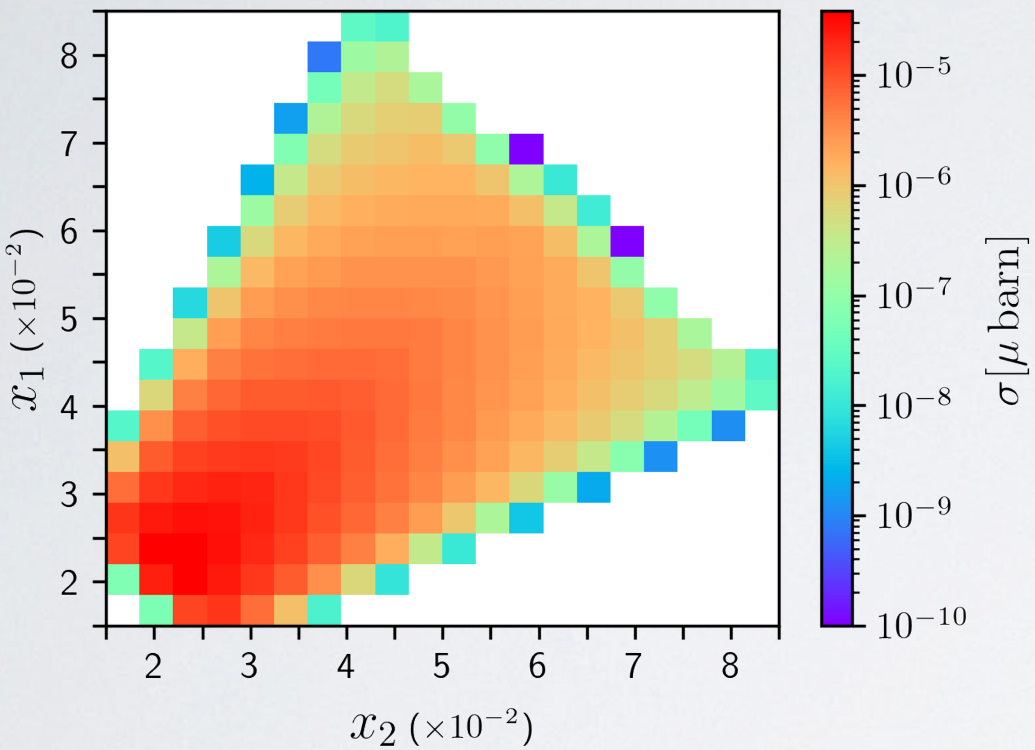
BACKUP SLIDES

EFFECTS OF REFITTING PDFS



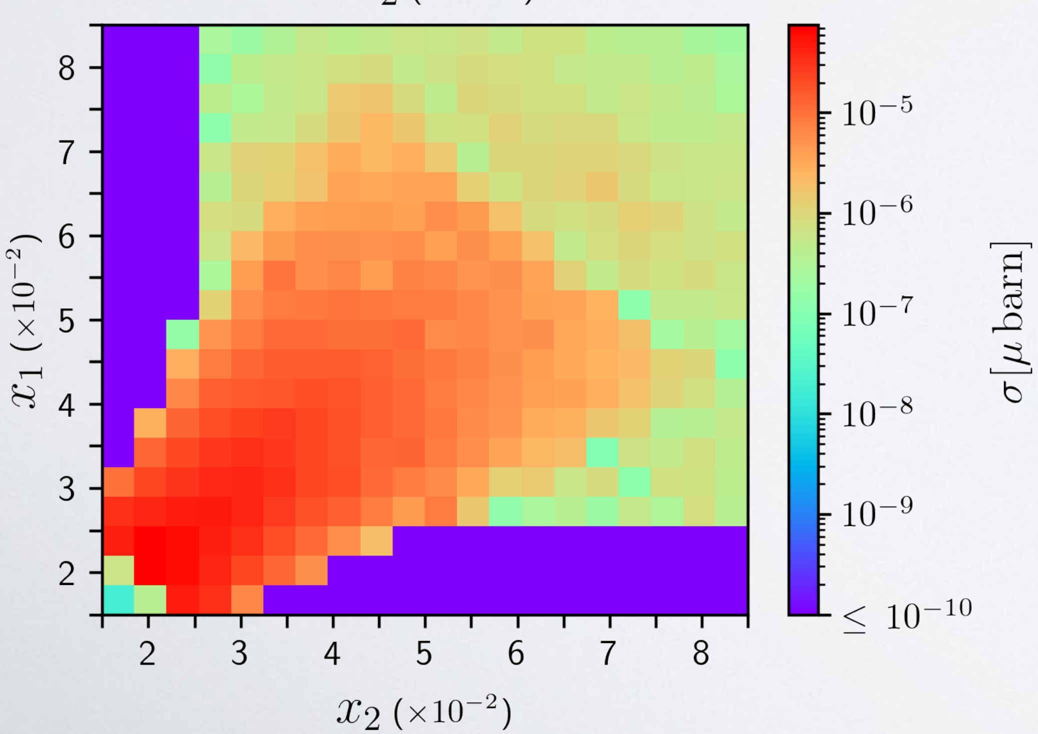
Considering the same FF (DSS2014), we find small deviations of NNPDF3.1 w.r.t. MSTW2008

PHOTON + HADRON CORRELATIONS



LO QCD

- Positive correlation
- Consequence of initial state symmetry
- Used as a cross check



NLO QCD + LO QED

NN PARAMETERS

- We try different architectures to find the best result.

Parameters	TEST 1	TEST 2	TEST 3
# hidden layers	2	4	3
# neurons/layer	50	100	100
tolerance	10^{-2}	10^{-2}	10^{-3}
max. number of iterations	10^8	10^8	10^9
# iterations w/o change	14,000	21,000	100,000

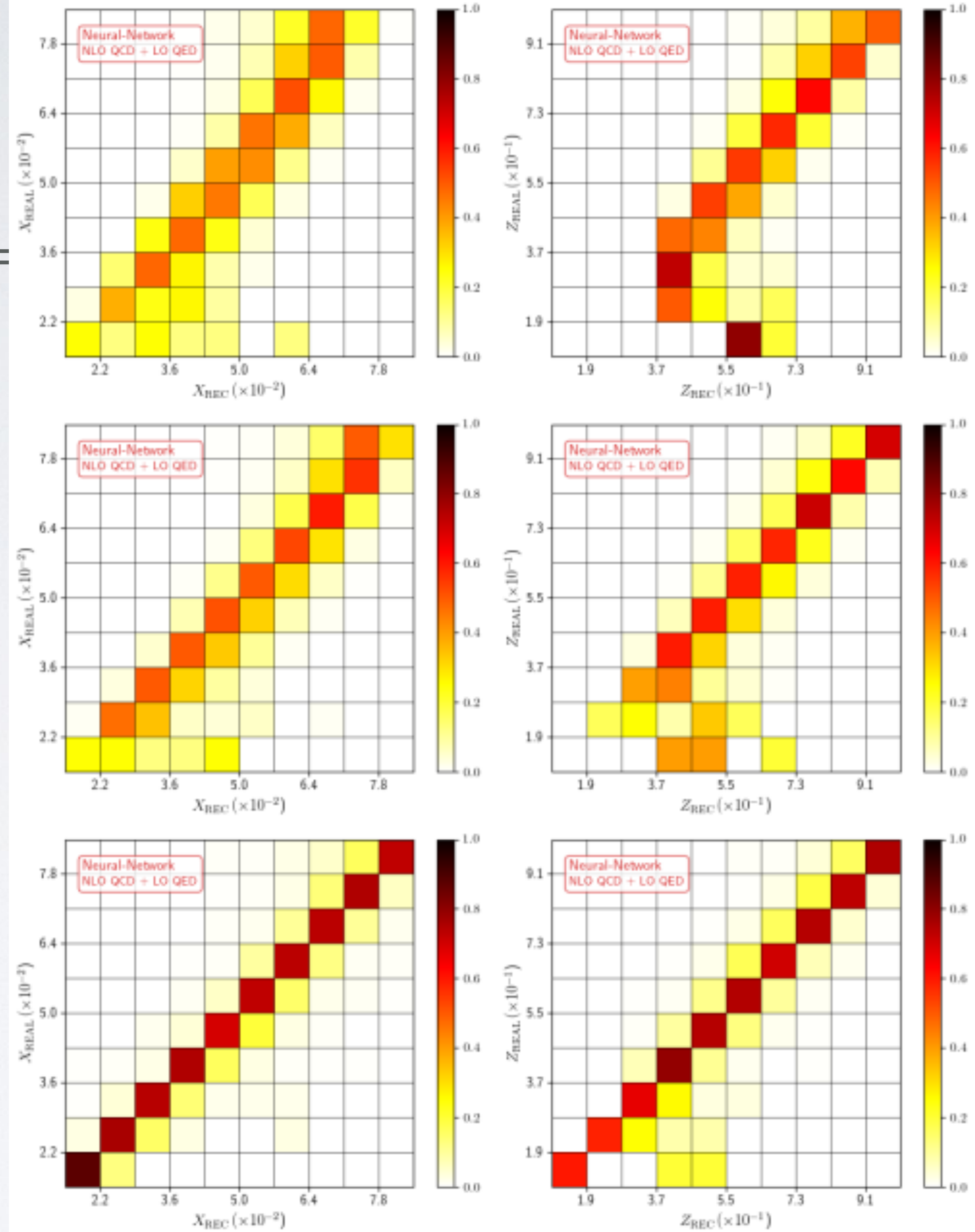


Figure 19: Comparison of the momentum fractions X_{REAL} vs. X_{REC} (left) and Z_{REAL} vs. Z_{REC} (right) obtained with MLP at NLO QCD + LO QED accuracy. The parameters for TEST1 (upper row), TEST2 (middle row) and TEST3 (lower row) are given in Table 3.

NN PARAMETERS

- Different configurations for each reconstructed parameter were taken.

	X_{REC} (LO)	Z_{REC} (LO)	X_{REC} (NLO)	Z_{REC} (NLO)
# of hidden layers	2	1	5	5
# of neurons/layer	200	100	300	300
activation function	ReLU	ReLU	ReLU	ReLU
# iterations	1×10^5	1×10^5	1×10^{12}	1×10^{12}
learning rate	1×10^{-3}	1×10^{-3}	1×10^{-4}	1×10^{-4}

Table 2: Architecture for the MLP best fit parameters for the reconstruction of the momentum fractions at LO in QCD: $X_{\rm REC}$(LO) and $Z_{\rm REC}$(LO) (second and third columns), and for the momentum fractions at NLO QCD + LO QED: $X_{\rm REC}$(NLO) and $Z_{\rm REC}$(NLO) (fourth and fifth columns).

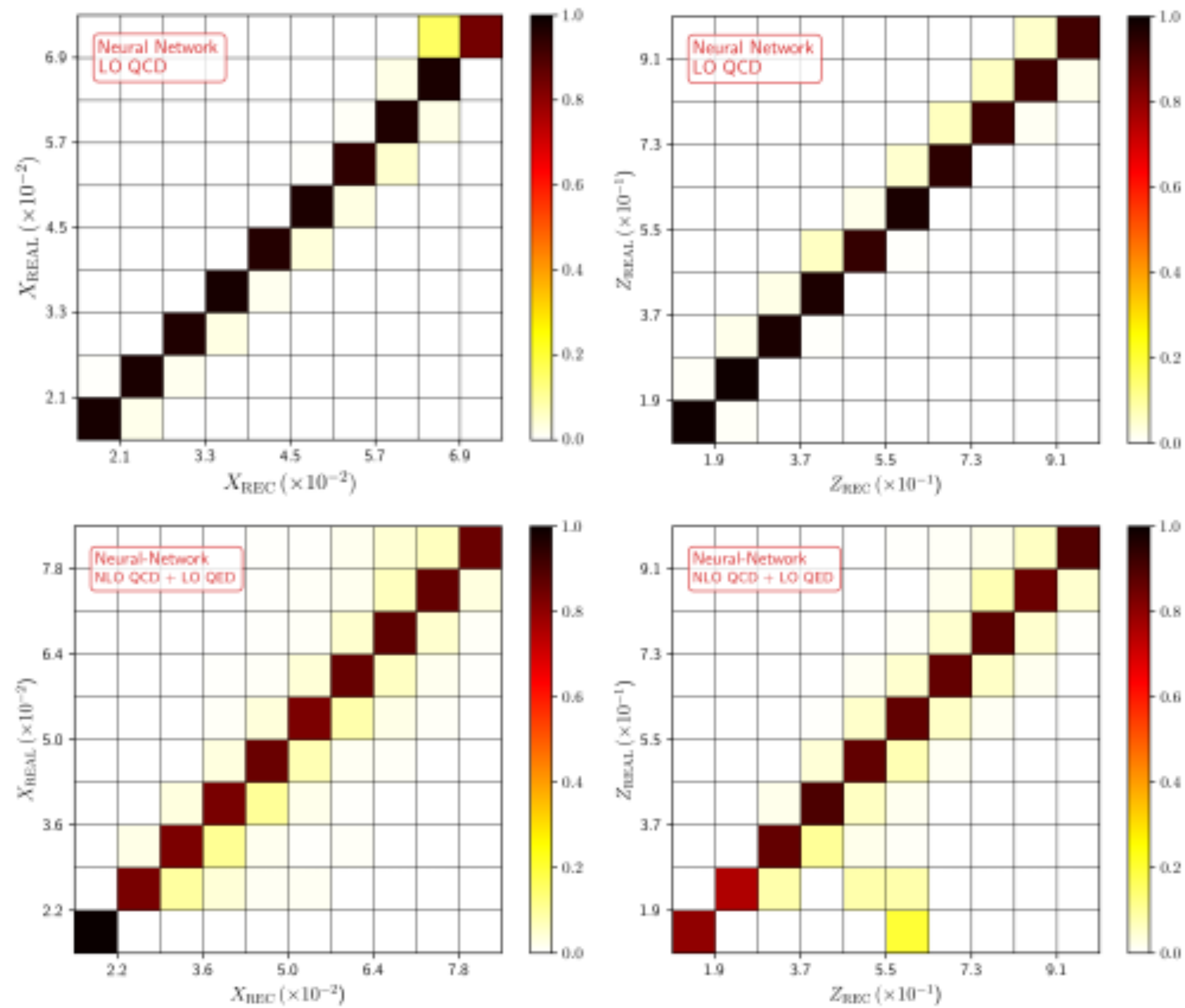


Figure 15: Left: Comparison of the momentum fractions $X_{\rm REAL}$ and $X_{\rm REC}$ obtained with MLP neural networks with the parameters given in Table 2. The upper (lower) row corresponds to the LO QCD (NLO QCD + LO QED) data set. Right: same as the l.h.s but for $Z_{\rm REAL}$ and $Z_{\rm REC}$.